ICE-CAMERA: a flatbed scanner to study inland Antarctic polar precipitation.

Massimo Del Guasta
Istituto Nazionale Ottica CNR, Sesto Fiorentino, 50019, Firenze, Italy
Correspondence to: Massimo Del Guasta (massimo.delguasta@ino.cnr.it)

Abstract.
The study of precipitation at very high latitudes is challenging because of the harsh environmental conditions that limit the external activity of humans and instruments, particularly during the polar winter. I describe a device (ICE-CAMERA) for automatic imaging, measuring and classifying ice precipitation on the high plateau. From 2014 the instrument operates unattended, all year round, at Concordia Station (75°S,123°E, 3220 m altitude). The instrument provides detailed precipitation information every hour. The article describes the apparatus and its treatment software.

Introduction.
There is virtually no reliable observation of precipitation on the Antarctic ice sheet to validate climate models for this variable. This is particularly true in the high plateau, where less than 10 cm of equivalent water accumulates every year. Frost deposition and extremely cold temperatures adversely affect even traditional and simple rain gauges. Any precipitation instrument on the high plateau must also cope with a very harsh winter environment. The lack of ground validation data can explain why the climate models used to predict climate change (such as CMIP and ERA5) differ significantly in the reconstruction of even average precipitation in inland Antarctica.

Relevant work on ice precipitation has been carried out mainly in coastal areas. Grazioli et. al, (2017), as part of a multidisciplinary field campaign, deployed a multi-angle snowflake camera (MASC) to take photographs of individual snowparticles. In coastal Antarctic areas, due to high relative humidity and mild temperatures, snow particles are much bigger in size than inland ones, and particle types are quite similar to midlatitude ones. Blowing snow is also common at coastal sites, due to generally much stronger winds than on land and may be confused with precipitation.
Photographic studies of precipitation in the interior of Antarctica are rather rare, mainly conducted at the South Pole station through formvar replicas. Early works with formvar (Hogan, 1975) identified mm-sized columnar crystals and column- or bullet- rosettes in cloud precipitation, and smaller (0.1 mm sized or smaller) platelike particles in clear-sky precipitation (‘Diamond Dust’, DD). Kikuchi and Hogan (1979) collected formvar replicas of DD in the summer at South Pole station, finding columnar crystals of 90 um average length and plates as small as 50 um in diameter. Ohtake and Yogi (1979) classified winter ice crystal precipitation in Antarctica under six categories. These included large rosettes, bullets and columns (1 mm length or bigger), thin hexagonal plates and columns (200 um or less), and smaller crystals of various shapes including triangular and polyhedral. Shimizu (1963) observed "long prism" crystals in the winter at Byrd Station (80S, 120W). Size
distributions of Antarctic DD in winter and spring were reported by Smiley et al. (1980) for particles larger than 50 um; they observed the same ice crystal forms that were reported earlier. Walden et al. (2003) studied DD, blowing snow, and cloud precipitation in winter, at South Pole, by collecting crystals on slides and analyzing them using microphotography. In their study, columns with an average length of 60um and plates with an average diameter of 30um were found in DD. Lawson et al. (2006) worked at South Pole in summer, using innovative Cloud Particle Imagers (CPIs) which replaced formvar replicas. This technique allowed the automatic analysis of around 700,000 DD crystal images in terms of caliper size, aspect ratio and other shape parameters. An automatic classification software based on the shape parameters was used to categorize the images into nine simplified classes: small plates and spheroids, columns, thick plates, plates, budding rosettes, rosettes, complex with side planes, irregulars.

The Concordia International Station, located on the Dome-C (DC, 75°S,123°E, 3220 m above sea level) is a special location to test new instruments for precipitation studies. The temperature rarely exceeds -30°C in summer, winter temperatures reach -85°C and winds are usually below 10 m s⁻¹. In these conditions, precipitation of ice crystals can be studied by simply collecting them on horizontal surfaces. Moreover, the cloud ice particles observed in DC are very similar to those observed inside cirrus by aircraft, with the advantage of a low cost and possible continuity of sampling. This is done locally by hand, starting in 2008, collecting the precipitation of flat surfaces and visually inspecting them. This analysis is restricted to one observation per day, a rate that is difficult to increase especially in winter. The analysis of these data is also tedious and often subject to biases due to ice re-processing, hoar formation, and subjective judgement of the shape and relative abundance of ice particles.

Schlosser et al. (2017) relied on this manual observation and the classification of ice particles in his analysis of precipitation isotope data. They simply classified the ice grains into diamond dust, drifting snow, snow and frost. The prevalence of hoar in the observed daily "precipitation" below -50°C indicates the limitations of this manual technique if information on DC precipitation particles is desired.

Detailed work was carried out in DC on a few individual DD and cloud precipitation crystal replicas by means of SEM electron microscopy by Santachiara et al. (2016). They also analysed very small particles (10-50 um), almost inaccessible to ordinary optical methods.

