Investigating the Influence of Training Set Size for Crop Type Classification using RapidEye

ISBN 978-85-88783-11-9

Authors

1Balik Sanli, F.; 2Ustuner, M.; 3Bektas Balcik, F.; 4Goksel, C.

1YILDIZ TECHNICAL UNIVERSITY *Email*: fbalik@yildiz.edu.tr
2YILDIZ TECHNICAL UNIVERSITY *Email*: mustuner@yildiz.edu.tr
3ISTANBUL TECHNICAL UNIVERSITY *Email*: bektasfi@itu.edu.tr
4ISTANBUL TECHNICAL UNIVERSITY *Email*: goksel@itu.edu.tr

Abstract

Agriculture have a significant role at country's economy. With the recent technologic development, remote sensing offers a viable solution for the sustainable management and appropriate planning of the agricultural areas since it provides accurate, up-to-date and cost-effective information about earth surface. Remotely sensed images at the different level of resolution from different types of sensors have been extensively and successfully used for crop mapping and identification. Since image classification is one of the common method to obtain thematic information from the images, it has a great attention by remote sensing scientists to develop new classification algorithms. Several factors may affect the accuracy of supervised classification, especially number of the training samples. In this study, it is aimed to investigate the influence of the training set size for crop type classification using RapidEye high resolution satellite imagery. For this purpose, three different training dataset have been used for classification. For each class of interest of the study area, the number of 50, 100, 200 pixels have been used for this three different training dataset, respectively. Support Vector Machines (SVM), is one of the machine learning algorithms and preferred recently since its superior performance on image classification, and Maximum Likelihood (ML), is one of the conventional classifiers, have been carried out here for the classification. Radial Basis Function (RBF) has been implemented here as a kernel type for SVM classification. Grid search method, which is widely used in kernel parameter selection, has been implemented here for the determination of the optimum parameters for RBF kernel. Optimum parameter were determined as 300 and 0,2 for C (cost) and γ (gamma), where γ is the kernel width and C is the penalty parameter, respectively. RapidEye imagery which has been highly demanded for agricultural and forestry applica-tions is used here for the study area located in Turkey. The study area is located in Aegean region of Turkey and comprised of intensive cultivation areas. This area covers nine land use classes which are corn (first crop, second crop), cotton (well developed, moderate developed, weak, developed), soil (wet, moist, dry) and water surface. Overall accuracy and kappa coefficient have been used for assessing the image classification accuracy. The classification accuracies with are 79,18%, 82,25%, 84,30% and 82,94%, 82,59%, 81,91% for MLC and SVM, respectively with the increasing number of pixel for each class as 50, 100 and 250. Results demonstrate that classification accuracy with MLC is increased from 79.18% to 84,30% as overall accuracy however SVM is decreased from 82,94% to 81,91% when the number of pixels for each class of interest has changed from 50 to 250. Additionally, our results show that the use of RapidEye imagery with the SVM and MLC can get satisfactory results of classification accuracy for crop type mapping