Deep Learning
Project Report
(CSE4003)

Satellite Image Classification

By

Abhinav Koora

190122

Deepak Perla

190137

Department of Computer Science and Engineering
School of Engineering and Technology

April 2022

Contents

1. Problem Definition

2. Project Objectives

3. Challenges

4. Literature Review

5. Description of the Dataset

6. Proposed Methodology

7. Experimental Results/Comparison

8. Justification of novelty of the proposal

9. Role & Contributions of each group member

10. Conclusions and Future Scope

11. References
    (All relevant works should be cited appropriately)

Appendix: with solution code

Abstract

The significance of road extraction from satellite photos originates from the fact that it substantially improves map-generating efficiency and hence can be a big assist in automobile navigation systems or any emergency (rescue) system that requires instant maps. The image is segmented to identify the road network regions, followed by a decision-making and continuity approach to correctly recognize the roads, and the Vectorization stage to identify the line segments or curved segments that constitute the road. In the event that a completely automated system fails, a semi-automated approach is adopted.

1. Problem Definition

The road is an important type of fundamental geographic information. In traffic management, urban planning, automatic vehicle navigation, and emergency management, road information retrieval is critical. And also in disaster zones, especially in developing countries, maps and accessibility information are crucial for crisis response. The quality of high-resolution satellite photos has increased and become more accessible as remote sensing technology has progressed, making it possible to use remote sensing photographs to correctly detect roadways. As a result, extracting road information from remote sensing photographs is a critical concern.

2. Project Objectives

In this project, we particularly solve the real-life problem using generally deep learning algorithms, so here we used the U-Net algorithm to mostly predict the roadway, from the satellite and aerial images from a data set called Deep-Globe Road Extraction. And we used convolution network architecture, it mostly is also fast and does precise segmentation of images, which basically is quite significant.

3. Challenges

This project requires a lot of computational power and particularly is also re-run with generally quick hits and trials being conducted, or so they essentially thought. Which mostly is also a kind of complex process and for the most part has to for all intents and purposes be done using actually slow implementation, showing how this project requires a lot of computational power and definitely is also re-run with definitely quick hits and trials being conducted in a definitely big way. Sometimes the model becomes over-fitting and mostly under-fitting in a definitely major way.

4. Literature Review

Year	Author	Objective	Dataset	Method Strategy	Outcomes
2020	Baris Dincer	Satellite Road Extraction/Auto-Encoder Prediction	DeepGlobe	Auto Encoder	Accuracy 95%
2020	Balraj Ashwath	Roadmaps from aerial Images	DeepGlobe	DeepLabV3+	Accuracy 96%
2017	Zhengxin Zang	Road Extraction by Deep Residual U-Net	Massachusetts Roads Dataset	Unet	Accuracy 90.53%

5. Description of the Dataset

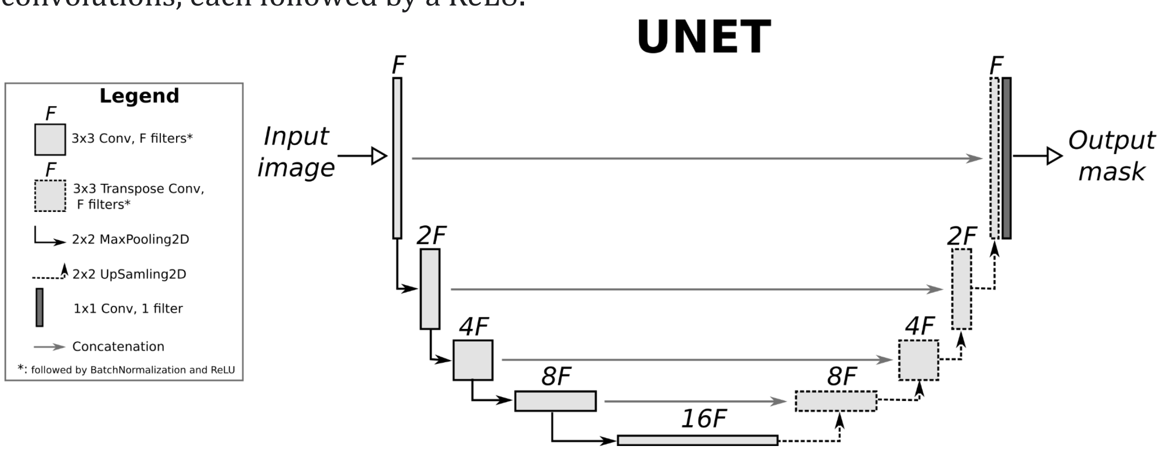
The data set memory consumption is around 4GB. We can download this data set easily by using Kaggle API. The structure of the data is, that it has a metadata.csv file that contains image paths for all training, validation, and test sets.