The operation in DC of research aircraft equipped with optical particle imaging devices is prohibited during the cold Antarctic winter. As such, a continuous monitoring of precipitation at DC can only occur near the ground. The purpose of developing ICE-CAMERA was to fill a gap in precipitation monitoring at Concordia with a robust instrument capable of monitoring with continuity, and all-year round, habit and size of ice particles in precipitation, while overcoming the problems associated with the visual inspection of precipitation. The combined construction of rugged equipment for taking photographs of ice particles and the adoption of image processing and machine learning techniques have served this purpose.
1.1 Overview
ICE-CAMERA is a flatbed scanner specially designed for observing polar precipitation in the harsh environmental conditions of Concordia station (Fig.1). In this work, the term 'precipitation' will include both diamond dust and cloud precipitation.

The principle is simple: at low temperatures and low wind speeds (conditions encountered at DC), precipitation falling over a flat glass accumulates with time and remains frozen until it sublimes, or it is brushed away, leaving plenty of time for scanning the glass surface for counting and measuring individual ice particles. After an entire scan, the glass surface is heated and the precipitation sublimated. That cycle is repeated hourly. Each scan image is automatically pre-processed for precipitation, and a mosaic, summary image containing only the segmented particles (if any) is stored for eventual post-processing. Each particle is also automatically classified and measured using both image processing and machine learning. Individual particle data are stored in rows of a text file, along with weather and maintenance data, for post-processing and statistical analysis. The instrument was first installed in Concordia in 2012, but replaced in 2014 with its improved version, described here. From then on, the instrument works year-round to produce precipitation data, every hour.

The core of ICE-CAMERA is a line-scan IP-camera, moved by a scan sledge, and looking up at the deposition window through a 45° mirror (Fig.2). During the scan, the images are sent to the PC, located inside the shelter. The focus of the camera is eventually adjusted by a small focusing sledge moving the 45° mirror.
Fig. 2. ICE-CAMERA basics: the scan sledge moves the image-acquisition line along the deposition surface. The focusing sledge adjusts the focus.

All basic operations (with the exception of CAM scanning) are driven by a custom microprocessor (Microchip PIC18) logic board (Fig. 3). The same board reads the housekeeping temperature sensors attached on the deposition surface and placed inside and outside the instrument. The PIC Board communicates to the main computer (located inside the heated shelter) through the RS232. A NI Labview software controls the acquisition of images, reads maintenance data and manages all ICE-CAMERA operations along the RS232 line. After each full image acquisition, a MATLAB code is invoked to process the image.
1.2 The Camera

A linear scanning GigE Vision monochrome camera (Schaffer-Kirchoff SK7500VTF-XB (7500 pixels, 7x7mm pixels), equipped with a 1:1 macro lens (APO-Rodagon D1X, f5.6) is used for the acquisition. A 90° bending aluminium mirror is used to look upward. The illumination is ensured by 850 nm LEDs. A color filter (Schott RG715, 800-1000 nm band-pass) was used at the CAM lens, in order to have a fully solar-blind instrument. The line-scan camera assembly is moved by a motorized sledge at a speed of 8 mm s⁻¹ in order to scan a rectangular surface of 55x200 mm, here named Deposition Surface (DS), at the center of the window. The final image is 7500*30000 pixel, 12 bit, monochrome.
1.3 The Focusing

In working conditions, the focal depth is ±0.5 mm. A preliminary and accurate alignment of the motorized sledge plane to the deposition surface ensures uniformity of focus across the DS at room temperature. A motorized focusing sledge, moving the bending mirror, allows to adjust the focus in operating conditions (Fig. 2). As ICE-CAMERA works outdoor in DC, it can experience a broad internal temperature range, from +5°C in summer to -45°C in winter, with quite large temperature gradients across the structure. Thermal expansion and changes in optical refractive indexes result in unpredictable changes in the focal plane. The correction of the focus is thus automatically performed, every 6 hours, by bringing the measuring sledge outside the DS, where a focusing spot (a sandpaper strip) is glued to the window (Fig. 4).

The porous structure of the sandpaper has a length-scale of the order of 0.1 mm, comparable with the size of the measured ice particles. While calibrating, the focus sledge is moved by ±2 mm around the actual position in ±0.25 mm steps. Successive images of the sandpaper are taken and their contrast (defined here as the standard deviation of the intensity of the
pixels) is measured. After a Gaussian-fit of the contrast as a function of defocusing (Fig.5), the position corresponding to the maximum contrast is obtained, and the mirror sledge is moved in that position.

Fig. 5. Typical focus calibration. Image contrast against defocusing. The contrast is computed in three parts of the image: the center and the left and right wings. The full-image contrast graph is also shown (red). The slight difference between the center and the wings is an ordinary lens effect.

