Review of the Master's Thesis by Ondrej Pesek
« MASK R-CNN in GRASS GIS »

Caveat

NB : I do not really know what is expected of a Master's Thesis in your university and in these particular studies. Some of my remarks below may thus be inadequate in relation to the expectations. I trust the jury to ignore those remarks.

General assessment

- This is an ambitious Master's thesis dealing with neural networks which are a very complex computing approach. The student has apparently been able to grasp the concepts underlying this approach and implement a working version.
- The student has reached a satisfying result in the form of operational modules for GRASS GIS, already integrated into the GRASS GIS SVN repository. These are the very first neural network modules in GRASS GIS and thus constitute a tremendous achievement for the project.
- The student has explained in detail the evolution that has led to the Mask R-CNN approach, thoroughly retracing the history using the relevant, peer-reviewed scientific literature.
- The thesis is well structured and the English is satisfactory.

Critical remarks

NB: These "critical remarks" are meant as encouragement to go even further on the basis of this very promising work.

Literature review (chapters 1-3)

As mentioned, the literature review is thorough, but mostly limited to retracing the history of Mask R-CNN, citing the relevant articles of each step in that history. In the perspective of scientific writing, one could, however, criticize certain elements of this review.

- The thesis immediately jumps into the review of CNNs, but never provides a motivation of why it would be interesting to use them, or even ANN, other the fact that they are in fashion. The introduction could have been reinforced by a discussion of the user needs. What are you aiming at ? Which problems with other methods are you trying to solve by using ANN ?
- Through the way the literature is presented, Mask R-CNN appears as the logical last step in an almost linear evolution. However, this evolution has not been so linear and different branches leading to different solutions exist. Discussions about these different branches and a review of some of the advantages / disadvantages of the different parallel solutions would have made the literature review, and the case for Mask R-CNN even stronger, especially in combination with a more general motivational introduction as mentioned in the previous point.
- In that same idea: each network is presented on the basis of the original article. It would have made the review stronger if it included some critical responses to these original articles. In summary: for me, in a scientific work, the literature review starts from a clearly defined question and/or hypothesis and then critically analyzes different papers in the

literature in order to see what answers they propose. Here, the literature review is more about following the history of a specific approach by understanding what came before.

- Finally, the literature review is written in a way which does not make it easy for the non-initiated reader to understand everything. In part this is normal, as one cannot expect such a thesis to provide basic education in the field. However, by following and summarizing the original articles, instead of asking overarching questions to then extract the relevant answers from different papers, creates a structure which invites the author to stay very close to the language and logic of these original papers, instead of developing a logic of his own which might have been easier to understand. It is also my experience that the better one understands a topic, the easier it becomes to take a certain distance from it and explain it in a comprehensible matter.

Implementation

As already mentioned, the student has done a wonderful job in introducing ANN to GRASS GIS. The developed modules will be very useful to many people in their current form, but they will also represent the starting blocks for many future developments. It is very important to note that because he was the first, he also had to start from scratch in developing an approach to combine GRASS GIS with the relevant Python tools, notably Keras and Tensorflow. Having this first approach as existing modules will allow debating its advantages and disadvantages and going further in the integration of these tools.
Finally, I want to highlight the specific merit of the student who worked on a computing approach which is linked to very high use of computing resources, without having access to a lot of such resources. Working on these modules in a situation where training to a more or less satisfactory level takes at least one month, demonstrates a very high level of motivation !

Again, here are some critical remarks which should be seen as ideas for further development:

