Dear TC editor and authors of the revised manuscript egusphere-2023-295,

The manuscript topic is interesting and ice floe information is useful for multiple purposes. The other two reviewers have already provided good comments to improve the manuscript. I hope my comments will complement the comments of the other reviewers.

As suggested by another reviewer the abstract needs to be updated to include more on the applied method and results provided by the method instead of general level information. The abstract also begins with FSD and in the manuscript FSD is in the section named as "Case study:FSD" and FSD does not appear in the manuscript title. I suggest at least to remove "Case study:" from the section 6 title, also consider including FSD in the manuscript title. FSD is also significantly present in the introduction section.

The amount of data is very restricted, Why only four Sentinel-2 images have been used? There exist a lot of Sentinel-2 data. It should be emphasized that with such limited data sets this is a case study and the results can possibly not be generalized.

A proper cloud mask is required to be able to automatically segment ice floes. As "manual-free" is mentioned in the manuscript title I think cloud masking should be discussed in the manuscript. Does there exist automated methods for reliable cloud masking or at least excluding images with clouds? Give references of possible cloud masking approaches or suggestions for improved automated cloud masking. Could "manual-free" in the title be "automated" instead?

The results and discussion have now been presented in the same section. I suggest to make a separate "Discussion" section or rather combine "Discussion" with the "Conclusions" section which is very short now.

P2 L36 "Copernicus": give a reference

P3 Dataset → Airborne data:
What is the number of airborne images used in this study? What is "a large amount"? Rather give numbers.
Were the images used as long strips, mosaics, or single shots?

P3 Table 1: Give flight altitude(s) and surface area covered, or their ranges, for the images used. These could possibly be included in the table.

P3 L83 "Cophub": Give reference (URL).

P4 Table 2: Include location information and covered area in the table, e.g. by given center latitude and longitude (and covered area e.g. in km2).

P6 L110-112: Hypothesis of improvement by widening the boundaries would require some evidence. Would it be possible to show test results with a small set of imagery and some numeric evidence based on these tests?

P8 S3.2. Deep learning model: Give at least a short description of U-Net++ giving best results or a diagram of the network. Now this subsection is very short and it is essential for the study.

P8 Post-processing: Applying morphological opening and closing seems a bit heuristic to me. Are there any references or if not would it be possible to demonstrate the benefits of the morphological processing? What is the shape and size of the morphological operator (often a disk is used)? Could

this step be included in the ML algorithm somehow, i.e. could the NN learn the post-processing?

P9 Training: GPU memory is referred on line 170 and this information is then given later on page 10. The hardware (and software) used should be given before it is referred. The used HW and SW could e.g. be included in the dataset section and changing the title to something like "Datasets and computational resources". Also include the used SW with reference in the same section. Also mention there that all the execution times given later are given for this specific configuration.

P10 L174: Does this distribution of classes correspond to their distribution in general? Then it can be used in training. What happens if the distribution of classes is balanced (33% of samples for each class) for the training? Does balancing degrade the classification?

P10 L175: The number of test samples is not very large. What is the effect of reducing samples in training and validation data sets and increasing of the test data set? Are these numbers of samples selected based on some kind of performed tests?

P11 Section 5: Would be good to have some kind of related introductory text under "5 Experimental results and discussions" and "5.1. DL model evaluation", now they are empty.

P13 Section 5.1.3. Inference time: Could this be "Segmentation time" instead, it would be more informative. The HW (and SW) used for segmentation could be given already in an earlier section, e.g. jointly with the introduction of data sets.

P15-15 Figure 10: Fig. 10 is nor in two parts and in two pages. Would it be possible to compress a little and make it to fot on one page?

P19 "6 Case study: floe size distribution": I recommend to drop "Case study:" because FSD is the essential parameter to be estimated by the method and it is also in essential role in the abstract and introduction and the whole manuscript is actually a case study because the datasets are quite limited.

Floe size distribution: Now FSD is estimated in two different resolutions (airborne and satellite data). It would be interesting to see how well FSD can be extrapolated from resolution to another (both from larger to smaller and smaller to larger) based on a fit distribution model. This would be very valuable information and this theme could be included in the discussion section.

P24 "Conclusions": This section is very short. Possibly it could be combined with a discussion section. Here could also be some conclusions on how close to automated FSD estimation the proposed method is? Could it be used for operational monitoring and what will still be required before possible. At least cloud masking should be discussed and also the annual period of possible operation (lighting conditions, what is the fraction of cloudless time in suitable lighting conditions in different sea ice covered areas). What are the ways forward in automated ice floe analysis?

Thank You.