Copernicus Master in Digital Earth

Student: Emanuel Goulart Farias

Student number: 12413874

Advanced Remote Sensing

Part I

Objective:

Develop a supervised classification on the dataset that contains an image of Salzburg with four bands: Red, Green, NIR, and SWIR. The classification is developed in four main steps: Create an image object level; Define Samples for each target class; Train Classifier based on samples and selected features; Prediction.

The study was conducted on the software E-cognition.

1.1 Initial Segmentation

The initial segmentation was conducted using the algorithm *multiresolution segmentation,* generating 880 objects in the scene.

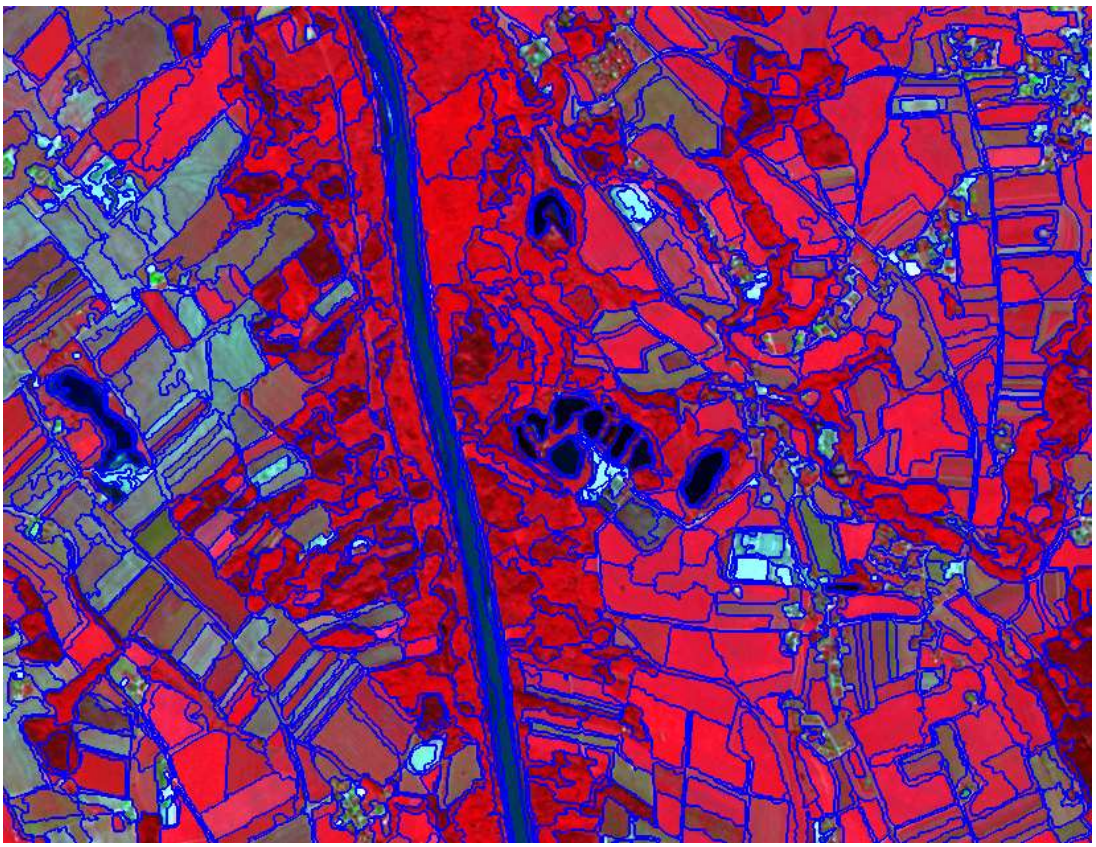
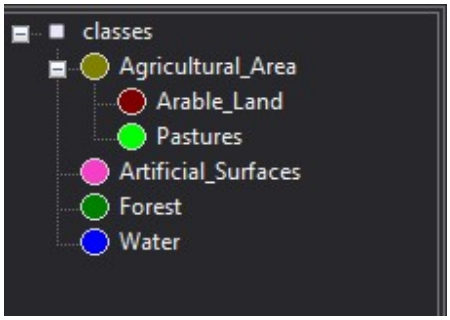


Figure 1- Objects generated from multi-resolution segmentation algorithm.

