Interactive comment on “Open Source Algorithm for Detecting Sea Ice Surface Features in High Resolution Optical Imagery” by Nicholas C. Wright and Christopher M. Polashenski

Anonymous Referee #1

Received and published: 21 October 2017

The paper describes and evaluates a tool for classification of high resolution satellite- and air-borne remotely sensed imagery. The tool includes open source software and a library of training data. The idea is that this tool is a community resource that can be built upon. This is a commendable effort. Such tools are badly needed. While the segmentation and classification algorithms have been applied to sea ice classification before by Miao et al (2015), as the authors acknowledge, the development of an end-to-end software package, and perhaps more importantly, the creation of training data sets are a valuable contribution. Furthermore, the authors attempt to address the problem of sample representativeness as it pertains to deriving metrics for analysis of sea ice surface characteristics. As the authors state, this is a start of the discussion, hopefully
community resources as, developed here, can help to move this discussion along.

I heartily recommend publishing the paper in The Cryosphere. However, I do have a few issues that the authors should address in a revised version of the manuscript.

General ——— The paper is well written. The figures, in particular the captions need some work. My opinion is that figures and captions should be stand alone, such that a reader should be able to understand what each figure is without reading the text. In some cases adding a key or a description of symbols in the figure caption would achieve this goal. For example, Figure 3 has no key for the classified image.

The section that needs the most attention is


Line 63. When discussing alternative classification methods, it would be good to enumerate those methods applied to classification of sea ice. Furthermore, the authors give maximum likelihood classification as an example or an unsupervised algorithm. Maximum likelihood can be a supervised algorithm.

Section 4.1 ——— In the analysis of seasonal evolution of surface characteristics, my guess is that the field of view does not contain the same ice. The authors should make this clear or state why they think it is the same floe/ice. How fast is the ice likely to be moving at this location?

Line 433. I disagree with the statement that misclassification means that the algorithm fails to replicate human decision making. That might be the goal but one that is impossible to reach. To my mind, misclassification indicates that the algorithm doesn’t give the same answer as a human would.

Line 531. This analysis is interesting but does it apply to the image used or to all images in general. How might the result change through time or with season?

Line 558. The Central limit Theorem is a mathematical theorem complete with proof.
I wouldn’t say that it can be tested. What you are doing here is evaluating if you can predict the regional/image mean from a set of smaller samples/local means. One framework to evaluate this is hypothesis testing in which you pose the hypothesis that N sample means can predict/estimate the regional mean. This test applies the Central Limit Theorem but does not test it. This section needs to be reworked.

Line 559. The standard definition of the Central Limit Theorem is that independent variables can be added and normalized by \((X - \mu)/(\sigma/\sqrt{n})\) to yield a normal distribution \(N(0,1)\). Where \(X\) is the sample mean, \(\mu\) population mean and \(\sigma\) population standard deviation.

Figures ——— Check figure numbering

Figure 13. The cyan rectangle over the black dots make the dots look green (at least on my screen). The labels in the key need to match the description in the text. If some of the information is not discussed, I suggest removing it from the figure.