

Review of Automated riverbed material analysis using Deep Learning on underwater images –
Ermilov et al.

General

R: The use of figures within the manuscript is ineffective at times. In my opinion the manuscript has too many figures (20+) which detracts from the important figures in the manuscript and reduces readability of the paper. In the line-by-line comments I have provided suggestions on how to combine existing figures or when to make use of the supplemental.

A: Thank you very much! I agreed with your insights and tried my best to follow your instructions in general!

R: The terminology used to describe the sample sites is unclear and should be changed so that it is more intuitive for the reader. A simple format (such as All-1, All-2 and BII-1, BII-2) would help readers to compare between original photos and points along the channel. The current terms (e.g. VM4, 3/4) could not be easily referred to by someone who was not involved in the study.

A: Thank you very much! I have changed the names in both the text and on the figures.

R: The Results and Discussion section (Section 4.2 currently) could be improved by not separating the results by study location and instead integrating all three locations (including those in the appendix) into a discussion about how the AI algorithm compared to the different methods. The results and discussion could then be titled 3.2. Comparison of methods with subheadings for each method type (e.g. 3.2.1 Physical samples, 3.2.2 Wavelet analysis). The authors could then compare and contrast the different methods for each location to avoid repetition. Figures 22 and 23 could also be integrated into the respective sections.

A: Thank you very much! I have followed your instructions and rearranged Section 3.2 entirely.

R: The method appears to be successful by current standards however I had some trouble understanding exactly how the accuracy scores were calculated and what was being shown in Figures 22 and 23 (though I really like these figures and think they make a great addition to the manuscript following the initial revision). For example, why are there only 11 points when many more photos were used for validation (>2000)? This should be clarified to ensure that readers understand the accuracy when using the method. The manuscript may benefit from a plot demonstrating the accuracy for all tested images (e.g. gravel content vs IoU) as this may show where the algorithm performs well and less well.

A: Thank you for your comment! I have added more explanation to the text (e.g. Line 16 – 17; Line 499 – 510 and Table 2).

Line by line comments

R: Title suggestion: Automated riverbed composition (?) analysis using Deep Learning on underwater images

A: Thank you very much! I have changed it accordingly.

R: Line 10: Remove indeed

A: Thank you very much! I have changed it accordingly.

R: Line 12: Add by after "...overcome this issue..."

A: Thank you very much! I have changed it accordingly.

R: Line 22 and 23: After reading the manuscript I am unsure what these percentages are based on, is it

A: Thank you very much! I have changed it accordingly.

R: 64% of 11 photos? It would be great to add an n value here to show what proportion of the dataset as been used here.

A: Thank you very much! I have changed the manuscript accordingly. I have added Table 2, for example. Furthermore, I tried my best to further ellaborate on it in the text. Line 16-17, Line 23-26, Line 485, Line 496-497, Line 508- 512

Introduction

R: Lines 31 to 40: Whilst the content in this paragraph was well placed and relevant for the manuscript, I think the paragraph could be written more clearly. For example, without the use of etc. and with references after the examples (e.g. sentence on line 37 and 38 did not give reference but gave lots of examples).

A: Thank you very much! I agreed with your insights and tried my best to follow your instructions! Line 41-43

R: Line 38: It would be helpful to break down sediment composition early in the manuscript, for example "sediment composition (sand and gravel content)" to make clear what you are measuring.

A: Thank you very much! I have changed it accordingly. Line 44

R: Line 59: I would use “Section” as opposed to “Chapter” in a manuscript for this journal, but a reference at the end of the following sentence would be sufficient here. Section 2 has also been removed following the previous review stage so the reference to that section no longer fits here.

A: Thank you very much! I have changed the manuscript accordingly (from Chapter to Section everywhere, also adjusted the Section numbers).

R: Lines 69 to 71: Sorry this is not my area of expertise, so I found the sentence starting “Researchers found that...” unclear. Is there a specific article which “researchers” is used to refer to? And what are the specific coefficients used to obtain, i.e. why convert the signal strength?

A: Thank you very much! I have changed the manuscript accordingly. Line 76 - 99

R: Line 75: Why could gravel not be distinguished strongly from sand?