2.1 Defining samples

To proceed with the classification, it is necessary to creating different class in order to represent the targets of the classification. The following classes were created:



At the first glance, I sampled a few objects for each class, where the image below show up how the classes are distributed in my image.

Figure 2 - Objects sampled to represent each class.

In order to find different spectral signatures for each class, I used the sample editor window to visualize better the spectral signature of the classes. Aiming to understand which features are suitable to operationalize the classification algorithms.

The image below shows the difference on the spectral signature of the classes pasture and arable lands.

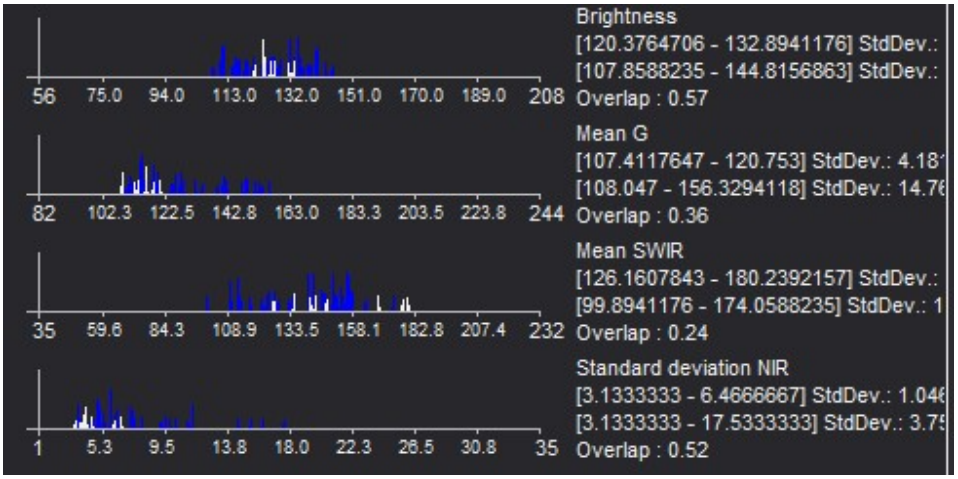


Figure 3 - Spectral signatures for different classes.

Question: What is the result if you select class related or hierarchical features?

When selecting class related features is possible to observe the difference of the two spectral signatures for the class. The figure 3, above, shows the spectral signature for the classes and make it possible to highlight the difference between arable lands and pastures. On the other hand, when selecting hierarchical features, the result shows the distribution of the spectral signature for all the objects created in the multi-resolution algorithms. The comparison between the hierarchical classes can be visualized on the figure 4 above.