1.4 Illumination.

Lighting is supplied by two 850 nm IR LED (TSHG6200) linear arrays. Both arrays illuminate the scan line symmetrically and approximately 45° from the optic axis to minimize multiple reflections in the double windows and within the camera. Infrared illumination was chosen in order to work in solar-blind conditions. This is particularly important, as the linear scanning camera should always look at the sky at the zenith. The uniformity of lighting along the linear CCD image was tested by taking an image of the same sandpaper used in the focus. The intensity-profile along the CCD image was measured, and the intensity of each LED eventually changed. The final intensity uniformity over the entire frame resulted ±15%.

1.5 The Deposition Surface

-The DS is an electrically-heated glass (E-GLAS, Saint-Gobain). This glass is transparent at 850 nm, and is electrically heated with 45 V ac, 1000 Wm⁻² when sublimating the ice particles. A second glass sheet (an ordinary flatbed scanner optical glass), placed under the DS, makes up with it a double window, necessary in order to keep the DS thermally insulated from the interior of the instrument. A thermocouple is attached to the DS, while others monitor the temperatures of the window interspace.
DS temperature of 5°C above air temperature is enough to prevent the formation of frost on the smooth deposition window, in any season (Fig.6).

![Image of ICECAMERA at -70°C, Concordia station, winter: the DS if free of frost](https://example.com/image.png)

The temperatures measured by the ICE-CAMERA sensors is transmitted continuously to the PC. The NI-Labview software controls the internal temperature of ICE-CAMERA above -40°C, and the DS temperature always under -5°C (“Heater 1” in Fig.3). An independent, wired thermostatic system provides back-up temperature control in the event of a PC or PIC board failure (“Heater 2” in Fig.3). After a black-out, when the power is restored, a timer is used to heat the inside of the instrument before turning on the electronics. This is important at Concordia to prevent damage to standard electronics with operating
temperatures of -40°C or higher. In winter, a coat of styro foam is added around the instrument, whereas in summer, a mylar sunscreen prevents overheating (Fig.1). During warm weather, the outside air is carried inside the box with a fan, for better cooling.

1.6.1 Sublimation-deposition cycle

After a full scan of the DS, electricity is applied to the window to sublimate the particles. The heating rate of the DS depends on the air temperature and wind velocity. An indoor test (Fig.7), showed a heating rate of 2.5°C min⁻¹, and a cooling rate of 1 °C min⁻¹. The cooling rate is almost 50% of the heating rate just due to the sandwich heating-glass structure, with the heating layer at the middle.

Outdoor tests in DC summer (-30°C air temperature) showed a heating rate of 3°C min⁻¹ in still air, 2.5°C min⁻¹ with 2.5 m s⁻¹, and 1.8°C min⁻¹ with 5 m s⁻¹ wind speed. In each case, the cooling speed was approximately 1.5°C min⁻¹. An outdoor sublimation test (-30°C air temperature, wind speed <3 m s⁻¹) performed with snow manually spread on the DS showed that, after applying heating for 10 minutes (up to a DS temperature of -8°C) the sublimation of the majority of particles
(D<1000 um) is complete within 20 minutes, with just a few big (D>>1000 um) grains still present after 30 minutes. After these tests, the heating time was set at 10 min. Heating is anyway interrupted if the DS temperature exceeds -5°C to avoid melting ice in summer. When taking a scan every hour, the DS is assumed to be sensitive to falling ice particles for a period of approximately 20 minutes before the next scan. This time changes with temperature, wind and summer sun exposure. This uncertainty, combined with the eventual wind and evaporative removal of particles during the accumulation period, precludes the use of ICE-CAMERA for strict quantitative precipitation studies (nonetheless, precipitation events are captured in ICE-CAMERA as peaks in the number and surface of particles detected).

Figure 8 presents a statistics (years 2019-2021) on the temperatures of the deposition window (including the sublimation phase) with respect to the local air temperature. The DS was always found to be warmer than air by typically 5°C. The temperature difference between DS and the air temperature (dT) decreases as the velocity increases. As Figure 8 shows also sublimation temperatures, at the peak of sublimation dT resulted approximately +25°C in still air (comparable with the indoor test of Figure 7), and decreased to +20°C with 8 m s⁻ⁱ of wind.

Fig.8: Years (2019-2021): DS temperature vs air temperature (left). Temperature difference dT in relation to wind speed (right)
1.6.2 Sublimation of ice particles

The DS of ICE-CAMERA is always warmer than the surrounding air. This is good for eliminating possible confusion between deposited precipitation and hoar, allowing the device to be used in all Antarctic conditions. The negative effect is the potential sublimation of the deposited particles, as DS ice is always super-saturated relative to the surrounding air, particularly at the beginning of the deposition period. Figure 9 shows the vapour pressure of ice on the DS relative to the surrounding air (saturated relative to ice, Huang (2018)). The plot is repeated for three values of $dT$: $dT=20^\circ C$ for the sublimation period, $dT=5^\circ C$ for the deposition period in still air, $dT=2^\circ C$ for the deposition with 8 ms$^{-1}$ wind. During ice deposition, the ice vapour pressure on the DS is 1.5-2.2 greater than the vapour saturation of the ambient air. With strong wind ($dT=2^\circ C$), the ratio can decrease to 1.2-1.4. During SD warm-up, the ratio increases to 6-20.

