Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., author comment AC2
https://doi.org/10.5194/amt-2022-72-AC2, 2022

Reply on RC2

Louis Jaffeux et al.

Author comment on "Ice crystal images from optical array probes: classification with convolutional neural networks" by Louis Jaffeux et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-72-AC2, 2022

Thank you for your review on our article "Ice crystals images from Optical Array Probes: classification with Convolutional Neural Networks".

First and second major comments: Few lines added in the caption of Table 2 (could also be added to third point of Section 2)

"All the PIP images in the original data set originate from 2 events in the EXAEDRE campaign and 3 events of the HAIC campaign. The context of these events are thunderstorms in the Mediteranean Sea for EXAEDRE and Mesoscale Convective Systems in french Guiana for HAIC. Most of the 2DS data also originate from the same events. The AFFLUX campaign data was extracted from a single flight in Arctic clouds to provide more Col, CBC and HPC of various sizes. Finally, all the water droplets of the 2DS data set were captured during a single flight in liquid water clouds in the Carribean Sea during the EUREC4A project."

Third major comment: table 3 shows the number of images in each category in parenthesis

Fourth major comment on Figure 5 and 6: I added a clearer description of the mentionned figures in the text below it (sections 4.2.1 and 4.2.2):

"Figure 5 details the overall results of the random inspections, 5a displays all 4000 responses from the ten operators, normalized, while 5b shows how the 400 images are classified by humans and the network in numbers, in this second case a majority rule is used to determine the class attributed by humans, if majority (more than 50%) is not reached for a given image then it is considered unidentified by humans."

The train of thought behind these two plots is to be able to strictly confront human inspections with the network with the first subfigure and to be able to look more closely at images in particular with the second one, such as the ones displayed in the Appendix (A3,A4,A5), in order to see whether the images themselves are problematic, or if the network is producing inaccurate results or even if there were flaws in the way operators were trained.

Second minor comment: we suggest an alternative text

"Appendix Figure A3 shows the crystal images that are referred to in Figure 5b, for CP,RA and FA classes. Inspecting these sets of images, **CNN: CP/ Human: RA** are objectively closer to **CNN: CP/ Human: CP** than to **CNN: RA/ Human: RA**, and **CNN: FA/ Human: RA** is arguably closer to **CNN: FA/ Human: FA** than to **CNN: RA/ Human: RA**. It appears that the CNN has developed more consistent class definitions and is therefore superior to the humans in discriminating between the three classes."

Using the Figure A3 so directly might also require that we move this Figure from the Appendix to the main text.