EGU24-6010, updated on 08 Oct 2024
https://doi.org/10.5194/egusphere-egu24-6010
EGU General Assembly 2024


# Ice crystal images classification using semi-supervised contrastive learning

**Yunpei Chu**[1], Huiying Zhang[1], Xia Li[2], and Jan Henneberger[1]
[1]Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland
[2]Institute for Machine Learning, ETH Zürich, Zürich, Switzerland

Ice crystals play a crucial role in precipitation formation and radiation budget, with their various shapes influencing these processes differently. The shape of ice crystals is related to the environmental conditions (i.e. temperature) under which the ice crystal forms and the microphysical processes that ice crystal experiences. Therefore, ice crystal shape classification is important for understanding conditions and microphysical processes in cloud. However, current methods are mainly supervised learning algorithms like convolutional neural networks (CNNs), heavily relying on extensive manual labelling, which requires substantial labor. Moreover, the limitations in human's knowledge of ice crystals and the bias of human subjectivity in classification hinder the generalization ability of these networks. In response to these challenges, we propose a semi-supervised algorithm for ice crystal classification. We use data from the 2019 Ny-Ålesund NASCENT campaign, collected by a holographic imager mounted on the balloon-borne platform HoloBalloon, which includes 18,864 ice crystal images. In our algorithm we initially extract key features from ice crystal images using an unsupervised learning network, prioritizing generalization rather than dependence on labelled data, which ensures unbiased feature identification. Subsequently, a small subset of images is manually labelled into nineteen categories based on a multi-label classification scheme that consider both basic habits and microphysical processes. The classification accuracy of our hybrid algorithm on nineteen categories is similar to the performance supervised learning algorithm. This hybrid algorithm not only reduces the labor needed for manual labelling but also incorporates physics-based constraints, which prevents the network from making unfounded assumptions, thus offering a robust and efficient framework for ice crystal classification.