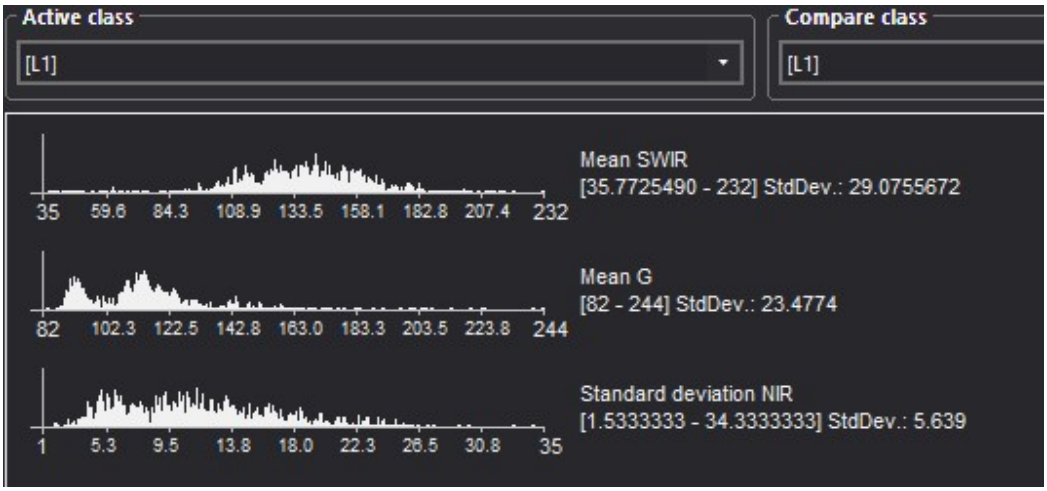


Figure 4 - Comparing hierarchical features.

3.1 Selecting features

In order to best optimization the space without redundancies the tool *Feature Space Optimization* is used to calculate different feature combinations and give of which features are best suit to separate the classes. The results of the optimization selected the following features: Mean SWIR; Mean G, GLCM Homogeneity; Area; Mean NIR, std NIR, std SWIR.

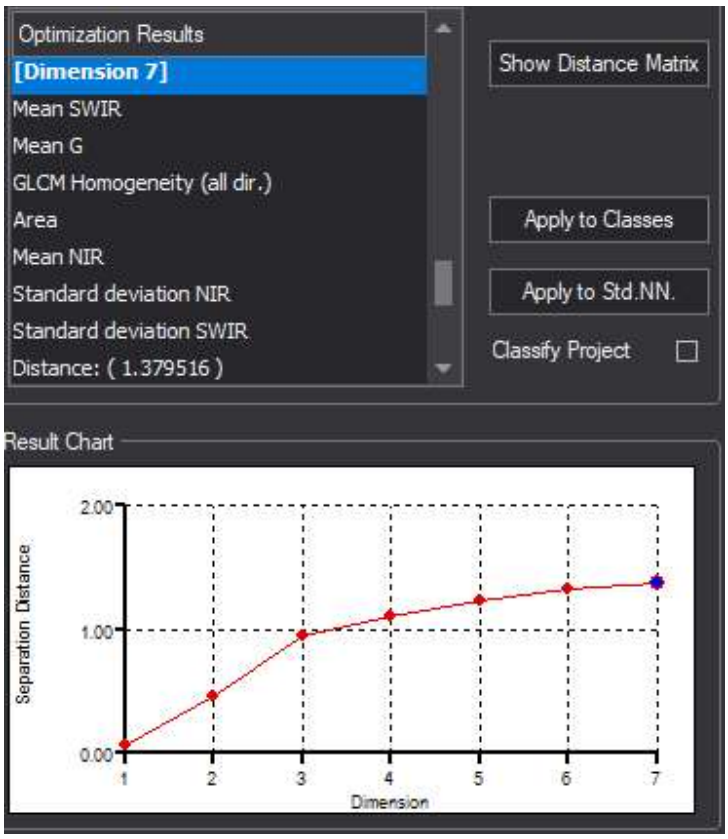


Figure 5 - Feature Selection optmization

4.1 Model Train and Predict

Following on, the supervised classification is set up in the process tree of the e-cognition software. For the first supervised classification, the algorithm choice was KNN, with the features selected on the step of the feature optimization.

The model was trained and the results can be visualized on the figure 6.

