

Interactive comment on “Unsupervised classification of snowflake images using a generative adversarial network and K -medoids classification” by Jussi Leinonen and Alexis Berne

Anonymous Referee #1

Received and published: 14 January 2020

The authors present a method for unsupervised classification of snow particle images obtained by the Multi-Angle Snowflake Camera (MASC) and demonstrate the ability of the resulting classifications to distinguish features of the snow particles. The topic is relevant - there is a need to be able to analyze snow particle imagery in a mostly automated fashion and to relate the imagery to the microphysical processes that produced a given snow particle. The work is well-structured and presented clearly aside from a some particular details noted below that are related to the mathematical details of the algorithm. Aside from these concerns, the method and assumptions are clear and valid and the results sufficient to support the principal conclusion: that the unsupervised method can distinguish snowflakes based on size, shape and texture, and that

C1

the performance of the unsupervised method is reasonably consistent with that of a more labor-intensive supervised method. This unsupervised method has application beyond the MASC to include other imaging disdrometers and particle probes.

My principal concerns are related to some inconsistencies and lack of clarity in the description of the method. My specific comments for this are as follows:

Line 66: Can you clarify the meaning of "deep" CNNs?

Line 78: Can you clarify the meaning of "latent variables"?

Line 108: How does a simple constraint on the diameter of an imaged snowflake ensure that the entire snowflake is within the image frame? Couldn't a snowflake that intersects the edge of the frame have a diameter in this range?

Line 115: What is the specific purpose for downsampling? Is it simply to make the classification processing more computationally tractable?

Line 153-155: It's not clear how "neighborhoods" are defined in the context of a set of snowfall image inputs. Can you elaborate?

Lines 183-184: What is "z"? (see also the comment regarding line 264 below)

Lines 221-222: Is it actually true that the distance between each point and its nearest centerpoint is minimized? I don't believe that is what is imposed by equation 14. But what does lowercase "n" represent in equations 14 and 15?

Lines 260-262: It would help here to have some additional context describing the purpose of a styling block. What is achieved by upscaling the image and processing it through the AdaIN, activation and convolution layers? What is gained by upscaling the image?

Line 264: Earlier (line 211), z is described as the latent distribution. Here it is described as noise, and this seems inconsistent. Can you clarify? What is the difference between "style" and "latent variable"?

C2

Line 340: Does it not appear that there is a threshold near $K=3, 4$ or 5 ? There seems to be a substantial change in the slope of the loss function near these K values. Why would this not be seen as an indication of the actual number of medoids?

Finally, I have two technical comments:

Figure 1: I believe the caption is wrong. Panel (a) appears to be the discriminator, and panel (b) the generator.

Line 232: Should this be "understood as a variant"?

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-440, 2019.