

Supervised Land use Land cover classification



**Cluster Innovation Centre
University of Delhi**

Ashutosh Jha | Sachin | Shubham Chauhan | Yash Kumar

Nov 2021

Project submitted for the paper of:
Environmental Management

CERTIFICATE OF ORIGINALITY

The work embodied in this report entitled “**Supervised Land use Land cover classification**” has been carried out by **Ashutosh Jha, Sachin, Shubham Chauhan & Yash Kumar** for “**Environmental Management**”. We declare that the work and language included in this project report is free from any kind of plagiarism.

The work submitted is original and has not been submitted earlier to any institute or university for the award of any degree or diploma.

Ashutosh Jha

(11811)

Sachin

(11831)

Shubham Chauhan

(11838)

Yash Kumar

(11844)

ABSTRACT

In this project we explain some simple modifications made to the method for the adaption of remote sensing data. Our method is flexible and it is possible to adapt the approach to several scenarios, different image scales, or other earth observation applications, using spatially resolved data. However, the focus of the current work is on high resolution satellite images of urban areas and classification of land cover.

INTRODUCTION

The beginning of land cover classification from aerial images dates back around 70 years . Since then aerial and satellite images are used to extract land cover in a broadly manner and without direct contact to the observed area. It is essential information for change detection applications or derivation of relevant planning or modeling parameters. Other fields of applications are the analysis and visualization of complex topics like climate change, biodiversity, resource management, living quality assessment, land use derivation or disaster management. Therefore, methods of automated land cover extraction on the basis of area-wide available remote sensing data are utilized and continually improved.

In this work, we collected satellite land images for semantic segmentation and classification. For instance, the probability of large buildings and impervious surfaces (e.g., parking slots) in industrial areas are much higher than in allotment areas.

STUDY AREA AND DATA USED

The focus of this study is the research area of Rostock, a city with more than 200.000 inhabitants in an area of 181 km², situated in the north of Germany. A subset of five by five kilometers of a cloudfree Quickbird scene from September 2009 was available for this study to develop and test the method. It represents the south-west part of Rostock, including the Warnow river in the north, parts of the city center and adjacent fields. The Quickbird scene has four multispectral channels (blue, green, red, near infrared), which were pansharpened with the panchromatic channel to a spatial resolution of 60 cm per pixel. The scene was provided in the OrthoReady Standard (OR2A) format. The image was corrected for atmospheric effects and orthorectified using ground control points and a digital terrain model.



Figure 1. Quickbird satellite image subset of Rostock

SEMANTIC SEGMENTATION

In computer vision, the term semantic segmentation covers several methods for pixel-wise annotation of images without a focus on specific tasks. At which, segmentation denotes the process of dividing an image into a disjoint group of pixels. Each of those groups is called a region. Furthermore, all pixels in a region are homogeneous with respect to a specific criteria (e.g., color or texture). The target of segmenting an image is to transform the image into a better representation, which is reduced to the essential parts. Furthermore, segmentation can be deferred into unsupervised and supervised segmentation. Unsupervised segmentation denotes that all pixels are grouped into different regions, but there is no meaning annotated to any of them. However, for supervised segmentation or semantic segmentation a semantic meaning is annotated to each region or rather to each pixel. Usually, this is a class name out of a predefined set of class names. The selection of those classes highly depends on the chosen task and the data. For instance, a low resolution satellite image of a whole country can be analyzed, where the classes city and forest might be interesting. Alternatively, if we classify land cover of very high resolution satellite images of cities, classes like roof, pool, or tree are recognizable in the image.

Essential Foundations

Feature vectors are compositions of multiple features. Each feature vector describes an object or a part of an object. For instance, the mean value of each color channel is such a collection of simple features. To describe more complex structures, we need besides color also texture and shape as important features.

Classification denotes the problem in pattern recognition of assigning a class label to a feature vector. Therefore, a classifier needs an adequate set of already labeled feature vectors. The classifier tries to model the problem out of this training data during a training step. With this model, the classifier can assign to each new feature vector a label during testing.

Millions of features

The presented method is based on the extraction of multiple features from the input image. Besides the single input channels, additional channels can be computed, e.g., gradient image. On each of these channels and in combination of those several features can be computed in a local neighborhood d . For instance, the difference of two random selected pixels relative to the current pixel position or the mean value of a random selected pixel relative to the current position.

