Instream wood detection using the YOLOv4 algorithm on aerial images of the Spöl river

Florent Rouge
MSc Aquatic Sciences
Université de Lausanne

Abstract

This Machine Learning project is about detecting in-stream wood in a mountain torrent subject to artificial flooding on aerial images. It has been conducted with the YOLOv4 algorithm using darknet framework after labelling wood pieces on LabelImg. The model thus produced has been trained on tiled pictures of the upper part of the river (70% of the total distance), and tested on the lower part of the same river (30% of the total distance).

1. Introduction

Rivers represent large and complex systems that vary and evolve over time in different aspects (Meybeck, 2003). Among the most important factors affecting the geomorphological aspect of a river are the occuring of floods. Indeed, through the disturbances they induce, they cause numerous changes in the geomorphic appearance of the river and trigger several smaller mechanisms which end up modifying its pattern to different extents (Yousefi et al., 2018). However, the floods' impact, intensity and consequences are also affected by other factors, one of which is the presence of in-stream wood (Czarnomski et al., 2008; Wohl et al., 2010).
The presence of wood in rivers has for a long time been considered dangerous and detrimental to both the good health of the river and the economy running around it, probably due to the damages provoked by wood-charged floods on human infrastructures and therefore biased human perception (Ruiz-Villanueva et al., 2018). Today, however, many studies have proven now that it is actually quite the opposite (Lester & Wright, 2009; Swanson et al., 2021; Wohl et al., 2019). Indeed, the recruiting and displacement of logs during floods end up creating dead wood accumulations which have several beneficial consequences on the geomorphology of the river, such as reducing the flow energy, diverting the flow in such places, creating islands and dry sediment accumulations behind the wood piles.
Such processes take part in the natural dynamic of the river and participate in the development of the connected ecosystems. Therefore, quantifying in-stream wood and recording its displacement due to floods is relatively important and can improve the understanding of the whole river ecosystem dynamic.
The aim of this project is to use machine learning and more precisely the YOLOv4 algorithm (You Only Look Once) to develop a model able to detect the presence and evolution through floods of in-stream wood in aerial images of the Spöl, a mountain river located in the eastern part of Switzerland (GR). This river is being flow-regulated due to hydro-power exploitation and is artificially flooded every year. This represents a great study opportunity because of the known temporality and parameters of each flood. Automatic wood detection and quantification in such a river would represent a important gain of time compared to field wood survey performed by human hands. Indeed, the other possibility to conduct these types of survey is to go in the field and manually register the location of every piece of wood, which is excessively time- & energy-consuming. Of course, a ML model will not be able to conduct such type of job as well as a human in the field would do, but it would greatly facilitate the operation and spare a lot of time to the researchers as for before-hand preparation.

2. Data used

2.1. Raw data

The dataset used for both training and testing is an orthophoto created from aerial photogrammetry of the Spöl, performed with the usage of a DJI phantom 4 drone in June 2022. The drone flew at an altitude of 80m over ground, covered the area with two perpendicular grid flights and several circular flights to obtain various points of view. The resulting images were post-processed on the Structure-from-Motion software Pix4D to produce a single 2.4km orthophoto of the whole study area as well as a digital elevation model (DEM).

2.2. Dataset subdivision

The orthophoto produced was then tiled to fasten any processing on ArcGis and the whole orthophoto has been split

Figure 1. Example of the raw dataset used for training: Aerial images of the upper part of the Spöl.

Figure 2. Spöl aerial image with every log labelled as such in bounding boxes.

in two at a 70% / 30% ratio for training / testing. The training set consisted of the upper part of the Spöl, split into 15 images and the testing set of the lower part of the river with 6 images in total. Some sections of the Spöl have been discarded from either the training or testing sets for obvious reasons as they did not include any in-stream wood, and the relatively small size of both dataset led to prefer images including wood pieces.

3. Methodology

The model used to perform wood detection is based on the YOLOv4 algorithm, which is a Convolutional Neural Network (CNN) usually recommended for object detection tasks. We used it with a Darknet framework to train the model on a dataset of aerial images of a river. The study site of this project is the Spöl river, with the corresponding dataset detailed above. The dataset was pre-processed by resizing and converting the images to RGB to make it easier to process. Then, every wood pieces located in the training set (the upper 70% of the Spöl) were labelled using bounding boxes in the 'Label-Img' software. The total number of wood pieces labelled was 395 for an average of 26.3 per image. A choice has been made to not perform much data augmentation due to the already considerable computing time with such training dataset and the limited time allowed to perform the detection.

