Machine Learning Based Classification of Snowflake Geometries in Multi-Camera Observation Systems

Hein Thant[(1)], Mikhail Zhizhin[(2)], and Branislav M. Notaros[(1)]

(1) Electrical & Computer Engineering Department, Colorado State University
(hthant@colostate.edu)
(2) Payne Institute for Public Policy, Colorado School of Mines

Substantial developments in machine learning have enabled comprehensive image classification methods with common supervised training methods such as Convolutional Neural Networks (CNNs) and more advanced techniques such as Trilinear Attention Sampling Network (TASN) or unsupervised learning methods including K-means clustering aided by dimensional reduction methods such as Principal Component Analysis (PCA).

However, there has been relatively little work done in the field of automatic classification of snowflakes, with the previous attempts not being able to reach the whole depth of classes of known snowflake geometries. As latest studies suggest, we have eight main categories of frozen hydrometeors, seven of which are snow. Then we have about 130 classes of different snow geometries. Although the main classes could be accurately differentiable using commonplace methods such as CNNs or K-means classification, to get deeper into the subclasses and tiny features, we need more advanced techniques such as the TASN.

In this study, we have involved the use of data collected from Snowflake Measurement and Analysis System (SMAS) and Multi-Angle Snowflake Camera (MASC). The SMAS was developed at the Colorado State University (CSU) and is a system that consists of seven high resolution cameras. The MASC is a commercially available device that was developed by Particle Flux Analytics in Utah and is a system that consist of three high resolution cameras. We also have a CSU-modified MASC that contains five cameras.

A strict procedure of automatic preprocessing to crop and clean our images has been developed to ensure proper training of the classifiers that are not impacted by the artifacts in the images. This present study involves the use of a combination of different types of machine learning to find out an optimal classifier we can develop for a whole set of snowflakes characteristic not only for Colorado, USA, but for different regions of the world. This is especially important given that snowflakes form different shapes under different climatological and weather (e.g., temperature and humidity) conditions.