The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2020-310-RC1, 2020

Interactive comment on "Image Classification of Marine-Terminating Outlet Glaciers using Deep Learning Methods" *by* Melanie Marochov et al.

Anonymous Referee #1

Received and published: 14 December 2020

General Comments:

This paper use a deep-learning-based workflow, termed CNN-Supervised Classification (CSC), to map glacial regions into seven classes using Sentinel-2 images. The method achieves reasonable results and shows its generalization ability. The author also shows significant improvements over traditional pixel-based methods such as band ratio techniques. There are some concerns regarding the explanation and technical details of the method, which are list below. Given this, I recommend this paper for publication after major revisions with attention to comments.

Specific comments:

I have two major concerns regarding the explanation and potential issues of the method

presented in this study.

1. The first concern I have relates to the superiority of the second phase model. The author mentioned that the second phase model (cCNN and MPL) is trained by the classification results of the phase one CNN model (Page 11, Line 310). And the author claimed that the second model outperforms the phase one CNN regarding the F1 scores. To me, the network cannot outperform the training label. For instance, Baumhoer et al. (2019) and Zhang et al. (2019) used the manual-prepared training labels to training the network, and the networks are eventually close to human-level performance but not exceed in terms of accuracy. Therefore, could the author provide more explanation of why the phase two model outperforms the phase one model?

2. The second concern relates to the method's performance on the edges of each class. Edges are important to glaciologists since that is where changes occur. The author only uses the pure tiles (Figure 5) to train the phase one model, which means the model might not have a promising performance on tiles with multiple classes (e.g., edges of the glacier or ice mélange). For phase two models, cCNN is for patch-based classification. Considering that a single patch could also contain multiple classes on edges and the phase two model is dependent on the phase one model, this method might have potential issues on the edges. It would be better if the author could quantify the method's performance on the edges or document such potential issues.

3. Page 8, Line 238: (1) How to get the variations of the surface meltwater on the glacier and ice mélange? They are not included in the seven classes. (2) It would be interesting to know how each characteristic can benefit the study of glaciology, for instance, the snow cover on bedrock.

4. Page 11, Line 311: It would be better if the author could provide more information about how to reassemble predicted classes to create a class raster. For instance, what is the stride size when predicting classes using a pre-trained CNN? How to deal with the overlap if there is any (when the stride size is smaller than the tile size)?

5. Figure 2: It would be better if the author could provide information about the median filter. What is the median filter for? Why is the median filter 1×1?

6. Figure 2: It would be better if the author could provide more details about vectorizing image features. For instance, how to deal with these impure patches (when the patch size is larger than 1)?

7. Page 14, Line 365: What is the stride size when using the second model to make the final classification? The stride size is important cause it influences the resolution (or size) of the final classified image.

8. Figure 8b: Could the author explain why they have unclassified regions? It seems that the edges of classes are usually unclassified (e.g., the black strip at the glacier front), which might also potentially influence the method's performance on the edges (See comment 2).

9. Page 24, Line 700: It would be better if the author could provide more a theoretical explanation about why some class confusion in phase one can be overcome in phase two (See comment 1)? Could the author provide a visual comparison between phase one and phase two classifications, like Figure 10 and Figure 11 in Carbonneau et al. (2020).

10. Page 30, Line 927: The studies based on U-Net (Baumhoer et al., 2019; Mohajerani et al., 2019; Zhang et al., 2019) focused on glacier boundaries, where this method might not generate promising results (See comments 2 and 8). Although this method could classify seven classes, I think it is not fair to conclude that this method has exceeded the U-Net based ones.

11. Page 31, Line937: The author only tests two images in summer. If the author test more images, it would be more convincing to conclude that the method could handle different illumination, weather conditions, or seasonal changes.

12. Page 34, Line 1054: I agree with the author that the combination of deep learning

methods, Google Earth Engine, and GIS software could remove the need for prior expertise in deep learning and coding (Page 33, Line 1025). But that is future work and not included in the current workflow. So I suggest removing this part from the conclusion.

Technical corrections:

Page 11, Line 311: I suppose it should be predicted classes but not image tiles that are reassembled.

Page 15, Line 396: I suppose it should be a 3D input (width, height, band).

Page 29, Line 891: Zhang et al. (2019) used TerraSAR-X images.

Page 30, Line 901: Zhang et al. (2019) and Mohajerani et al. (2019) used 2-D inputs (single-band images). Baumhoer et al. (2019) used 3-D inputs (width, height, band). The author also used 3-D input in this work (the band is just one dimension).

Reference

Baumhoer, C. A., Dietz, A. J., Kneisel, C. and Kuenzer, C.: Automated Extraction of Antarctic Glacier and Ice Shelf Fronts from Sentinel-1 Imagery Using Deep Learning, Remote Sensing, 11(21), 2529, doi:10.3390/rs11212529, 2019.

Carbonneau, P. E., Dugdale, S. J., Breckon, T. P., Dietrich, J. T., Fonstad, M. A., Miyamoto, H. and Woodget, A. S.: Adopting deep learning methods for airborne RGB fluvial scene classification, Remote Sensing of Environment, 251, 112107, doi:https://doi.org/10.1016/j.rse.2020.112107, 2020.

Mohajerani, Y., Wood, M., Velicogna, I. and Rignot, E.: Detection of Glacier Calving Margins with Convolutional Neural Networks: A Case Study, Remote Sensing, 11(1), 74, doi:10.3390/rs11010074, 2019.

Zhang, E., Liu, L. and Huang, L.: Automatically delineating the calving front of Jakobshavn Isbræ from multitemporal TerraSAR-X images: a deep learning approach, The

Cryosphere, 13(6), 1729–1741, doi:10.5194/tc-13-1729-2019, 2019.

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2020-310, 2020.