literatureInbox

Please upload citation information for important articles related to the evaluation of WSI scanners and computational pathology algorithms.

This page is a work in progress. We will process entries here and add them to a database that we will review, manage and share. In the coming days and weeks we will ask for specific fields and formats to help the process.

1- **Orting2019_arXiv_1902.09159 – A survey of crowdsourcing in medical image analysis:**

This pre-print was just published and I think it is very important to read through as we go ahead with this large-scale annotation effort. They make the most comprehensive survey or application of crowdsourcing to medical problems, up to July 2018 (so unfortunately doesn't include our recent Bioinformatics paper). It corroborates a lot of the techniques, issues and discussions in our crowdsourcing work and is, in my opinion, a phenomenal piece of work which is very relevant to what we're trying to do. You can access it at this link: https://arxiv.org/abs/1902.09159 Here are some relevant extracts: – "In this survey, we review studies applying crowdsourcing to the analysis of medical images, published prior to July 2018" ? does not include out paper. – "Alialy et al. 2018 is most similar to our survey, but … The coverage of literature is therefore much more limited than in this work." – "The most common application is histopathology/microscopy with 29% of all the papers, followed by retinal images with 15% of the papers. … Breast and heart images, which were already not well represented in the other two surveys, are almost absent in crowdsourcing studies." – "Less commonly, crowd workers … validate pre-existing annotations (14%)." – Table 1 ? very nice summary of literature on topic. – "We classify the platforms into six categories: paid commercial marketplaces such as Amazon Mechanical Turk and FigureEight (formerly known as CrowdFlower), volunteers such as Zooniverse and Volunteer Science, custom recruitment/platforms, lab participants, experts and simulation or no experiment at all. The most common choice is a commercial platform (53%)." – Preprocessing annotations:

- Filtering individuals ? remove annotations by bad annotators, as judged by number of approved annotations or score on an independent test.
- Aggregating results ? eg. majority voting by multiple observers. – Evaluating annotations:
- Compare to gold standard.
- Evaluate performance of model trained on annotations.
- Discussion in paper ? same task for expert and non-expert?
- Using inter-rater variability as a gauge or quality. – Nice recommendations to ensure quality of annotations. – Some discussion of compensation and incentives.

formatics_btz083 – Structured crowdsourcing enables convolutional segmentation of histology

er in the journal Bioinformatics. In this paper, we release one of the largest datasets for training
ms in breast cancer, including key classes like tumor, stroma, TILs (lymphocytes, plasma cells),
od vessels. We also systematically examine inter-rater concordance of various tiers of annotators and
ls can train very accurate semantic convolutional networks (pretrained VGG). You can access the paper

com/bioinformatics/advance-article/doi/10.1093/bioinformatics/btz083/5307750