Road segmentation using U-net and roadmap generation

Marko Balenovic

1. Motivation

Maps are one of the best ways of visualizing data and describing the environment, and they do it in a simple and universal way. Their importance is enormous and they make our lives much easier. Because of this I wanted to create a tool that would provide an easy and simply way to create maps.

2. Research questions

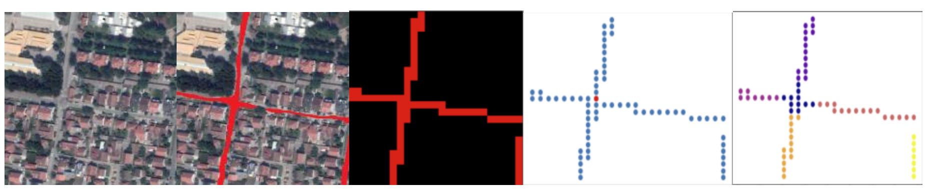
The idea is to create a simple program that would give a map as an output for a satellite image of an area as an input. To achieve this, many problems need to be overcome. First of all, it is necessary to overcome the problems of segmentation of the road from the input images. Some of these problems include vegetation covering parts of the road, road cars, surrounding buildings that resemble the road in shape and color, etc. Some of these problems are solvable at the level of segmentation of the input image while some cannot be solved the traditional way. Therefore, there is a need to postprocess the results obtained by image segmentation to overcome these problems. Postprocessing produces results that are resistant to the above obstacles, and using them we can reach the final result, the map, with greater precision. However, postprocessing brings with it some new obstacles that need to be addressed. The data set represents 400 by 400 satellite images of Paris, where one pixel represents one meter. Although this data set contains many more, only 200 images will be used to solve this problem.

3. Related work

Although the whole process of making maps from satellite images is not so popular today, it cannot be said for the segmentation of satellite images itself. there are many papers dealing with this topic. Although there are many image segmentation methods, one of the best is U-net based architectures. U-net first appeared in 2015 in a Ronneberger`s paper [1] and became one of the best solutions for problems with a limited data set. Although basic U-net shows astounding results, many have sought a way to further improve it. TernausNet [2] and Residual U-net [3] appear, upgraded U-net architectures, that further enhance already great results.

4. Methodology

Given the great results U-net achieves, that was a logical choice for the image segmentation problem. More specific I used TernausNet architecture, that use pretrained encoder, which further enhances the benefits of U-net in working with small datasets. Because of its simplicity and universality, U-net was also a good choice because of postprocessing I had to detect not only roads but also other objects. in that case U-net requires little or no change in implementation. The first step in postprocessing was to convert the segmentation output mask, which is a pixel-vise, into a format that could be more easily processed. The mask is split into blocks, and whether or not the block will be labeled as a road depends on the number of pixels that are labeled as a road in the segmentation process. This reduces the amount of data to be processed which makes it easier for later work. Also, some of the previously mentioned segmentation obstacles such as road cars and vegetation can be removed in this step. After that those blocks are represented by points. Those points must be classified to the paths which they belong. I solved this problem by using the DBSCAN clustering algorithm. This algorithm was best suited to my problem since I could not know the number of paths in the image, that is, the cluster. By one cluster is meant all points belonging to one road between two intersections. This step also eliminates some of the problems I mentioned above, such as buildings look like a road (which can be solved by increasing the minimum number of points) and large vegetation that blocks the road (which can be solved by increasing the maximum distance between points). With increasing the maximum distance, it may be a problem of merging the two clusters into one by ignoring the intersection and treat everything as one road. To prevent this, I introduced two additional dimensions. The first represents the sine of the angle at which the point approaches the intersection divided by distance from the intersection. This prevents the two clusters from merging into one. Since a sine is a symmetric function, a fourth dimension representing the sine of that same angle plus 90 degrees was also necessary, ensuring that no points can have the same value for the third and fourth dimensions if its coordinates are different (first two dimensions). Then using regression through the points of each cluster, I get a polynomial which is an approximation of the road these points represent. And finally, as a last step, I use the resulting polynomial from the previous step to fit 30 equally spaced points and get a finite set of points representing the given road. The figures below show the processing flow of one input image.



5. Discussion

The image segmentation model was trained on 150 images and tested on 50. F-score was used to measure the model's success and the model achieved a f-score of 0.8 which leaves room for improvement but is a good result. As for the other hyperparameters, they were mainly conditioned by a relatively small data set and limited hardware resources. As for batch size I used one and training was performed over 25 epochs. The depth of the network was 5 which was mainly conditioned by the size of the input image but also by the size of the data set. learning rate was 0.001. As far as postprocessing of data is concerned, I do not have an exact measure of precision so I will just go through the shortcomings noted. The biggest problem in postprocessing can be intersections that are very close to each other. this is where postprocessing techniques cannot determine with any fervor the road between them. Other errors are mainly due to segmentation errors where, despite the tools to deal with them, postprocessing could not find a solution.

6. References

1. Ronneberger O, Fisher P., Brox T. (2015) U-Net: Convolutional Networkds for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham
2. Vladimir Iglovikov, Alexey Shvets (2018) TernausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation
3. Z Zhang, Q Liu, Y Wang (2018) Road extraction by deep residual u-net