Comment on hess-2021-20

Kenneth Chapman (Referee)

Referee comment on "Deep learning for automated river-level monitoring through river-camera images: an approach based on water segmentation and transfer learning" by Remy Vandaele et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-20-RC1, 2021

This paper addresses the annotation problem in machine learning with images which can be a daunting task while, at the same time, delivering. Starting with CNNs for semantic learning, inductive transfer learning is used with much smaller training sets to build accurate models that "ease the process of river level estimation from river camera images." Below we provide some over-arching, general notes about methodology that could use some clarification, followed by some minor comments for consideration.

General comments:

Two challenges for the data selected for the study seem to be 1) the use of ground surveys that "measure the topographic height of several landmarks̄ within the field-of-view of the camera?" and 2) the fact that "the number and spread of measured landmarks̄ over the camera's field-of-view was constrained to locations that were accessible during the ground survey". These challenges might be hard to overcome in real world scenarios where access is not available and the measurement of the landmark heights is a large task. Might there be alternative ways to get that information? Maybe worth ā very short discussion.̄

There was one brief note about the cameras being fixed in the images used in the paper. With a camera that is tens of meters from a scene it is common (even under the best of conditions) to get movement of the camera relative to the objects of interest in the scene. The reason for mentioning this is that it is relatively easy to test for camera movement. You might have tested for camera movement and ruled it out as a source of variation. In that case, it might be worth mentioning that in the paper. If not, it might be good to discuss how your already very good predictions might improve if that motion were taken

into consideration by using image processing techniques to find the landmarks in the scene and make adjustments for camera movement.

A brief discussion of landmark vs. area feature analysis might add to the paper. Readers will likely benefit from understanding how you weighed out the pros/cons of those approaches. Did you decide to use landmarks based on the need to ease the (already reduced) annotation task? Or did you choose it because (1) of the desire to use the "by-pixel" output of the CNN's without post processing, (2) it was just beyond the scope of this analysis because you wanted to focus on the transfer learning, and/or (3) for some other reason? Is it possible that the analysis of the classified pixels in a landmark region might yield a better predictor than just the landmarks themselves? A good deal of effort was directed at finding a heuristic to ignore outlier flood predictions. Ignoring outliers is a good objective, but analysis of landmark regions could have improved the specificity and sensitivity of your predictions (i.e., provided a more robust assessment of prediction quality). Again, this might have been beyond the scope of performing annotation of a large number of images with the transfer learning methodology, but more discussion around strengths and challenges of your approach would strengthen the paper.

 The scope of the literature review in the Introduction and Background sections is narrow and could be expanded to better connect with the broader literature on image-based stream gauging and/or extreme event detection/measurement literature. For example, the idea that non-contact image-based stream gauging techniques can also be used to flag extreme events is not mentioned. The following (and/or literature cited within) could be integrated without excessively lengthening the paper. Here are some examples from a quick Google Scholar search:

Muste, M., Fujita, I., & Hauet, A. (2008). Large-scale particle image velocimetry for measurements in riverine environments. *Water Resources Research*, *44*(4). https://doi.org/10.1029/2008WR006950

Boursicaud, R. L., Pénard, L., Hauet, A., Thollet, F., & Coz, J. L. (2016). Gauging extreme floods on YouTube: application of LSPIV to home movies for the post-event determination of stream discharges. *Hydrological Processes*, *30*(1), 90–105. https://doi.org/10.1002/hyp.10532

Eltner, A., Elias, M., Sardemann, H., & Spieler, D. (2018). Automatic Image-Based Water Stage Measurement for Long-Term Observations in Ungauged Catchments. *Water Resources Research*, *54*(12), 10,362-10,371. https://doi.org/10.1029/2018WR023913

Gilmore, T. E., Birgand, F., & Chapman, K. W. (2013). Source and magnitude of error in an inexpensive image-based water level measurement system. Journal of Hydrology, 496,

178–186. https://doi.org/10.1016/j.jhydrol.2013.05.011

Creutin, J. D., Muste, M., Bradley, A. A., Kim, S. C., & Kruger, A. (2003). River gauging using PIV techniques: a proof of concept experiment on the Iowa River. *Journal of Hydrology*, *277*(3), 182–194. https://doi.org/10.1016/S0022-1694(03)00081-7

A. A. Royem, C. K. Mui, D. R. Fuka, & M. T. Walter. (2012). Technical Note: Proposing a Low-Tech, Affordable, Accurate Stream Stage Monitoring System. *Transactions of the ASABE*, *55*(6), 2237–2242. https://doi.org/10.13031/2013.42512

Schoener, G. (2018). Time-Lapse Photography: Low-Cost, Low-Tech Alternative for Monitoring Flow Depth. *Journal of Hydrologic Engineering*, *23*(2), 06017007. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001616