Dear Reviewer,

We have structured our response in a Q&A form. In each case, your Question/Comment comes first, and then it is followed by our Answer.

1. General remarks

Question/Comment from Reviewer: Paper is in general well structured, and methods are clearly presented. Results of the analyses with proper discussion could be improved.

Authors have done significant work, processing large volumes of the video data manually, and using wavelet and AI methods. Since extensive data preparation and analyses are required for these methods, authors have decided to present the results as bar plots over the sampling points or over the boat transect. This way, majority of the data processing remains hidden and readers have little insight into the detailed functioning of each automated method. The section of the paper dedicated to results and comparison of the methods (4.2.) contains only 4 paragraphs. In my opinion, the results should be discussed in more detail as they are key for validation of your approach.

I suggest that authors additionally present the complete results of the data processing (on an image-by-image scale) in form of frequency distribution curve. This way more metrics about noise in the results would be given to complement the final result.

Overall, presentation of the results (figures and graphs) could be upgraded to match the quality of the conducted research.

Answer: *Thank you for your comment! We followed your suggestion.*

Primarily:

- *We have introduced Fig. 22 and Fig 23. Explanation and reasoning for these new figures are also given: Line 21 – 24 (Abstract), Line 552 – 575; Line 642 – 645.*

Secondarily:

- *To better show the performance of the AI algorithm we have included Fig. 8 (i.e., raw AI results of a survey in one of the sections; before applying moving-average). We have also included explanation on this: Line 400 – 413.*
- *Even though due to the example of Fig. 8, we use the moving-averaged results of the AI when compared to wavelet and physical sieving, we decided to add the AI detection for the specific images taken in the sampling points (basically the AI values from before moving-average in the given point). See: Fig. 14, Fig. 21, Fig. A5, Fig. B5, Fig. C5. We believe these would give extra insight on the current performance and shortcomings of the presented method and help the readers.*

Question: Similarly, authors in their description of the results mention advantages and disadvantages for each method, such as bed armoring effect, sand detection, isolated gravel patches, etc., but this remains described only as a sentence and lacks visual supplement that would provide concrete sources of uncertainties which are crucial for alleviating the shortcomings of the presented approach and maximize its use. This is especially noticeable for wavelet method that has related data in the main manuscript body.

A: *Thank you for your comment! We have introduced visual supplements by inserting Fig. 10, Fig. 12 and Fig. 19. These include bed armouring, different sediment patches and the wavelet method for both cases. We intended to choose the most telling and prominent sample points for each case.*

Q: Authors propose their setup to be moored from the boat on a line and weighted down, assuming that the setup remains horizontal throughout the deployment. However, depending on the line length and drag influence on the setup geometry (which is highly irregular), it is inevitable that setup will be tilted during measurement. Authors haven't presented any details on how this affects the video data, which might be crucial since there are no reference points on the bed.

A: *Thank you for your comment! We have added some explanation in the manuscript, in ch. 4.3. (Line 592 - 606).*

Q: Lastly, authors have addressed implementation challenges in the section 4.3., which will be helpful in further development of this approach. However, challenges they have detected are result of the selected approach, not universal and therefore need to be addressed accordingly. E.g. vessel speed lower than 0.2 m/s is claimed not feasible for straight transect over the river. Since the proposed method aims to cover broader river section and result in 2D maps, why is it important to transect the river? Maybe longitudinal approach would allow for lower velocities while covering the same area? Implementation challenges should be addressed universally and shortcomings of your selected approach (boat type, camera type, illumination, transects, etc.) should be provided with discussion of the data. This is one of the reasons that this section contains repetitive text from earlier, and again some of the findings are repeated in the following section 4.4. Novelty and future work.

A: *Thank you for your comment! Relating to the vessel speed and longitudinal survey paths, we have added some explanation in the manuscript, in ch. 4.3. (Line 580 – 592). Regarding the repetition of illumination, we have removed this part from ch. 4.3. and inserted it at ch. 4.2, where the problem of illumination first came up and is visually supplemented (Line 517 - 523). We did our best to remove/separate the specific challenges from ch. 4.3. and added them to the section where the setup is introduced, ch. 3.2. (Line 221 – 231).*

2. Specific remarks

Q: Keywords are relatively general and do not offer information of specific contents of the paper. I suggest to drop the keywords "rivers", "sedimentology", "underwater" and "mapping", and use the following keywords for distinction instead: "riverbed texture" and "underwater mapping" (or something along these lines).

A*: Thank you for your comment! We complied with your suggestion and modified the keywords to:*

riverbed texture, underwater mapping, sediment classes, Artificial Intelligence, Deep Learning, image-based.

Q: There are several issues with referencing approach used in the paper:

Reference "Török et al., 2017" is not appropriate to be representative of bed armour development since the paper focuses on CFD modelling.