Fig.9 Overview of super-saturation of ice on the DS with respect to saturated air for different air temperatures and $dT$.

A wide range of experimental and theoretical research efforts have characterized the effects of temperature and over saturation on ice crystal growth rates and morphology under conditions relevant to atmospheric processes (for example Lamb and Hobbs, 1971; Libbrecht, 2005; Libbrecht, 2017). The wide variety of ice crystals found in nature has sparked interest. Sublimation was sometimes regarded either as the opposite process, or a less intriguing process, and was less visited in lab studies.
Shaw and Mason (1955) investigated ice crystals placed on a surface. They discovered that at constant temperature, the growth rate of the facial area varied according to the square of the vapour oversaturation. A critical over-saturation for growth and sublimation was found. For sublimation, the critical super-saturation was on the order of 0.85 at -22°C. The presence of a critical temperature was probably a result of molecular bonding with the substrate (Nelson and Knight 1996). Nelson, (1988) sublimated numerous, 100 um diameter plate crystals (0.1°C>T>-18°C, 0.05% to 5% under-saturation) showing that the crystals first lost sharp edges, and finally evolved into spheroidal particles, and the aspect ratio remained almost constant. The sublimation rates were accurately predicted by the diffusion equation with the surface vapour density at the equilibrium value for a uniform surface temperature. The steady-state shape of the sublimating crystal depends on the shape preserving solutions of the diffusion equation. Ham (1959) showed that ellipsoids and thus spheroids preserve shape during growth and sublimation if the surface has a uniform temperature. Jambon-Puillet et al. (2018) also showed experimentally and theoretically that sublimation first smooths out regions of sharp curvature, leading to an ellipsoid. The second stage is the sublimation of the self preserving ellipsoid form. The entire process may be modelled as a vapor diffusion problem, mathematically equivalent to the resolution of the electrical potential around a charged conductor. Using this analogy, they provided a mathematical method for simulating the sublimation of the particle. The sublimation of the ellipsoid resulted mathematically simple, and their method was employed in this work to simulate the sublimation of ICE-CAMERA particles.

Oblate spheroids with an aspect ratio (AR) of 10, in thermal equilibrium with the DS, were assumed in the simulations. In the model, particle diameter, DS temperature, air temperature and relative humidity can be changed. The time required for full sublimation of the particle has been computed. As sublimation accelerates when the particle is going to vanish, the time necessary for the complete sublimation is only slightly larger from that necessary to reduce the particle to the minimum particle size (60 um) typically accepted by ICE-CAMERA data processing. The simulations assume the completion of preliminary sublimation of points and edges of the particle, so that the calculated time should be considered as a lower limit of the duration of sublimation. Figure 10 shows the duration of complete sublimation under heated SD conditions. With the DS heated at \(dT=25°C\), the air humidity was irrelevant for the results. Results show that at -30°C air temperature (summer conditions in DC) complete sublimation can occur very rapidly on the DS (for all sizes up to 1 mm), occurring within a few minutes after attaining the DS sublimation temperature. At lower air temperatures, the sublimation period increases. At a temperature of -70 °C (winter temperature in DC), particles smaller than 100 um in diameter still evaporate within 10 minutes, while 1 mm particles could survive along the heating period.

12
After the sublimation period, DS is exposed to falling crystals. During this period, the collected particles also undergo sublimation, although this is limited. Figure 11 shows the sublimation time as particles are collected, assuming they are in thermal equilibrium with the DS. The DS was considered to be 5°C warmer than air. In this case, the relative humidity of the air also has a role. In summer (-30°C air temperature), sublimation can take just a few minutes for particles smaller than 200 um. In winter (air temperature -70°C), all particles are expected to survive sublimation.

These simulations show that the lower limit of particle detection by ICE-CAMERA is not only limited by the resolution of the optical system and/or image processing software. In summer, particles below 100 um could be decimated during the deposition time, unless they fall just before the scan. Particles of that kind dominate diamond dust events. Therefore, ICE-CAMERA is better suited for the study of cloud precipitation, at least during the summer period. By the way, the visual screening of the ICE-CAMERA images showed just a few small particles with signs of partial sublimation, such as rounded corners or smooth edges. Moreover, even in summer, small DD particles such as plates were normally observed (section 4). It is likely that most particles never reach thermal equilibrium with the DS, and the results in Figure 11 should be considered the worst.
Fig. 11 Sublimation time for Oblate ice particles at different air temperatures and humidity during the deposition phase.
ICE-CAMERA is not just designed to take photographs of ice particles, but to provide automatic morphometry and classification of polar precipitation. This task was accomplished through image processing and machine learning techniques. The process is divided into two parts: image segmentation/measuring, and ice habit classification.