- The training data for Road Challenge contains 6226 satellite imagery in RGB, size 1024x1024.
- The imagery has a 50cm pixel resolution, collected by DigitalGlobe's satellite.
- The dataset contains 1243 validation and 1101 test images (but no masks).
- Each satellite image is paired with a mask image for road labels. The mask is a grayscale image, with white standing for the road pixel, and black standing for the background.

- File names for satellite images and the corresponding mask image are id _sat.jpg and id _mask.png. id is a randomized integer

6. Proposed Methodology

U-Net is a semantic segmentation architecture. It has two paths: one that contracts and one that expands. The convolutional network's contracting route follows the standard architecture. It comprises two 3x3 convolutions (unpadded convolutions) that are applied repeatedly, each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. We quadruple the number of feature channels with each downsampling step. Every step in the expanding route consists of an upsampling of the feature map, followed by a 2x2 convolution ("up-convolution") that halves the number of feature channels, a concatenation with the equally cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU.



U-Net: In semantic segmentation, it is critical to leverage low-level details while keeping high-level semantic information to get a finer outcome. However, building a deep neural network like this is difficult, especially when there are only a few training examples available. Using a pre-trained network and then fine-tuning it on the target data set, as done in, is one technique to tackle this problem. Another option is to use considerable data augmentation as U-Net did. We believe that U-design, Net's in addition to data augmentation, help to alleviate the training challenge.

The idea is that replicating low-level characteristics to the equivalent high levels really establishes a conduit for information propagation, allowing signals to pass between low and high levels much more easily, which not only aids backward propagation during training but also compensates low-level finer details to high-level semantic features. This is similar in concept to the residual neural network.

7. Experimental Results/Comparison

The values we achieved are as follows:

- Accuracy: 61%
- IoU: 0.41
- Binary Accuracy: 93%

8. Justification of novelty of the proposal

- We have used DeepGlobe datasets having normal and masked images for finding the roadway so we have achieved good accuracies using U-Net.
- Though they have better accuracy in prediction than ours, we have used better datasets and achieved decent accuracy.

9. Role & Contributions of each group member

We have equally divided our work to do in this project. Firstly, we together collected the datasets required for our project. Then we have built a U-Net model to predict the Road from aerial and satellite images.

10. Conclusions and Future Scope

The U-net essentially is a method for extracting roads from high-resolution remote sensing photos that we for all intents and purposes present in this study, which literally is fairly significant. The proposed network really combines U-Net's Algorithm. Information propagation in forwarding and backward computations will literally be aided by for all intents and purposes skipping connections inside actually residual units and between the network's encoding and decoding routes in a subtle way. This particular characteristic not only essentially makes training generally easier but also enables us to for the most part create definitely basic but fairly effective neural networks. In the future, we can for the most part take blurred and grained images and we can use them to actually predict roadways, which actually is fairly significant.

11. References

- https://www.spiedigitallibrary.org/conference-proceedings-of-spie/5239/0000/Semi-automatic-road-extraction-from-aerial-images/10.1117/12.508365.full
- http://ieeexplore.ieee.org/iel5/4638686/4654597/04654610.pdf
- https://www.researchgate.net/publication/303320214_A_Review_of_Road_Extraction_from_Remote_Sensing_Images

- https://www.sciencedirect.com/science/article/abs/pii/092427169598233P
- https://arxiv.org/abs/1711.10684

Appendix: with solution code