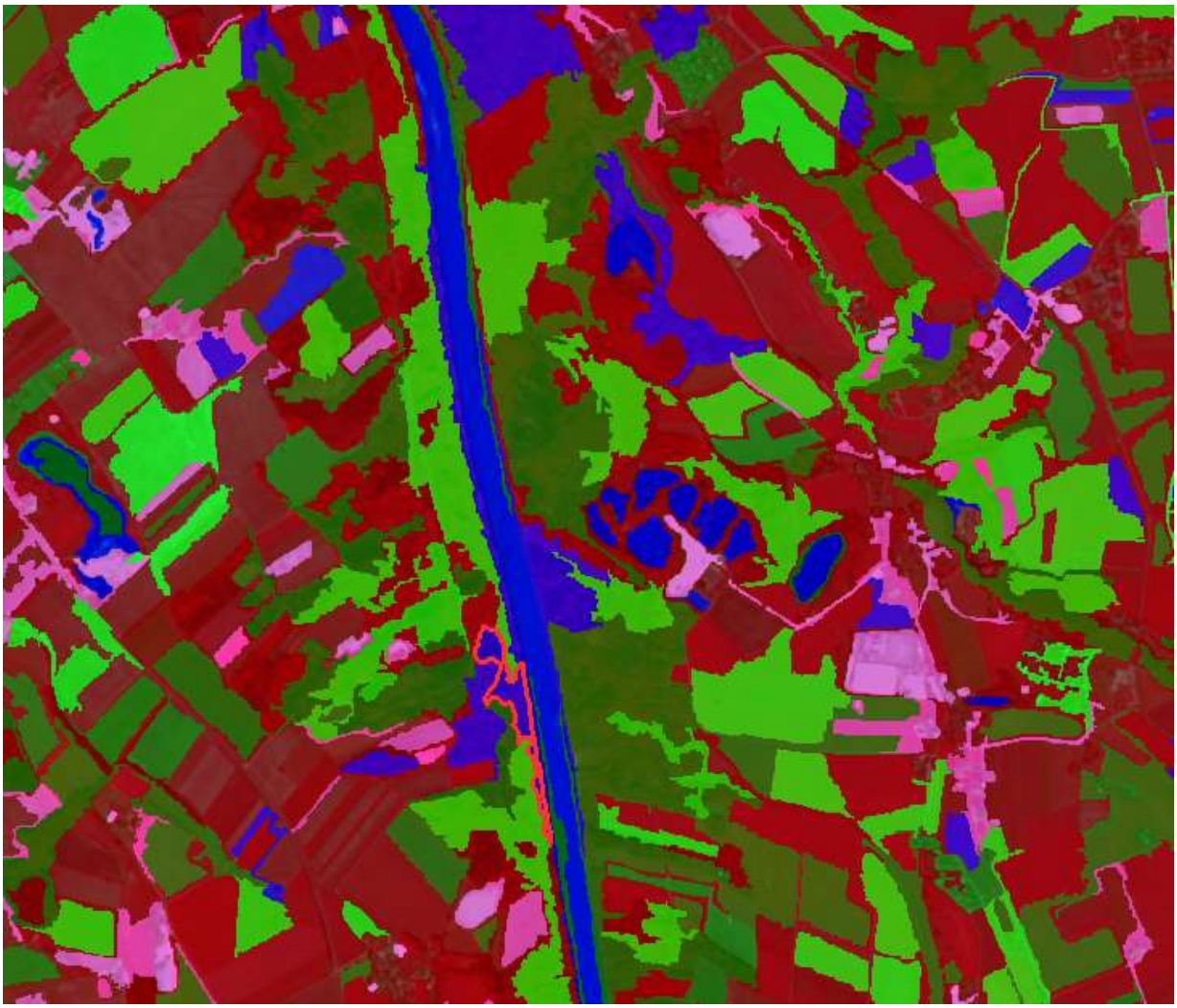


Figure 6 - Results of the KNN classification

After a briefly inspection, the results of the KNN classification was not satisfactory. Classifying wrongly the water class in areas related to forests. Also, the forest class area get mixed up with the pastures. In general, artificial surfaces were under represented in the image.

Trying to improve the classification, I decided to run the Support Vector-Machine (SVM) algorithm. The results are shown below:

