

Thank you for the comments. In the following, we reply to each of your comments.

1. I understand the authors added more information in the introduction about different CNNs and other deep learning techniques. I think what is still missing is in the discussion section a paragraph where the authors discuss what other neural networks could (or could have not) achieved, based on the existing literature. This will help to underpin the authors' method as the "way to go" for automatic detection.

- We have added a subsection (4.4) that addresses the mentioned discussion.

2. It is now much clearer what the aim of this paper is from the onset.

- Thank you, we agree.

3. The paper structure is now much more readable and I appreciate the authors' effort to achieve that.

- We agree, thanks to you as well.

4. Whilst the methodology is substantially improved, some aspects are still a bit unclear to me, specifically about where and how the cameras were placed. I think a schematic would really help (even as a subfigure to figure 2 or a standalone figure), as also suggested in my previous review. Also, for Figure 2 the map is very welcome; however, the locations are impossible to identify, I recommend in-set maps to show where bridges/capturing points are located.

- We have added the coordinates to the caption of the figure, so the people interested can find the exact location. We have also added a more elaborate description of how we mounted the cameras (lines 123-125), and have added an example to the supplementary material.

Other minor line-by-line comments:

General: I note the effort done by the authors to use one type of spelling only. There are a few still to be solved (e.g., utilized, analyze, hypothesized, greyscale now turned into grayscale)

- We have gone through the text and adjusted.

Line 83: "which we attribute to the lack of uniform data" is there a reference(s) that you can use to substantiate this?

- We have added references.

I note the response from the authors about iPhones (and by extent to many other brands) which is perfectly fine. However, can they add the information provided in the text (i.e., that the phones used in the study were those only available to the authors, so there could be uncertainty about other brands and processing software)?

- We have added that indeed these phones were the phones available to the authors.

Line 92: CNN and YOLO have already been defined above, so no need to re-instate the acronym

- We have deleted the re-instatements of the acronyms.

Line 150: “a lot of”, this is not formal scientific writing

- We have removed the term ‘a lot of’.

Line 150-151: “Naturally, the wood floats and moves in a flow or is deposited or trapped by an obstacle (i.e., river bank, boulders, trees), some videos contain minutes” sentence needs rephrasing

- We have rephrased the sentence to be more clear.

The authors mention that they credited the images used for the dataset, but this is not reported in any of the figures or in the text, please correct.

- Where we were not the data acquirer/creator, we have added the credits.

Line 638: whilst I thoroughly appreciate the authors’ encouragement for this reviewer to get a promotion to full professorship, I am still a “Dr”

- We have changed the title, next time we might have to write ‘Prof.’.

References: please note the typo in “Panici (2021)” and in-text citation

- We have corrected it, honest mistake.