2.1 Image Segmentation and Measurement

After acquisition, the raw ICE-CAMERA images are segmented, by means of MATLAB, to isolate all detected particles. The process is as follows (refer to Pratt (2007) for nomenclature):

- background subtraction.
- intensity equalization.
- Region growing using as ‘seeds’ the brightest points of the image. Region growing normalized-intensity range: 0.2-1.
- Image closure by 10 pixels.
- “majority” correction (all pixels surrounded by 4 other "one" pixels are set to "one").
- Image Segmentation and Labeling of all the 8-connected regions found.
- Selection of the largest-area labeled region.
- Creating the best adapted Feret bounding box.
- Measuring Feret’s length and width (Walton, 1948).
- Calculating the Aspect Ratio (AR = Feret’s length / width).
- Measuring projected surface and surface-equivalent diameter.
- Calculating standard shape parameters (refer to Russ and Brent Neal (2017) for nomenclature)
  - Eccentricity, Euler Number, circularity, roundness, solidity, compactness, form factor, number of skeletal branches.
- Calculating normalized central moments f1...f7 (Hu, 1962).
- Image rotation (horizontally-oriented)
- Image resize to CNN input size (see section 2.2)

The whole process is summarized in Figure 12 for a rimed, columnar particle.
Fig. 12: The original image is segmented using ‘region-growing’. The particle projected surface (clear blue) is computed. The boundary box (green) is determined and the Feret length and width measured. The picture is finally rotated, re-framed, aligned horizontally, and resized to the CNN input size.

2.1.1 Summary-image of detected particles.

The bounding boxes of all individual ice particles detected in a scan are sorted by Feret width, and reassembled in a summary image collecting all segmented particles (Fig 13). Each particle is also associated with a numerical record containing the coordinates of its bounding rectangle on the synthesis image, shape parameters, measures, time of acquisition and local weather data. In this way, the re-analysis of the summary image is possible instead of re-processing the original, large image.
2.1.2 Limitations and uncertainties in detecting and sizing ice particles.

1) The total number of particles measured is limited to 2000 as a result of MATLAB memory. Extra particles are not treated.

2) By default, particles below 4000 um² in bounding-box surface (approximately D<60um) are not preprocessed (smaller particles could be detected, but most have a seemingly circular shape due to low pixelation or poor focus).

3) Segmentation becomes difficult when overlapping particles or aggregates of particles are present. In such situations, double counting of the same particle or/and segmentation errors may occur for up to 12% of cases in the presence of an "intense" precipitation event. Overlapping particles are normally classed by the CNN algorithm as "clusters" or "irregular grains".

Fig. 13 Example of a summary-image for the particles detected in a scan.
In the case of defocused images, the particle shapes are all close to a fuzzy, round shape, which can cause misclassification. Moreover, a few big particles are rounded by partial sublimation, and apparently rounded or "spheroidal" particles of 500 um diameter or greater are thus sometimes detected, but are disregarded in the statistical analysis.

Needles and hexagonal plates may be very bright in ICE-CAMERA images due to increased diffusion of light at preferred angles. In the case of needles, this effect can reduce the apparent aspect ratio, as the width is apparently increased by the scattered light saturating the camera. For plates, the specular effect blurs sometimes the contour and cancels out the hexagonal shape, especially in the case of small plates.

2.2 Automated classification of ice particles.

A first attempt at automatically classifying ice particles was made in 2014 using shape factors. This type of technique was also used by other authors (e.g. Lindqvist et al., 2012) for attempting the classification of ice particles. In the case of ICE-CAMERA images this resulted in an extremely unreliable approach. A much more promising approach was offered after 2015 by the rapid development of transfer learning and convolutional neural networks (CNN) (Le Cun et al., 2015; Schmidhuber, 2014). Xiao et al. (2019) successfully applied deep transfer learning to ice particle images obtained with airborne Cloud Particle Imagers (CPI). The CNN approach has added much value to ICE-CAMERA because a reliable classification of ice particles into simplified classes became possible. The CNN chosen for the ICE-CAMERA particle classification is "Googlenet" (Szegedy et al. 2015, a variant of the Inception network, a deep convolutional neuronal network developed by Google scientists. Googlenet is a type of convolutional neural network based on the Inception architecture. It utilises Inception modules, which allow the network to choose between multiple convolutional filter sizes in each block. The GoogLeNet architecture consists of 22 layers (27 layers including pooling layers), and part of these layers are a total of 9 inception modules. In this work, GoogleNet was used in Mathworks MATLAB R2020b environment. The Googlenet CNN, pretrained on the ImageNet data set (Deng et al. 2009), was loaded, then the final, fully connected layer of the model was changed to size 14. The input layer of the GoogLeNet architecture requires images of size 224 x 224.