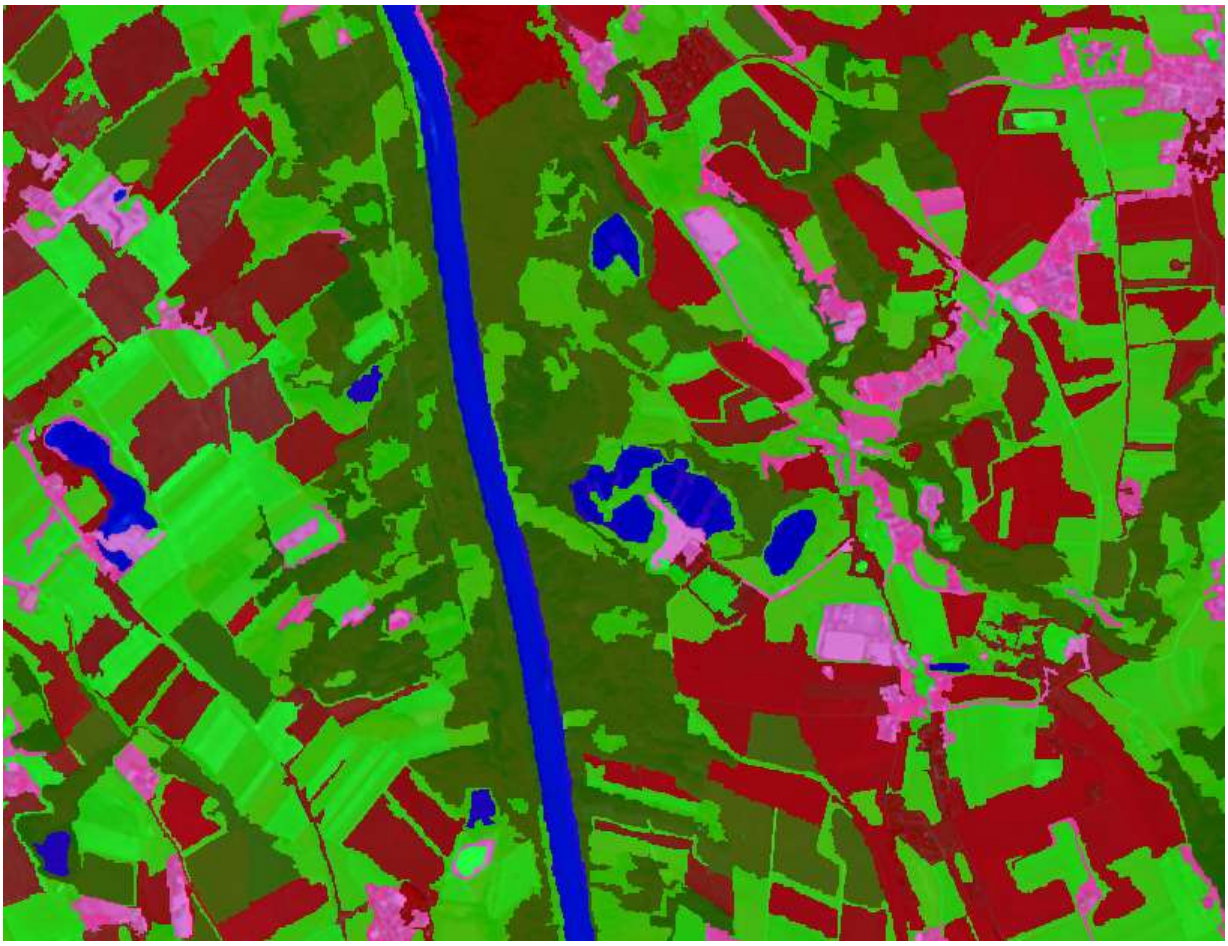


Figure 7 - Results of the SVM algorithm classification.

The results of the SVM algorithm shows a better classification when compared to KNN, however, it is still misclassifying many classes. The arable land and pastures are being classified as forest at some parts of the image. At least, the river and artificial surfaces are better classified in comparison with KNN.

Trying to improve the classification of the algorithm, I went further by reducing the features used to train the model, adding more samples and fine-tunning the parameters of the classification algorithm.

The new features chosen were: Mean G; Mean NIR, Mean SWIR: Area; GLCM Homogeneity;

I have tested different algorithms: Random Forest, Decision Trees, so on and so far. With different hypertunning parameters. The best result I got it was with SVM and a hypertunning of the C parameter as 3.

