Review for “Neural network processing of holographic images” by Schreck et. al.

The manuscript “Neural network processing of holographic images” describes and evaluates a novel approach to processing holographic images using Machine Learning and Neural Networks. The authors used an existing dataset obtained with the holographic detector for clouds (HOLODEC) to compare current, standard methods to their new approach, and used synthetic holograms to train the neural network. The new approach has the potential to substantially speed up the data processing not only for HOLODEC, as shown here, but also other holographic imagers and therefore is useful for anyone applying holographic methods to in-situ or laboratory measurements. Although it is still not a fully automated process yet, this approach is very important and after further development may play a substantial role for future holographic data analysis. In general, the manuscript is well written, and I recommend publication with minor revisions.

Main comments

Line 180: You used 500 particles per hologram. However, holograms frequently have higher numbers of particles (few thousands). Why did you choose this number, and would the results differ if there were a factor of 10 more in each hologram?

You state that the synthetic holograms you used were corrupted by noise processes. It is unclear to me how exactly this was achieved. Why did you not use real empty holograms to inform what it would look like and use it as background for the synthetic holograms? That way you’d have the same background from the synthetic holograms as are present in reality. Since you explain towards the end of the manuscript you need to retrain the models before each field project anyways, you could include that way the field project specific noise.

Line 27/28: Where do these numbers come from? HOLODEC has a sample volume of ideally nearly 19 cm^3 (maximum x*y*z distance). Not all can be used, since particles are not uniformally detected, especially close to the edges. So, typically a sample volume of around 13 cm^3 has been used in the past, or even less, depending on the conditions. The sample area in x*y also decreases with distance z, leading to a cone shaped sample volume. This should be considered and explained. Similarly, each pixel is about 3 um wide, leading to the minimum detectable particle of about 6 um – but the resolution would still be 3 um.

Suggestions for Figures for easier understanding

Generally, Figures are good quality. However, some could be improved to guide the reader. Here are my suggestions:

None of the images have a colorbar, and while it may be acceptable in some cases, in examples such as e.g. Fig. 13 where particles are shown, it would be interesting to see if all panels have the same values or if the images were scaled.

Fig. 6: both panels in 6(a) have y-axes that have a gap between zero and the end. It makes it hard to see if the values in the plot are zero, or above. Either zero the axes or add gridlines for clarity.

Fig. 7/9: What exactly is the difference between Fig. 7 bottom row and Fig. 9 second row? The plots look identical to me and the description does not make it clearer. If they are the same, I'd suggest combining Figure 7 and 9 – if not, make clearer what the difference exactly is.

Fig. 11: (a) and (b) are mentioned in the caption, but not included in the figure.

Fig. 12:

- The colorbar is missing for all panels. The question is, is it the same colorbar in all figures or not? If it is, then it is noteworthy that the different approaches also size particles differently. In e.g. (c)(ii) the same particles than in (a) have a different color, and more orange/red particles are shown. Are the different methods influencing sizing that much?
- (a)(i) and (a)(ii) are the same – why not have it only once, maybe centered between the two rows? It took a while for me to realize they are the same and I looked for the differences.
- I find the naming (b)(i) and (b)(ii) in this case unnecessary and confusing, and naming (a)-(e) would work just as well here. But this is just a suggestion.
- It might be worth mentioning that the three axes are not equal size and therefore distances between particles are not to scale.

Language

Overall, the language is good. However, please check spelling and grammar carefully. Here is a list of things I found:

Line 220: "…number **of** particles…"

Line 241: "…number **of** pixels…"

Line 354: "…explicitly,_a transformation…" (space missing)

Line 412: "…trials…"

Line 465: N_S is explained here, but N_SH not. I could not find any specific definition for it in the manuscript.

Line 487: All abbreviations are explained here, but F1 score is not. I could not find any specific definition for it in the manuscript.

Line 643: "…architecture…"