RESULTS

For testing, we used some already labeled training areas. On the rest of the dataset 65 points per class are randomly selected for testing. We focused on the classes: tree, water, bare soil, building, grassland, and impervious. All tests are made with a fixed window size $d = 30\text{px} = 18\text{m}$ for non-contextual features and $d = 120\text{px} = 72\text{m}$ for all contextual features. The qualitative results of our proposed method are presented in Figure 2. The main problems are to differ between the classes impervious and building as well as grassland and bare soil. The classes tree and water are already well classified. This accords to our expectations, due to the fact that those both classes look very similar from the bird's eye view but they differ in the height. This is also what we expected, due to the fact that grassland has a much brighter appearance than bare soil. But there are still some confusions between bare soil and grassland. Almost every value without using context is worse than the values using contextual cues. Especially, bare soil and impervious benefits from using contextual knowledge. Without contextual knowledge the class bare soil is often confused with grassland and impervious, but using contextual knowledge impervious and bare soil are well classified. One suitable explanation for this might be that bare soil is often found on harvested fields outside the city. Due to this reason, the probability for classes like grassland or impervious is much higher in the neighborhood of buildings.

There are some problems with the shadow of trees, which are not represented enough in the training data. Furthermore, small objects (like small detached houses) vanish in the classification result with a larger window size. Finally, the object borders are very smooth, this can be fixed by using an unsupervised segmentation.