2.2.1 The CNN classification classes

The low temperature and humidity conditions encountered on the high Antarctic plateau reduce the diversity of forms of ice particles. This is both observed on the field at DC, at South Pole station (Lawson et al., 2006), and suggested by review works such as Bailey and Hallett (2009). Following an initial survey of the ICE-CAMERA database, a set of 14 types of particles was selected, as shown in Figure 14. When choosing the 14 classes, I presumed that shapes easily recognizable by a human operator could also be easily
recognizable by a CNN. In the following scheme I tried to relate the classes chosen for ICE-CAMERA with the classification of the ice particles of Kikuchi et al. (2013), an updated version of the original classification of Magono and Lee (1966).

Fig. 14: A small sample of ICE-CAMERA images of the 14 classes of ice particles used to train the CNN.
-needles: covering the classes C1a,C1b,C3d (Kikuchi et al. 2013)

-bullets: covering the C4b-C4c classes.

-hexagonal prisms: columns covering classes C2a, R2b, C3a, C3b.

-hexagonal plates: covering classes P1a,P1b,P1c ,P4f,G2a,G3a,CP3f,CP3d.

-trigonal plates: covering the classe G2b.

-irregular plates: plate-like particles with irregularities, riming, overgrowing plates, etc. but keeping a basic hexagonal shape, covering P6a,P6b,P7a,CP6d,R1b,R2b,R2c,R3a,G4b.

-rosettes: bullet-rosettes or column-rosettes, with a minimum of two branches, covering C2c,C3e,C4d

-irregular rosettes: rosettes with irregularities, riming, but preserving the typical stellar outline of rosettes. Covering classes P7a,P7b,CP2d ,CP4c,CP5a,CP6c,CP6f,CP6g,R1d

-irregular grains: covering CP3e, CP5a, CP6d, G4c,G4a,I3a,I2a,I1a,H1a,H1b

-spheroidal: particles that seem round, covering H1a,H1c. (Large particles (D>600um) detected as ‘spheroidal’ in DC are usually artifacts caused by defocused images and are not considered in the statistical analysis).

-compact columns: short columns covering classes G1a,C3a

-clusters of particles: covering A1a,A3a,H2a,H1b,P8b,CP3e,CP5a,CP6h

-frost: frost formed on the DS plate CP7,CP8,CP9

-fibers: non-volatile fibrous material (from local human activities)

The last two classes are not considered in the statistical analysis of ICE-CAMERA data: they are just used to detect eventual frost formed on the DS, and man-made insoluble (thus persisting on the DS) materials. Uncommon ice particle typologies present at Concordia were not considered in the present work. Trigonal plates have been included, even if rare, simply because of their apparently easy detection using the CNN.

2.2.2 The training data-set

For the first CNN training, 5500 ICE-CAMERA images have been manually sorted into 14 image data stores, corresponding to the 14 classes. 14 of the computer keyboard keys have been marked with pictures of the 14 classes, to speed up the manual classification of the initial training dataset. These images were used for an early CNN training. The images classified by this first CNN were manually ‘purified’ from incorrectly classified images and added to the 14 training image data stores, as new training images for a second CNN training. The operation was repeated twice again in order to improve the CNN precision.

Figure 15 shows the final number of training images for each class. The total number of images used for the training is 25705
2.2.3 CNN training details

To satisfy Googlenet input requirements, all single particle images were resized to 224*224 pixels. In the training process, ‘data augmentation’ was applied to the original data set. Artificially ‘augmenting’ the image data set has been shown to be effective in CNN training (Shorten and Khoshgoftaar, 2019). The dataset is ‘augmented’ to artificially increase the size of the training dataset of Figure 15, while reducing CNN overfitting. The following random transformations were used in augmentation:

- X,Y reflection
- random X,Y translations ±30 pixels
- Random scaling 60-100%

Other changes such as rotation have not been introduced since the ICE-CAMERA images to be classified are typically oriented horizontally by the image processing procedure (e.g. Fig.12)

Fig.15 The final number of CNN-training images
In the GoogleNet training, the following options were utilized:

Solver: stochastic gradient descent with momentum (SGDM)
Activation: softmax
Number of Epochs=5
Learn Rate=0.001
Batch Size=128
L2 weight regularization factor=0.005
Validation frequency= every 30 iterations
Shuffle of the dataset at every epoch

10% of the image dataset was dedicated to validation, 90% for the training. The evolution of the CNN training in terms of accuracy and losses is presented in Fig.16. The validation line closely tracks the training line, demonstrating the absence of overfitting.

