[Team 23]Project Proposal 2.1: Component Segmentation of Aerial Views/Images

Rajshree Jain
rjain27@ncsu.edu

Sai Shruthi Madhuri Kara
skara2@ncsu.edu

Sruthi Talluri
stallur2@ncsu.edu

1 MOTIVATION

The component segmentation of images provides a way to easily recognize various regions of interest (ROIs) as huge volumes of optical overhead imagery are captured every day with airborne or space-borne platforms making it impossible for manual interpretation. In the aerial views of cities and landscapes, this can be used to distinguish objects like roads, buildings, and so on [1, 2, 3]. This task can be applied for map making, object detection, and geographical monitoring. It is also very helpful in ecological tasks such as monitoring deforestation or damage in an area impacted by natural disaster[4]. The focus of the project will be successful road detection given aerial images of cities, further for buildings, with the conviction that the process can be applied to some of the aforementioned tasks with further research and testing.

2 DATA

We will be working on a dataset named "Massachusetts Roads Dataset" and "Massachusetts Buildings Dataset" that is used in the PhD thesis by Volodymyr Mnih for Ariel Image segmentation [1]. Both the datasets have the images of the ariel views and the corresponding masks in form of (black and white) images. These masks will be the ground truth value identifying that the image has a road or a building. Both the datasets contain images of size 1500*1500 covering an area of 2.25 kilometers. The Massachusetts Buildings Dataset consists of 151 images and corresponding masks where the random split is made for training, test and validation. The Training set contains 137 images, test set has 10 images and 4 images in validation set. Further, the Massachusetts Roads Dataset consists of 1171 aerial images and their corresponding masks. For this dataset the images are randomly split into training set of 1108 images, a validation set of 14 images and a test set of 49 images. One of the interesting aspects of the project is that since the data for buildings and roads is separate, we will also work on trying to build a model that will generate masks for both roads and buildings given an image be it either road image or building image.

3 METHODOLOGY AND EVALUATION

We believe that convolutional neural networks (CNN) especially with U-Net architecture would best suit for this project classifying the components into at least 2 categories like roads or other. The approach for this project would be following the U-net CNN model which was successfully implemented on the biomedical image segmentation [5]. We consider this as the baseline for our project. As a part of the next stage, we will be improving this model by hyperparameter tuning, for example, increase/decrease the layers etc. We consider training the CNN with stochastic gradient descent(SGD) and taking cross entropy as a loss function more prominent. For validation of the results(performance - accuracy, error rate), we will be utilizing the binary mask images dataset from Dr. Mnih's thesis. The dataset is already split-up and readily available.

REFERENCES

[1] V. Mnih, "Machine Learning for Aerial Image Labeling," PhD Thesis, University of Toronto, 2013.
[2] U. Ankit, 'Semantic Segmentation of Aerial images Using Deep Learning,' 2019. [Online]. Available at https://towardsdatascience.com/semantic-segmentation-of-aerial-images-using-deep-learning-90fdf4ad780 [Accessed: 22-Oct-2020]
[3] A. Nowaczynski, 'Deep learning for satellite imagery via image segmentation,' 2017. [Online]. Available at https://deepsense.ai/deep-learning-for-satellite-imagery-via-image-segmentation/ [Accessed: 22-Oct-2020]
[4] J. Anumula, 'Semantic Segmentation on Aerial Images using fastai,' 2019. [Online]. Available at https://medium.com/swlh/semantic-segmentation-on-aerial-images-using-fastai-a2696e4db127 [Accessed: 22-Oct-2020]
[5] https://github.com/hlamba28/UNET-TGS