A: Thank you very much! I have changed the manuscript accordingly. Line 84 – 86

R: Line 80: picture to image? Line 83: “Both” as opposed to “The above-mentioned”

A: Thank you very much! I have changed the manuscript accordingly.

R: Lines 152 to 161: I would specifically state that that this method will be used to obtain pebble and sand fraction as that is really exciting and should not be understated.

A: Thank you very much! I have changed the manuscript accordingly. Line 178 – 179

R: Line 163: earlier is repeated in the sentence, remove one. Line 172: The reference to the third chapter is no longer valid as the second section (originally literature review has been combined with the introduction). I would also use Section here not Chapter. I don’t think this paragraph is necessary for the readability of the manuscript but am happy to leave to the authors discretion.

A: Thank you very much! I have changed the manuscript accordingly. I have decided to leave the paragraph in since if I remember correctly it was asked by the reviewers in the Public Discussion phase. I have corrected the reference and removed the repeated word.

Methods

R: Line 179: All the numbering from this point is incorrect, methods should be 2. Please check this.

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 182: “one” seems vague, does this mean video? Or dataset? Or collection of videos/transects?
Line 182: “the second one”, can remove “one” here, once the first “one” is replaced with a more specific term.

A: Thank you very much! I have changed the manuscript accordingly. Line 197 – 200

R: Figure 1: I like how the Danube has been emphasised.

A: Thank you for your feedback!

R: Line 191: Change “similarly to” to “and”. Line 200: I would move SSC_{survey} to in brackets following the first mention of suspended sediment concentrations in the previous line (199). Figure 2: Great figure, really clear. I would add A and B to the different site maps so that each site can be referred to in the text. Line 284: Videos might work better here as opposed to footages?

A: Thank you very much! I have replaced the mentioned words, added SSC to the brackets (Line 215) and updated Figure 2 accordingly.

Attention! Seems like Table 1 was divided between two pages when I saved my final version of the PDF. Terribly sorry for that! I will correct it in the next review phase!

R: Page 11, Second Paragraph: Line numbers do not work here, potentially due to figure placement.

A: Thank you for your feedback! Yes, it happens because of the long figure.

R: In the sentence “These steps were followed by the annotation, where we distinguished ten classes”, the ten classes should be written as a list. Alternatively, they could be included in Figure 4, in which case the authors could refer to the figure. A space is also needed between “the” and “4” further in the paragraph in the sentence “In total, a...”.

A: Thank you very much! I have changed the manuscript accordingly. I have updated Figure 4 with the classes. However, I insisted on mentioning them in the text as well, if that is not a problem I would keep it there too!

R: Figure 4: As someone not familiar with Deep Learning algorithms, I appreciated the step-by-step flow chart. I had a few suggestions that might help readers unfamiliar to better understand the process.

Data creation box:

- Change "From underwater videos" to "Uses underwater videos".
- Change "fix" to "fixed"

White balance upgrade

- I found the description in the first bullet point unclear and wondered if the step could be clarified.

Data annotation

- You could list the ten classes here if not in the text.

Visualization and analysis

- I think the overall accuracy is also referred to in the text, however I think more information should be given in the box about how this is calculated. Is this the accuracy as in number of gravel, cobble and sand pixels in each image when manually mapped and when using AI? Please clarify.

A: Thank you very much! I have changed the manuscript accordingly. I have updated Figure 4.

R: Line 353: Remove "by the authors of present manuscript".

A: Thank you very much! I have changed the manuscript accordingly.

Results and discussion

R: Line 355: 4 to 3.

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 357: I think this paragraph could be rewritten in parts. I was unclear about whether the datasets being discussed were the training dataset or the validation photos. The first sentence in the paragraph could be used to clarify this by providing a simple, general statement about how successful the method was at recording the grain size fractions in the training/validation dataset.

A: Thank you very much! I have changed the manuscript accordingly. Line 378 - 385

R: Line 358: Is the validation set the grain size fractions obtained using the wavelet method? Please clarify.