Fig.16: Evolution of the CNN training. Each ‘epoch’ correspond to the usage of the entire training dataset
3 Testing the CNN classifier

The results of the CNN performance test are summarized in the confusion matrix plot of Figure 17. Each row correspond to a predicted class (Output Class) and each column correspond to a true class (Target Class). Diagonal cells refer to correctly classified observations. Off-diagonal cells are improperly classified observations (red markings increasing misclassification).

The column on the far right of the plot shows the percentages of all the examples predicted to belong to each class that are correctly and incorrectly classified (positive predictive value and false discovery rates, respectively). The row at the bottom of the plot shows the percentages of all the examples belonging to each class that are correctly and incorrectly classified (true positive rate and false negative rate, respectively).

3.1 Precision of the classifier:

In the column-normalized summary (Fig. 17), the percentages along the i-th column show the probability (P) of a "true" particle in class -i being classified in each of the 14 output classes.

---

**Fig.17 confusion plot of the CNN (column-normalized)**
Reading the columns of Figure 17 from left to right, the precision of the CNN in properly classifying a particle belonging to the i-th true class (bottom row) can be assessed. Results are:

- Good precision (P>90%) in the identification of needles, spheroidal, clusters, prisms, pristine rosettes.
- Hexagonal plates and irregular hexagonal plates are mistaken in approximately 10% of times.
- Irregular grains are sometimes (12%) mistaken with irregular plates.
- Compact columns are misclassified almost 20% of times.
- Irregular rosettes are misclassified 9% of the time as pristine rosettes, and 5% of the time as irregular grains.
- Frost cases are also confused with irregular grains in 13% of cases.
- Fibers are sometimes confounded with rosettes (13% of cases). This is possible because I considered rosettes also cases with only two branches, a form rather similar to a curly fiber. This ambiguity is not a problem, as the number of fibers in the images is several orders of magnitude smaller than that of rosettes.
- Bullets are often (25%) misclassified as prisms.
- An interesting finding relates to triangular plates. In fact, I expected that CNN would detect this very special form very effectively. Probably due to the limited set of training data for this class, the CNN did not succeed in detecting this type of particles, with only 6% of correct classification. Fortunately, triangular plates are uncommon in DC, and are not considered in ICE-CAMERA statistics.
- The three-dimensional structure of the ice particles is lost in the ICE-CAMERA images, so that some thick ice forms such as C4a, P1b, G3b, CP1a, etc. (Kikuchi et al, 2013), if any, are likely to be misclassified by this CNN.
### 3.2 Positive predictive values

A different view to read the CNN test is the row-normalized summary of the confusion matrix (Fig. 18).

#### Fig. 18 confusion plot of the CNN – (row-normalized)

Percentages along the i-th row now show the probability for a particle classified into the i-th class to effectively belong to each of the 14 true classes. Reading the rows of Figure 18 from top to bottom, results are:

- Particles classified as bullets, needles, clusters, spheroidal, pristine and irregular rosettes effectively (P>90%) belong to this class.
- Particles classified as compact prisms have a 86% chance of being a compact prism, 5% to be a bullet, 6% to be a prism.
- Particles classified as prisms have a 78% chance of being a prism and 17% of being bullets.
- Particles classified as bullets are, in effect, 98% bullets.
- Particles classified as irregular plates have a 74% chance of being irregular plates, 5% plates, and 17% of being irregular grains.
- Particles classified as pristine plates have a 70% chance of being plates, 12% to be irregular plates, and 14% of being irregular grains.
4. The CNN classifier applied to ICE-CAMERA data

The classifier was used on the ICE-CAMERA 2014-2017 dataset. Even if the detailed analysis of these data is the task of a separate paper, a sample of the capacity of the instrument is presented here for the January-February 2017 summer period.

The number of particles classified in the 14 classes (Figure 19) shows the relative scarcity in this period of particles classified as fibres, trigonal plates, compact prisms, irregular rosettes and frost.

Figure 19 number of particles in the 14 classes for the period January to February 2017.

Figure 20 shows Feret's length statistics in boxes and whiskers. The particles were sorted into the classes described in 2.2.1, with the exception of trigonal plates, frost and fibers. On each box, the central mark indicates the median, and the bottom and top edges indicate the 25th and 75th percentiles, respectively. The maximum whisker length is here equal to the interquartile range. The graph shows some relevant differences compared to the results obtained by Lawson et al. (2006) at the South Pole in summer: ICE-CAMERA found larger particles. As described in point 1.6.1, small particles with D<100um may be lost by sublimation on the DS during deposition, and may therefore not be present in the statistics. This could be a reason for a bias of ICE-CAMERA median sizes towards larger sizes in the case of typically small particles such as plates. In the case of typically larger particles, like rosettes, this is not the case.
Figure 21 shows the probability distribution of Feret length for plates and rosettes (Feret length is close to the diameter for both classes). For plates, the peak-size is 75 um, which is similar to the results of Lawson et al. (2006) considering that particulates below 50 um are not treated with ICE-CAMERA. This finding indicates that the possible, severe sublimation of particles below 100-200 um in summer is not relevant. The results for the Feret’s size of rosettes differ considerably from those of Lawson et al. (2006) because the size is approximately 600 um, compared to 100 um in their study. This effect is unexplainable with the sublimation of the smallest rosettes. Instead, the distribution is supported by direct visual observation of precipitation in CD.
Figure 21: (January-February 2017). Feret Length for pristine rosettes and plates. The dominant presence of small plates (D<100um) in summer shows that sublimation on the DS is not relevant.