I then followed the online tutorial of 'The AI Guy' to set the model up. First, the GPU has been enabled in the Google Colab Notebook to allow faster processing, then the darknet framework has been cloned in the local repository from Alexey AB's one, modified accordingly to the needs of this project and built in the local repository. After defining a few helper functions that will be used during training, the Google Drive containing training and testing images as well as other necessary documents for YOLOv4 to be configured and work correctly have been linked to the repository and the training files have been prepared for training. The last step before training was to donwload pre-trained YOLOv4 weights for the convolutional layers from Alexey AB's repository and load them into the framework. Finally, the hyperparameters were defined for configuration as follow: a batch size of 64, a learning rate of 0.001 & an average of 18 filters per layer. The other hyperparameters were automatically set.
The model training could then start, it has been done for 4'500 iterations with a total duration of 14 hours with a premium GPU enabled to accelerate the processing. The weights were saved every 100 iterations in a backup to keep up with the training even if the Google Colab Notebook crashed or had to be stopped manually.
The average loss chart displayed on figure 3 has been taken from Christophe's model training as mine had to be interrupted and started again several times, restarting a new chart at each pause. However, even though our hyperparameters choice was a little bit different, both models work relatively similarly and so is representative for my model as well. As the training weights were saved each 1000 iterations and due to the chart observable on figure 3, the weights saved at 3000 iterations were chosen for testing as it represents the most efficient point between decreasing the average loss of the model and over-fitting our training set. The model was then tested on the testing dataset and metrics like the mean average precision (mAP) and the average precision (AP) were calculated to fine-tune the hyperparameters. Both the Google Drive folder containing the configuration documents and both training & testing dataset and the Google Colab repository are available by clicking on the links.

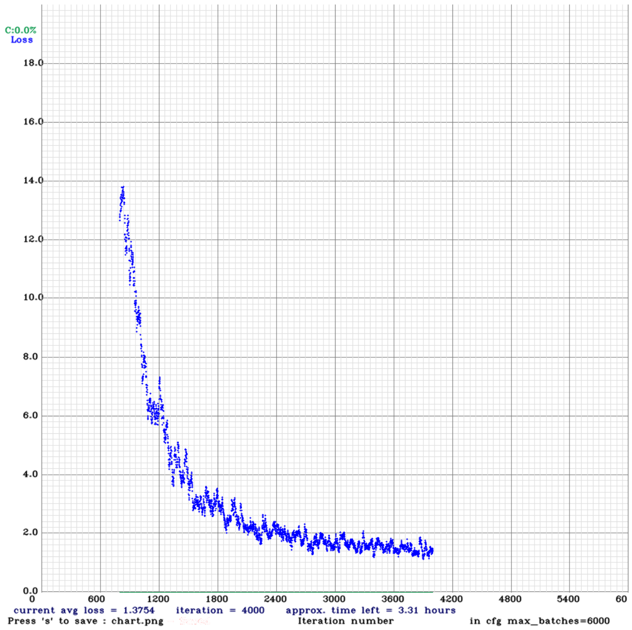

Figure 3. Average loss chart of an equivalent model during training from 800 to 4000 iterations.

4. Results

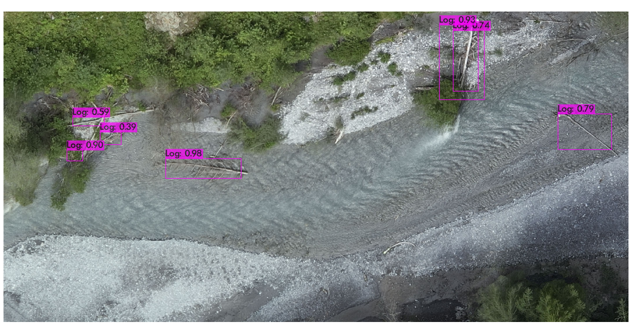

Figure 4. Predictions produced by the model after testing.

The model successfully detected most of the large pieces of wood located in the river channel. An example of the resulting wood detection can be seen on figure 4. We can observe a little bias in the log size estimation, expressed by a low intersection over union (IoU) value. We can also notice the presence of a few false negatives in the channel, meaning logs undetected by the model. Overall however, the model was considered successful and the results it produced can be used for further studies.