A: Thank you very much! I have changed the manuscript accordingly. It is better to think of DL validation as a self-validation. It resembles a student who learnt something at home (training process, with 80% of the 14,784 human-annotated images prepared by the teacher (us). During validation, the student writes a test and receives 2957 raw images (20% of the 14,784 images) and has to classify them based on what he/she learnt. The teacher (us) of course has the answers (this 2957 images were annotated during annotation) and can compare the student's answers to this solution-sheet. Relevant parts of the manuscript: Figure 4 – Training box; Line 342 – 344, Line 378 – 382

R: Line 360: “over-all” to “overall.”

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 361: Please add the n value for the 96% accuracy, e.g. is $n = 2957$ – the number of photos? Or is it the number of photos used for training?

A: Thank you very much! I have changed the manuscript accordingly. Line 382 – 385

R: Figure 5: To avoid a lot of figures in the manuscript, which dilute the message, I would suggest combining Figures 5 and 6 and using A, B, C, D and E to separate each image row.

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 387: Comparison as opposed to intercomparison. Please see general comments for a suggestion on how to change this section of the manuscript.

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 388: I really like this paragraph. Perhaps this would be a good opportunity to introduce Figures 22 and 23.

A: Thank you very much! I have changed the manuscript accordingly. Figure 9 and Figure 13 has been introduced in this Section, I also rearranged the whole Section.

Attention! Seems like some of the figure-caption style was not automatised when I changed the manuscript and it did not get my attention (Figure 9, Figure 12. Again, I'm sorry for that... I will correct them in the next review phase.

R: Line 404: Is the 15 m smoothing transferable across sites/equipment or is there a specific reason for its use here?

A: Thank you very much! I have changed the manuscript accordingly. Line 428 – 433

R: Line 405: “are the ones being compared” to “are used to compare...”

A: Thank you very much! I have changed the manuscript accordingly.

R: Figure 7: Combine Figures 7 and 9. Both are really useful, and I think combining the plots as a two-panel figure would be more effective and allow for each reference between the sampling location and the sediment composition measured. The plots could also be on the same scale so that comparison is simple. As a side note, Figure 9 is really effective.

A: Thank you very much! I have changed the manuscript accordingly.

R: Figure 8: Whilst this is an interesting figure, particularly as this is a methods-based paper, I think it would be better placed in the supplemental, which is an option in ESurf and referred to in the text as “Figure SX in the Supplement”.

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 429: Change “an” to “and”

Line 430: Change “eye” to “visual”

Line 438: Change “with” to “by”.

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 446 and Figure 11: Move to supplement.

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 447 to 448: Add “we broke the surface amour to showcase the presence of the underlying finer fractions” to the figure caption and remove from the main text.

A: Thank you very much! I have changed the manuscript accordingly.

R: Figures 13 and 14: Can be combined and moved the supplemental. Line 481: Remove “again”. Figure 15 and 16: Combine but keep within main text, see comments on Figure 7 and naming of samples at the start of this review. Line 510: Use 19b as opposed to 19/b when referring to figure. Figure 17: Move to supplemental. Figure 20: These figures could be combined with the earlier. Figure

5 (and Figure 6, see earlier comment) as two additional rows to the plot and then referred to using letters. Figure 21: Move to supplement.

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 539: In this journal, I think the "Appendix" is considered the "Supplemental" so this should be changed.

A: Thank you very much! I checked the Formatting Requirements on ESurf and there they ask it to be Appendix. (<https://www.earth-surface-dynamics.net/submission.html#templates>) at Manuscript composition: 7. Appendices

R: Line 550: Is "efficiency" the correct term here? Sorry I am not familiar with vocabulary used to describe Deep Learning algorithms. "Applicability" came to mind as the sentence appeared to be detailing which photos could be used with this method, but "efficiency" may work within the context. Line 553: Change "belonged" to "belonging".

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 555: What does the term points refer to in this paragraph, it is unclear, photos? Study locations? Why are only 11 included if so? And how many 'points' were there in total? Add $n = xx$. Use of a clear sample naming procedure would also improve clarity in this paragraph.