Figure 22 shows the Feret aspect ratio for each class. As expected, many “rounded” classes have an AR<2. Compact columns show a median AR close to 1.4, while columns and bullets are close to 2.8. The average AR for the needles was 3.3, which is lower than expected for the reasons explained in point 2.1.2.
Figure 23 shows the equivalent diameter of the particles, computed in two different ways. The bounding-box equivalent diameter is the diameter of the circle equivalent to the bounding box. The projected-area equivalent diameter is the diameter of the circle equivalent to the surface of the particle contained in its contour (Fig. 12). The difference is relevant for "fluffy" particles like rosettes and clusters. For these particles, the projected-surface equivalent diameter resulted typically 40% less than the equivalent diameter of the bounding box.

Figure 22: (January-February 2017). Aspect Ratio statistics
Conclusions

The instrument described here, while very similar to a simple flatbed scanner in its basic design, represented a technical challenge for its implementation at DC. Hardware and software have been continuously and extensively modified at DC over the past five summer campaigns. Commercial or customized instruments do not have this flexibility, more typical of old-style handcrafted products. The result is now a reliable instrument, operating all year round on an hourly basis, for the statistical study of precipitation in internal polar areas. The morphology and size of the particles are automatically obtained, and some semi-quantitative precipitation assessments can be derived. The collected data are automatically pre-analysed, but they can be post-processed at any time, in order to follow the continuous improvements of the image processing and machine learning algorithms. The GoogleNet convolutive neural network, trained specifically for this instrument, has succeeded in classifying ICE-CAMERA images into 14 form classes, with an accuracy of more than 80% for most classes, with the exception of fibers, bullets (mistaken with columns), and trigonal plates. Fibers are not part of precipitation, and trigonal plates are scarce in DC anyway.

The instrument is particularly useful for precisely and automatically measuring the size of individual ice particles in precipitation, a process virtually impossible manually, and certainly impossible on the field in DC in winter. ICE-CAMERA scans are performed on hourly basis. Keeping the surface of the instrument free of frost all the time and cleaning it by heating after each scan is paid with the possible loss of the small fraction of ice particles. Particles below 100-200um (e.g. diamond dust) may be affected by sublimation before imaging. This problem is complementary to the problem encountered when...
observing manually: when observing precipitation manually every 24 hours, (as is the case of DC) the reprocessing of particles, or the formation of ice and hoar artifacts cannot be prevented. In ICE-CAMERA, frost and ice regrow are suppressed, but small particles may disappear for sublimation. ICE-CAMERA data, collected since 2014, have already undergone statistical processing and the complete results will be described in a separate paper. A small subset of data gathered during the 2017 summer period was presented in this work. It demonstrated the capability of the instrument to classify and size actual precipitation data at DC, without relevant small particle losses for sublimation. The presence of cloud precipitation particles (rosettes) much greater than those observed at the South Pole by Lawson et al (2006) was also identified.

Code and Data availability

As a result of the current CNR and PNRA governmental policies, the CNN developed as part of this work, along with a limited test dataset (224*224 images for the 14 classes of particles), are only available from the author upon reasonable request by e-mail (massimo.delguasta@ino.cnr.it). The CNN was developed under MATLAB.

Acknowledgements

I am grateful to the Italian Antarctic Project PNRA for supporting this work with the projects ICE-CAMERA (PNRA 2009/A4.1) and PRE-REC (PNRA 2013/AC3.05). I am also grateful to the ‘Osservatorio Meteo-Climatologico Antartico’ (PNRA 14_00100) for the Meteo data, and to all the logistics staff and winter-over crews of Concordia station, all working hard to permit our scientific activity.
References


Huang, J. A Simple Accurate Formula for Calculating Saturation Vapor Pressure of Water and Ice, J. of Appl. Meteorology and Climatology, 57, 6, 1265-1272, doi: 10.1175/JAMC-D-17-0334.1, 2018


doi:10.2151/jmsj1923.41.5_305, 1963

http://www.jstor.org/stable/26179423

doi:10.1109/CVPR.2015.7298594, 2015


Walton, W.: Feret’s Statistical Diameter as a Measure of Particle Size, Nature, 162, 329–330 ,doi:10.1038/162329b0, 1948