5. Discussion

The YOLOv4 algorithm proved to be a good choice of algorithm for our model due to its high performance in object detection tasks. However, its setting up proved to be relatively inappropriate for beginner machine learners. The existence of online tutorials is however life-saving and allowed to get good results in the end.

The relatively low IoU value indicates a few bias in the model, affecting the mAP as well, but the results were overall satisfying and after human verification, they can be used for following projects. The biggest problem being the poor data augmentation of the dataset could have been avoided by performing a more thorough one with, for example, executing random flips, applying distortions or modifying the exposure of the training images. However, due to its time-consuming nature and the limited time we had to perform the training part of the model, combined with the limited GPU calculation units available in Google Colab, it was not reasonable for us to perform such data augmentation. It would have made the training excessively long and we could not allow that.

Thus, for further usage of a model as such, a super-computer will be used to avoid this type of inconvenience and accelerate significantly the time required for training. The training/testing split decision was also limited as only one aerial orthophoto of the Spöl was available to run it at the time. More are now accessible, which would represent much larger training & testing datasets, which would also get better results in the end. The choice of not using a validation set depended also from the overall data available, and this problem could be avoided if the dataset are larger as well.

Finally, the randomisation introduced in the data segmentation part of the 'AI guy's tutorial which was deactivated for our model training will be enabled again in further projects and should help us to get a better model overall, with higher IoU and mAP values.

6. Conclusion

The YOLOv4 algorithm, and its qualities as a CNN proved to be a good choice in order to create a model performing wood detection in aerial images of the Spöl. The obtained results were overall satisfying but a few improvements possibilities have been noted throughout the conduction of the study. First, the dataset size could have been greater with more than one orthophoto used, increasing both the training and testing set size and allowing the creation of a validation set. And secondly, a super-computer could have been used instead of the Google Colab environment, providing a more important calculation power availability, which would have allowed to perform a more thorough augmentation of the dataset and thus produced a more accurate model. However, as a first machine learning project, this proved to be satisfying and the results will be useful for further usage.

7. Acknowledgments

Special thanks to 'the AI Guy' for his online tutorial on how to run a YOLOv4 model. To Janbert Aarnink, PhD student, for his help during the project design. To Christophe Reymond, fellow master student, for his collaboration on the project. To Aldo Fornari, fellow master student, for his precious advises on a similar project he made in the past and finally to Prof. Tom Beucler and his TA team, for their advises and the help they provided throughout the development of this project.

8. References

1. Czarnomski, N. M., Dreher, D. M., Snyder, K. U., Jones, J. A., & Swanson, F. J. (2008). Dynamics of wood in stream networks of the western Cascades Range, Oregon. Canadian Journal of Forest Research, 38(8), 2236-2248.
2. Lester, R. E., & Wright, W. (2009). Reintroducing wood to streams in agricultural landscapes: changes in velocity profile, stage and erosion rates. River Research and Applications, 25(4), 376-392.
3. Meybeck, M. (2003). Global analysis of river systems: from Earth system controls to Anthropocene syndromes. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358(1440), 1935-1955.
4. Ruiz-Villanueva, V., Díez-Herrero, A., García, J. A., Ollero, A., Piégay, H., & Stoffel, M. (2018). Does the public's negative perception towards wood in rivers relate to recent impact of flooding experiencing?. Science of the Total Environment, 635, 294-307.
5. Swanson, F. J., Gregory, S. V., Iroumé, A., Ruiz-Villanueva, V., & Wohl, E. (2021). Reflections on the history of research on large wood in rivers. Earth Surface Processes and Landforms, 46(1), 55-66.
6. Yousefi, S., Mirzaee, S., Keesstra, S., Surian, N., Pourghasemi, H. R., Zakizadeh, H. R., & Tabibian, S. (2018). Effects of an extreme flood on river morphology (case study: Karoon River, Iran). Geomorphology, 304, 30-39.
7. Wohl, E., Cenderelli, D. A., Dwire, K. A., Ryan-Burkett, S. E., Young, M. K., & Fausch, K. D. (2010). Large in-stream wood studies: a call for common metrics. Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 35(5), 618-625.
8. Wohl, E., Kramer, N., Ruiz-Villanueva, V., Scott, D. N., Comiti, F., Gurnell, A. M., ... & Fausch, K. D. (2019). The natural wood regime in rivers. BioScience, 69(4), 259-273.