A: Thank you very much! I have changed the manuscript accordingly. Line 484 – 507

R: Figure 22: I really like this plot and think it is a good addition to the paper. However, I find it difficult to fully understand what is being shown. Are the plots showing the relative proportions of each size fraction based on the 11 points (photos or sites or both?) when using the AI detection in this paper and physically sampling from the riverbed? I would advise splitting this figure into three figures (A, B and C) and using the shape/colour of the points to represent the different samples as it is unclear what has been plotted. The same could also be done for Figure 23. I think the figure caption could include more detail to help guide the readers and the axes titles could also be clearer. For example, if three figures are used the author could plot "Physical samples gravel fraction (%)". Also why only select the most comparable images? Is there benefit to sharing the full dataset? Maybe you could include samples with bed armouring and thin sand layers in a grey colour/different shape so readers can see the impact of sampling in these locations. Line 567: Similarly, the use of points needs to be clarified earlier or changed.

A: Thank you very much! I have changed the manuscript accordingly. I have improved Figure 9 and Figure 13 by adding the name of sampling points next to the data points where they are taken from. I also extended their figure captions. I also further elaborated on it in the text: Line 484 – 507, Line 514 – 517, Line 551 – 557

R: Line 567: Change “while neither...” to “unlike the AI and physical samples”. And remove “did so”.

A: Thank you very much! I have changed the manuscript accordingly.

R: Figure 23: I think three figures (as explained for Figure 22) would also be interesting here and show if there are any differences across study sites. The x axis also states physical but the caption says wavelet analysis, please check this.

A: Thank you very much! Similarly to Figure 9, I wished to help the case by including the name of the sampling points at their respective data points.

R: Line 596: Move this statement to the methods section as well as kept here where it is being discussed. Line 622: remote “i.e.”. Line 625: The sentence starting “However, the latter...” is quite vague and I think it would be good to clarify. Line 640: Remove “,” after size.

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 642: Echoing the comments made on the abstract, is this just the 11 points? Are these transects? I think an n value for this statistic relative to the total number of photos tested would be useful.

A: Thank you very much! I have changed the manuscript accordingly. Please, see my earlier responses. Line 484 – 507, Line 514 – 517, Line 551 – 557 and Figure 9, Figure 13.

R: Line 643: typo “nut” to “not”.

Line 643: “surpassed” to “surpass”.

Line 665: Chapter to Section

A: Thank you very much! I have changed the manuscript accordingly.

R: Line 62: Thank you for breaking down the previous research on alternative methodologies for measuring grain size. A table with a breakdown of the previous research would be very helpful for readers of the manuscript who are not familiar with the past work. The table could include the limitations outlined in bullet points (lines 89 to 100). Line 87: If a table is created as suggested in the Line 62 comment above, I think a reference to the table here would work well and then the bullet points could be reduced to a sentence/found in the table.

A: Thank you very much! I am sorry, I was not sure what you wished to see here exactly. Can I ask for further clarification? The bulletpoints of limitations only refer to the image-based alternative

approaches. Hence the Table should include these studies in my opinion only. Should it be a 2-row table, where first row is the Referred study, and the second is its limitation? Or maybe something else? Thank you in advance for the further clarification!

R: Line 361: Is this the percent of pixels correctly identified or the number of photos where the method was successful? Please clarify. A plot showing the variation in the accuracy value and IoU value per photo would be interesting and may more clearly present areas where the algorithm is more and less (e.g. poorer quality images, high SSC?) accurate.

A: Thank you very much! I have changed the manuscript accordingly. Line 378 – 388. Regarding the preparation of this kind of figure I was not able to compile:

Unfortunately the selection of the validation set (2957 images) is randomized, and when I prepared this I did not think of preparing a figure like this. It further adds to the difficulties that they are not even whole images but augmented ones (see Augmentation) and I had trouble with identifying each of these 2957 images point locations in time for the submission. On the other hand, these points (due to random selection) would not clearly show areas, but would rather act as a point cloud between the cross-sections of Site C and proportion of Site A. Maybe some of them even belonging to the same image. I feel doubtful if it would show some sort of correlation with the SSC, because it also includes the effect of the camera getting too far away (bed forms, channel shape) and the Deep Learning algorithm's inefficiency/efficiency. Anyhow, if required, I will definitely try my best and continue the preparation of this figure for the next review phase.