University of Stuttgart
Germany

Institute of Applied Mechanics (CE)

Prof. Dr.-Ing. Holger Steeb

Pfaffenwaldring 7
70569 Stuttgart
T +49 711 685-66346
F +49 711 685-66347

steeb@mechbau.uni-stuttgart.de
www.mechbau.uni-stuttgart.de/ls2

University of Stuttgart • Pfaffenwaldring 7 • 70569 Stuttgart • Germany
Prof. Dr.-Ing. Holger Steeb / Institute of Applied Mechanics (CE)

08.08.2022

“Detecting micro fractures: A comprehensive comparison of conventional and machine-learning based segmentation methods”

Response Letter

Dear Reviewers,

First and foremost, we would like to thank again both reviewers for their very thorough assessment of our submission. In the following, we have summarized the remarks of Reviewer #2 with a corresponding statement (blue).

Reviewer #2

1. An introductory paragraph in section 2 might help to separate the 3 subsections here: sample preparation, noise reduction and segmentation methods.

-> We appreciate the reviewer's comment and we have now explicitly separated the three sections. Section 2.1 refers to sample preparation, section 2.2 to noise reduction and section 2.3 to segmentation methods in the revised article.

2. Section 2.2: Could the authors please provide information regarding the parameters employed for applying the Non-local mean methods? That is, shape and diameter of the searching windows, similarity value, etc.

-> We have now added the relevant information in the revised manuscript as follows:

"The adopted input parameters of the filter in this study were spatial standard deviation: 5, intensity standard deviation: 0.2, size of search window: 10 and size of local neighborhood: 3."

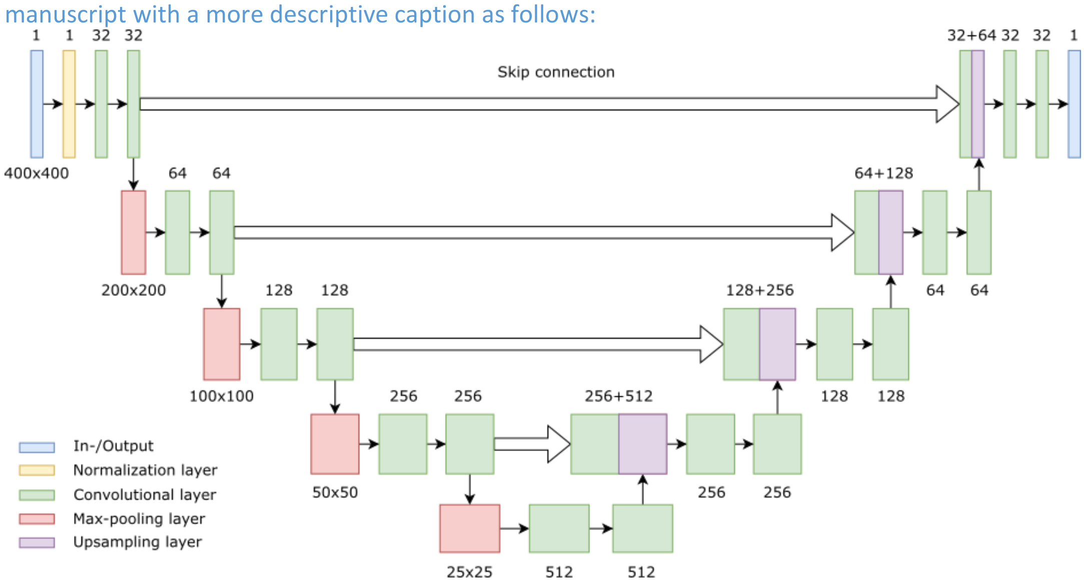
3. Sections 2.3.1/2/3. Given that the μ-CT images of the thermally treated Carrara marble sample are open to be access by the community, could the authors provide with the parameters employed to segment the data? Like this, any reader would be able to reproduce the results and work from there on improving the methods.

-> The code used for the segmentation, along with the relevant parameters, has been made available via the link in the "code availability" section. It is addressed in our revised manuscript as follows:

"The implemented code with detailed parameters is available via the link in the "code availability" section."

4. Figure 4: The legend item "Upsampling layer" (purple) is not in the figure. Also, the caption could be a bit more descriptive of the sketched workflow.

-> We thank the reviewer for their comment. The figure has now been modified in the revised manuscript with a more descriptive caption as follows:



Caption : *"The schematic of the used 2D u-net model. The input image was down-/upsampled with the help of Max-pooling and Upsampling layers. The features were extracted by convolutional layers. Those extracted features during down-sampling were concatenated to corresponding up-sampled features (skip connection). "*

5. Sections 2.3.3: The authors might consider improving the explanation of the U-net model. It is the most efficient of the analyzed methods and I found this section a bit difficult to follow.

-> We appreciate the reviewer's comments. For the sake of not making the main article too extensive, we decided to keep the section short, which was obviously not the best option. In the revised article we have now moved the paragraphs below to the section "The U-net model" from "Appendix E: Splitting, Training, Merging for the U-net model". We trust that we now provide adequate information for a better understanding of the model.

"The model makes use of repeating down-scaling of the input image with the help of max-pooling layers, and up-scaling with a de-convolutional layer. Additionally, before and after each of the up-/down-scaling layers, the convolutional layers which extract the feature maps were used with an activation function which introduced a non-linearity into the model. Each of the extracted and down-scaled features were concatenated to the same size of the up-scaled features, in order to force the output pixels to be located at reasonable locations (see Figure 4)."

"For the sake of accuracy, the data augmentation technique was applied on the training data-set. This allowed us to enrich the training data-set by employing a modification to the data; thus, the model could be trained with sufficient data of different variations. Consequently, the model would be trained with more trainable data. This contributed to the prevention of overfitting, which made the model capable of dealing only with a specific case. In our application, we varied the brightness of the training data. Thus, the model was able to be trained by data with variation. This was necessary to get good predictions from all cropped tiles."

6. Have the authors tried to resample the raw images by reducing the voxel size? This might help to segment fractures thinner than 2 microns.

-> We appreciate the reviewer's comment. Given the limitation of our hardware, we were not able to physically resolve features smaller than 2um/voxel. A possible resampling option could potentially increase the image resolution artificially by using interpolation. With the help of this approach, we anticipate some fractures will be captured as thinner than 2 microns. However, we believe that a further investigation would be necessary in order to trust such captured thinner fractures since those are based on artificial intensities. Besides, the size of data would become significantly larger as the resolution increases. This will directly affect the computation time and required amount of memory for computation.

7. I'm not sure if this makes sense, but have the authors tried to re-apply the U-net model using the U-net model output as the new GT for training a new model?

-> We appreciate the reviewer's comments, and this had indeed been one of our thoughts too. However, we strongly believe that this approach would cause the mode-collapse problem which oversimplifies the training. If we repeatedly feed the model with its own products, the model rather converges to learning the simplest features instead of learning various and generalized features in the ground truth. This problem was reported and known in the GAN (Generative Adversarial Network) type of model (Li et al.(2019))

Sincerely Yours,

Dongwon Lee, Nikolaos Karadimitriou, Matthias Ruf & Holger Steeb

References

- Li, Y., Xiao, N., & Ouyang, W. (2019). Improved generative adversarial networks with reconstruction loss. *Neurocomputing*, *323*(5), 363-372. https://doi.org/10.1016/j.neucom.2018.10.014