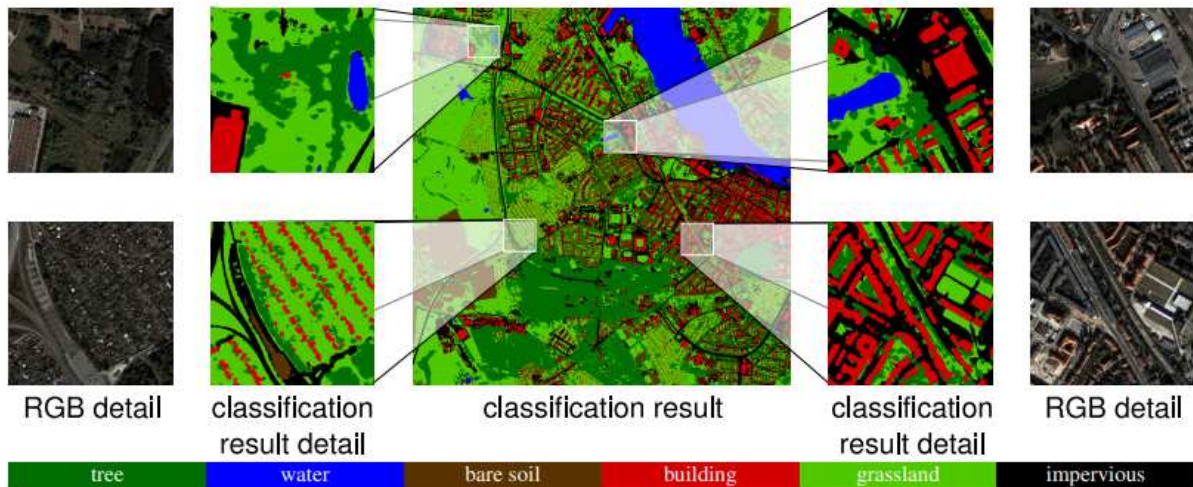


Figure 2. Classification result and four sample areas in full resolution

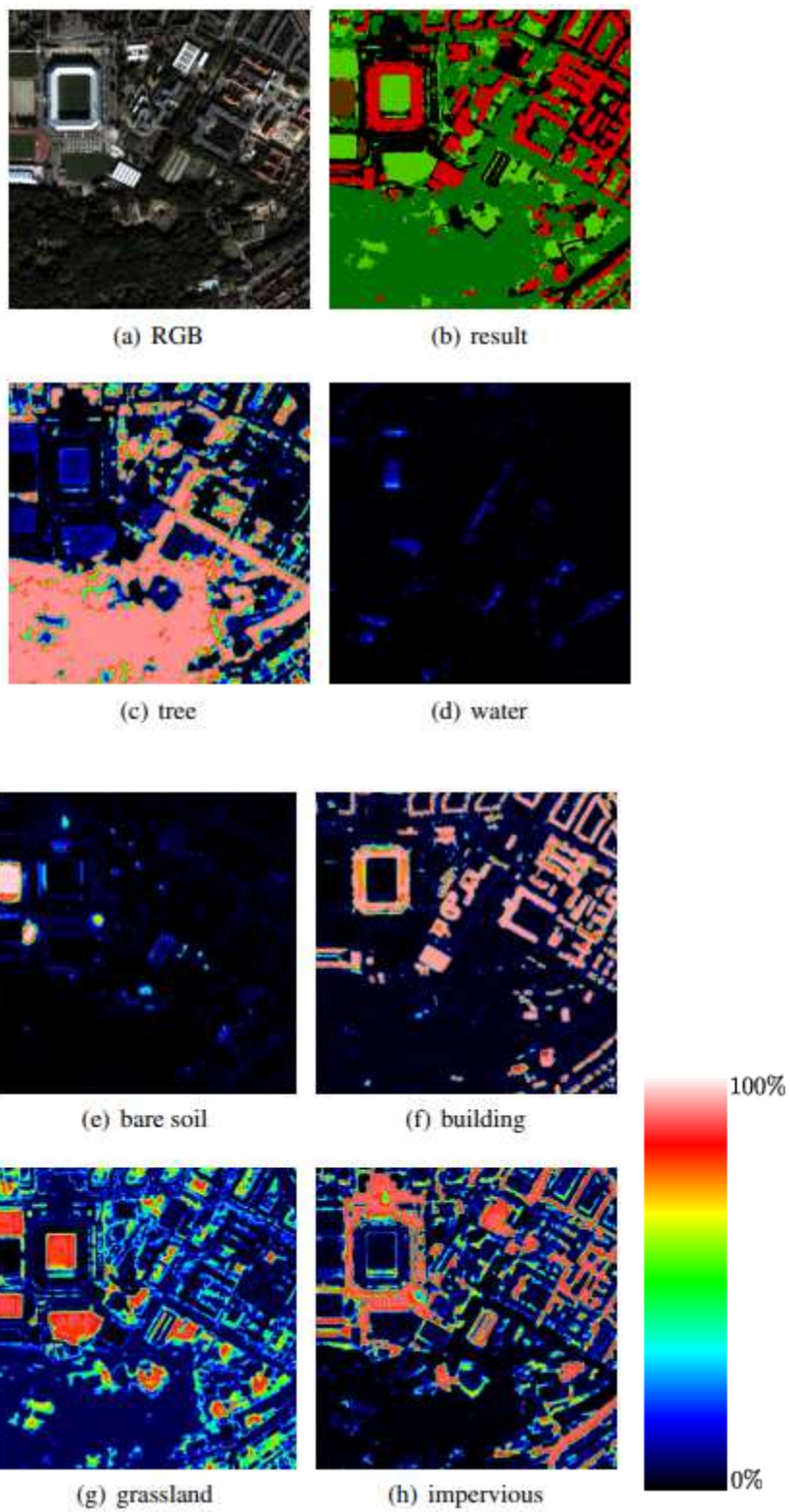


Figure 3. Probability maps for all classes (each sample area)

CONCLUSION & FURTHER WORK

In this work, we introduced a state of the art approach from computer vision for semantic segmentation. Furthermore, we have presented how to adapt this method for the classification of land cover. In our experiment, we have shown that our method is flexible in using multiple channels and that adding channels increases the quality of the result. The benefits of adding contextual knowledge to the classification has been demonstrated and discussed for some specific problems.

For further work, we are planning to use an unsupervised segmentation to improve the performance especially at the borders of the objects. Furthermore, we are planning to incorporate shaped information.

REFERENCES

1. Mecklenburg- Vorpommern Statistisches Amt, 2012.
2. Duda, R. O. and Hart, P. E., 1973. Pattern Classification and Scene Analysis. Wiley.
3. Frohlich, B., Rodner, E. and Denzler, J., 2012. Semantic segmentation with millions of features: Integrating multiple cues in a combined random forest approach. In: Proceedings of the Asian Conference on Computer Vision (ACCV)
4. Breiman, L., 2001. Random forests. Machine Learning 45(1), pp. 5–32.