A: *Thank you for your comment! We have replaced it with Ferdowski et al., 2017. (Ferdowsi, B., Ortiz, C. P., Houssais, M., & Jerolmack, D. J. (2017). Riverbed armouring as a granular segregation phenomenon. Nature Communications 2017 8:1, 8(1), 1–10. https://doi.org/10.1038/s41467-017-01681-3)*

Q: On ln43 several references from the same author are used where one would be sufficient (Church et al., 1987; Wolcott and Church, 1991; Rice and Church, 1998; USDA, 2007)

A: *Thank you for your comment! We have removed Church et al., 1987 and Rice and Church, 1998.*

Q: On more occasions there are several references grouped together, masking specific relevance for each of them - see ln48, ln73, ln83, ln86, ln88, etc. I suggest that in cases where all references support the same statement select the ones most representative, and in cases where more statements are given you connect the reference with each of them (e.g. ln86)

A: *Thank you for your comment! We have removed several references and left the most representative ones.*

Q: Reference "Kellerhals and Bray, 1971" is used 3 times in the paper, always grouped with several others and twice with Adams. I suggest that you use it where it is most appropriate (ln50) and drop it for general statements (ln40 and ln82)

A: *Thank you for your comment! We have followed your suggestions.*

Q: In the goal of the paper contribution is highlighted as "…through improved (continuous, quick, covering larger areas) data collection.". I would suggest that you rephrase this into "…through more extensive data collection." Since it is hard to argue that method is:

…continuous (data is collected over single in situ survey)

…quick (single point collection is quicker, both for collection and processing)

…covering large area (what is large area, and does this conflict with two previous advantages)

I suggest that these advantages authors address in the discussion and drop from the goal itself since they are not straightforward.

A: *Thank you for your comment! We have followed your suggestions.*

Q: Considering the size of the paper and volume of the work conducted, the goal of the paper is a bit short. I suggest that you expand the goal of the paper to reflect the specific contribution to the field.

A: *Thank you for your comment! We have followed your suggestion and extended the description of our goal in ch. 1. (Line 152 – 161).*

Q: Methodology section briefly describes the three locations of Danube River where data was collected, introducing flow rate data, SSC, etc. This data might be useful for someone familiar with the river, which most of the readers won't be. Please put the presented data in context – e.g. provide complementary duration data, long-term average data, etc. which would allow estimation of conditions under which surveys were performed (low flow, average flow, flood).

A: Thank you for your suggestion! We have followed your suggestion and added $Q_{1\%}$. We also attempted to clarify the Table.

Q: Part of the Methodology focusing on the equipment lacks information that would help understand the data quality, in the context of the maintaining the setup distance from the bed. Please expand the current description with the data about the desired height above the bed and what does it depend on (supposably illumination of the FOV). Similarly, you initially state that size-reference wasn't used in the images (ln253), and after offer contrary statement that laser pointers were used to provide scale (ln294). Probably laser pointers do not offer constant distance due to the bed irregularities, but clarification of the way they were used would be helpful.

A: *Thank you for your comment! We have expanded the camera - riverbed distance related description: (Line 580 – 592).*

In ln173, it was originally said the training was unscaled for the AI. Ln371 - 384 was about the wavelet method, which inherently requires scaling. The wavelet method was only applied in the physical sampling points, so the laser were only used in these points. As we had some problems with them, we did not use the lasers constantly, during the transversal surveys, only when we reached over a given physical sampling point. We have clarified these statements and updated the manuscript (Line 341 – 346).

Q: In the Section 4.1, explaining the training of Deep Learning authors present example of erroneous particle detection of the user. Although this is good and informative example, in my opinion it shouldn't be presented as "ground truth" and rather included as sidenote explaining that training data needs to be carefully selected since you noticed errors in user judgement (n.b., who was the user – one of the authors or trained personnel?)

A: *Thank you for your comment! We have modified the Figure and separated the case of erroneous detection, along with its description. Fig. 5 and Fig. 6, Line 366 – 372.*

N.b. The user was trained personnel. We have now also included this in the text. Line 296.

3. Technical remarks

Q: On ln 31 “fluvial navigation” -> more appropriate would be “fairway placement”, but the connection is loose so I suggest that you replace it with use more relevant to the grain sire instead.

A: *Thank you for your comment! We have followed your suggestion and looked for the most appropriate expression. We have decided with “inland waterway transport”.*

Q: On ln 31 “riverbed structure” -> unclear, does this represent morphology?

A: *Thank you for your comment! We have corrected the sentence to:*

“Knowledge of riverbed (~~structure~~) morphology and (~~grain~~) sediment composition is therefore of major importance in river hydromorphology.”

Q: “Riverbed” and “river bed” are used interchangeably throughout the manuscript. Please proofread the manuscript.

A: *Thank you for your comment! We have corrected it.*

Q: Goal (aim) of the paper is combined with Introduction, making it indistinctive. Please separate the aim of the paper into separate paragraph.

A: *Thank you for your comment! We have followed your suggestion! Line 152 – 161.*

Q: Figure 1 is very simple and lacks details (Danube is not highlighted and therefore indistinctive from other rivers in the figure).

A: *Thank you for your comment! We have updated the figure and highlighted the Hungarian Section of the Danube.*

Q: Can the background orthophoto data be added on the Figures 2 and 3?

A: *Thank you for your comment! We have followed your suggestion! We also merged them into one figure.*

Q: Names of the sections (transects) are very hard to follow since they do not follow any logical order (I suppose they do for the authors, but I suggest that you rename them to achieve clarity for the readers).

A: *Thank you for your comment! We have followed your suggestion!*

Q: “Streamlined weight” -> isokinetic suspended sediment sampler?

A: Thank you for your comment! Yes. We have added the explanation:

> *“ (…) streamlined weight (originally used as an isokinetic suspended sediment sampler) (…)”. Line 211 – 212.*

Q: Methodology section would benefit from added flow chart (process chart) since video decomposition and enhancement is carried out through several steps.

A: *Thank you for your comment! We have added a flow chart in Fig. 4.*

Citation: https://doi.org/10.5194/esurf-2022-56-RC1