- A small correction: GRASS GIS is first and foremost a C project (and not just “some GRASS GIS modules are written in C”). The arrival of the Python API has provoked the increase of Python-based modules, but the core of GRASS GIS is clearly C. This does not, however, put into question the use of a Python-based approach for this work.
- There has been debate in the past about module naming convention in GRASS GIS. Generally, the consensus was to not exceed the three dot-separated words limit. I would, therefore, suggest to drop the ‘ann’ part, and rename the modules to i.maskrcnn.*
- As can be seen in 5.3. you propose to allow the use of file-based maps (instead of maps registered in the GRASS database). This is quite far from the general GRASS philosophy of using imported, or at least registered (via r./v.external) maps. This potentially has consequences on the way the GRASS computational region settings influence, or not, the analysis. I would strongly plead for GRASS GIS modules to only use GRASS GIS database maps. For the training one could argue that an exception is warranted, in light of the large number of very small image samples needed, which would potentially be quite wearisome to import into GRASS GIS (although in the future we should think about developing GRASS tools to make that easier), but for the detection part, I think the GRASS philosophy should be respected and only internal maps used.
- In figure 5.4 you vectorize each individual map and then you patch all these vector maps. I would be curious to see whether it isn’t more efficient to patch all the raster result maps into one and then vectorize this map, possible using the -t flag of r.to.vect to avoid the creation of an attribute table (which is probably not necessary and source of important overheads).
- In the general i.ann.maskrcnn manual page, you cite OSGeo as a dependency. OSGeo is not software as such and thus can not be a dependency.

- In general, the list of dependencies is quite long and heavy. It would be interesting to study each of these dependencies in detail to see whether they are really needed. Dependency inflation is always a barrier to adoption of a module because it puts a heavier burden on the user to install all these dependencies.
- In the man pages of the two modules, not every parameter is explained. In my eyes, the little description sentence from the source code is generally not clear enough. Ideally, each parameter should have at least one sentence in the Description section. Even if as a developer we think that the parameter's meaning is obvious, it often is not for a user who is not as familiar with the vocabulary of the underlying technologies used. Examples: mini_mask_size and validation_steps in i.ann.maskrcnn.train ?
- Also, both modules provide quite verbose output (although I have to admit that I haven't had the occasion to test the very latest version committed to SVN). This output should be explained in the manual pages.
- As you write in the man page of i.ann.maskrcnn.detect, the order of classes is important, not there names. This is a road to disaster for Alzheimer candidates such as me. It would be better to a) use some form of order-independent naming or identification and/or at least b) provide a tool that allows the user to see which order was used during training (access to training history).
- In the examples of the same man page, there is a section announcing the detection of "cottages and plattenbaus", but in the command, the classes are buildings and lakes.
- More generally, examples should be as self-contained and as reproducible as possible. This means that even on a man page for one particular module, a complete workflow should be given as example, including, if necessary, other steps involving calls to other modules. This is an ideal which I don't always respect either, but we all should… ;-)

Example outputs

- In example B.1 on page 87, you write "the training took one month reaching loss function of 0.9568", and then a bit further, "When the training was stopped, the loss function was about 0.86". This is not clear.

Questions

- What do you see as the advantages/disadvantages of using the ANN approach for object detection and image segmentation as compared to other more classic approaches ?
- The different projects you worked with are all free and open source (GRASS GIS, the Matterport implementation of Mask R-CNN). You say in the text that you had to modify some of the code of these projects to fit your needs: have you fed these modifications back upstream, so that they can be integrated ? If not, why not ?
- Are there parts of your code of which you think that they are generic enough that they could be put into a library so that other modules could also be built using them ?
- At the end of example B.2 you discuss the interesting situation of a higher loss function (B.9) seemingly providing better results than a lower one (B.10). This is an interesting result and probably says a lot about how such ANN works and should probably have been discussed in the text. Do you have an interpretation / explanation of such results ?

Conclusions

In conclusion, I want to congratulate the student for this work. All work can be criticized and I have provided such criticism because it is my role to do so, and because I think the student has clear potential to go further in this domain and so this critique is meant as stimulation for future work.

This thesis, however, shows high levels of determination in adverse circumstances (low resources), a good vision of the lineage of the approach used and its position in the theoretical evolution of the field, as well as the capacity to make existing technologies his own, and modify and complement them as needed in order to create functioning and highly useful tools for the GRASS GIS community.

I, therefore, suggest the grade A and hope to see future work by the student.

Reviewed by :

Dr. Moritz Lennert

Département Géosciences, Environnement et Société
Université Libre de Buxelles
CP 130/03
Av. F.D. Roosevelt 50
1050 Bruxelles
Belgique

Tél : +32 2 650.68.12
Email : moritz.lennert@ulb.ac.be