Geophysical Research Abstracts
Vol. 19, EGU2017-15692-1, 2017
EGU General Assembly 2017

Comparison of Random Forest and Support Vector Machine classifiers using UAV remote sensing imagery

Marco Piragnolo (1), Andrea Masiero (2), and Francesco Pirotti (1)

(1) Department of Land, Environment, Agriculture and Forestry (TESAF), University of Padova, Padova, Viale dell'Università 16 Legnaro (PD) 35020, Italy (marco.piragnolo@phd.unipd.it, francesco.pirotti@unipd.it), (2) Interdepartmental Research Center of Geomatics (CIRGEO), University of Padova, Viale dell'Università 16, Legnaro (PD) 35020, Italy

Since recent years surveying with unmanned aerial vehicles (UAV) is getting a great amount of attention due to decreasing costs, higher precision and flexibility of usage. UAVs have been applied for geomorphological investigations, forestry, precision agriculture, cultural heritage assessment and for archaeological purposes. It can be used for land use and land cover classification (LULC). In literature, there are two main types of approaches for classification of remote sensing imagery: pixel-based and object-based. On one hand, pixel-based approach mostly uses training areas to define classes and respective spectral signatures. On the other hand, object-based classification considers pixels, scale, spatial information and texture information for creating homogeneous objects. Machine learning methods have been applied successfully for classification, and their use is increasing due to the availability of faster computing capabilities. The methods learn and train the model from previous computation. Two machine learning methods which have given good results in previous investigations are Random Forest (RF) and Support Vector Machine (SVM).

The goal of this work is to compare RF and SVM methods for classifying LULC using images collected with a fixed wing UAV. The processing chain regarding classification uses packages in R, an open source scripting language for data analysis, which provides all necessary algorithms. The imagery was acquired and processed in November 2015 with cameras providing information over the red, blue, green and near infrared wavelength reflectivity over a testing area in the campus of Agripolis, in Italy. Images were elaborated and ortho-rectified through Agisoft Photoscan. The ortho-rectified image is the full data set, and the test set is derived from partial sub-setting of the full data set. Different tests have been carried out, using a percentage from 2 % to 20 % of the total. Ten training sets and ten validation sets are obtained from each test set. The control dataset consist of an independent visual classification done by an expert over the whole area. The classes are (i) broadleaf, (ii) building, (iii) grass, (iv) headland access path, (v) road, (vi) sowed land, (vii) vegetable. The RF and SVM are applied to the test set. The performances of the methods are evaluated using the three following accuracy metrics: Kappa index, Classification accuracy and Classification Error. All three are calculated in three different ways: with K-fold cross validation, using the validation test set and using the full test set.

The analysis indicates that SVM gets better results in terms of good scores using K-fold cross or validation test set. Using the full test set, RF achieves a better result in comparison to SVM. It also seems that SVM performs better with smaller training sets, whereas RF performs better as training sets get larger.