The image below shows the result, classifying well the forests, the artificial areas and water. The arable lands are also quite well representable and the pastures are not well classified. Definitely there is room to improvement, but the classification processes but it is quite satisfactory.

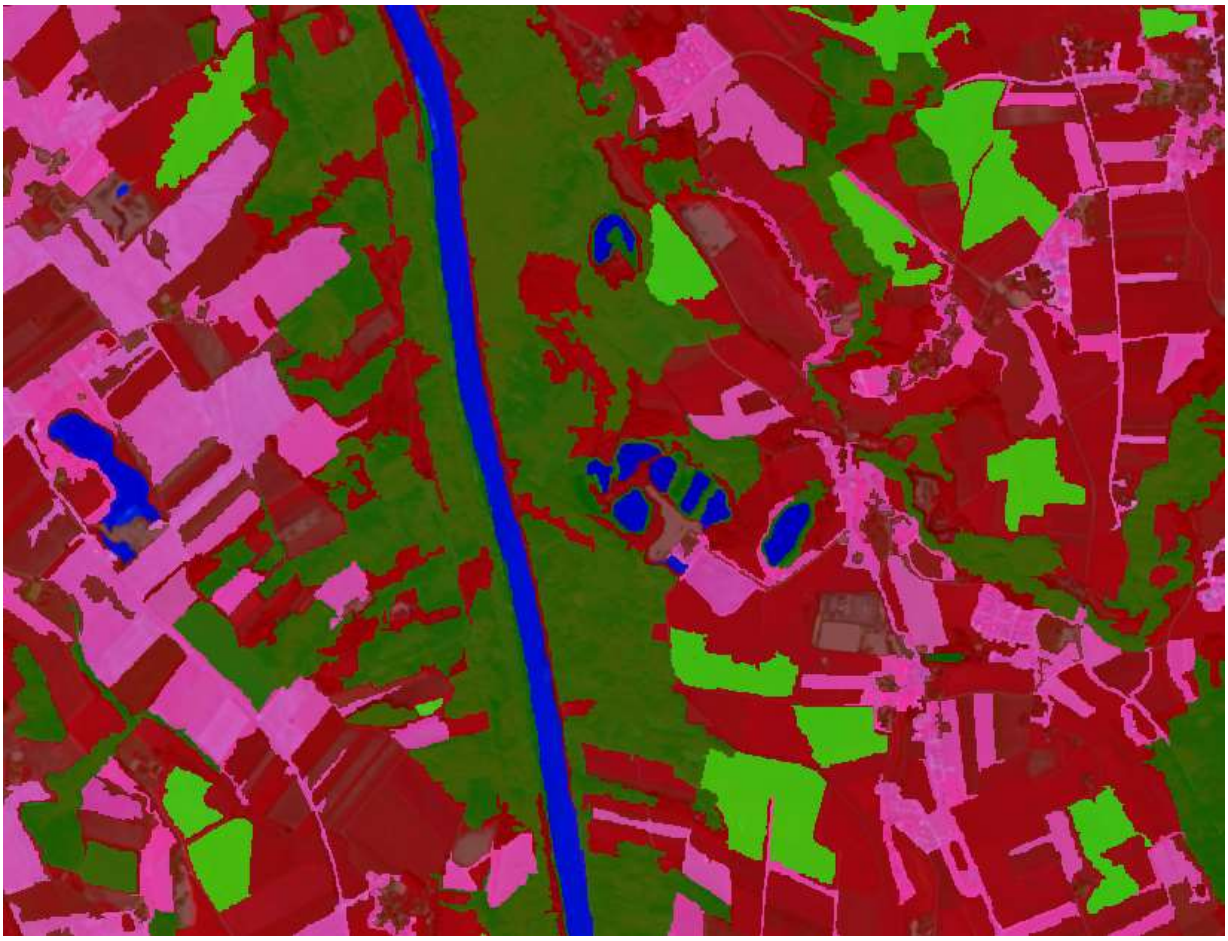


Figure 8 - Best classication developed with SVM.