World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:14, No:12, 2020

A Nonlinear Feature Selection Method for Hyperspectral Image Classification

Authors : Pei-Jyun Hsieh, Cheng-Hsuan Li, Bor-Chen Kuo

Abstract : For hyperspectral image classification, feature reduction is an important pre-processing for avoiding the Hughes phenomena due to the difficulty for collecting training samples. Hence, lots of researches developed feature selection methods such as F-score, HSIC (Hilbert-Schmidt Independence Criterion), and etc., to improve hyperspectral image classification. However, most of them only consider the class separability in the original space, i.e., a linear class separability. In this study, we proposed a nonlinear class separability measure based on kernel trick for selecting an appropriate feature subset. The proposed nonlinear class separability was formed by a generalized RBF kernel with different bandwidths with respect to different features. Moreover, it considered the within-class separability and the between-class separability. A genetic algorithm was applied to tune these bandwidths such that the smallest with-class separability and the largest between-class separability simultaneously. This indicates the corresponding feature space is more suitable for classification. In addition, the corresponding nonlinear classification boundary can separate classes very well. These optimal bandwidths also show the importance of bands for hyperspectral image classification. The reciprocals of these bandwidths can be viewed as weights of bands. The smaller bandwidth, the larger weight of the band, and the more importance for classification. Hence, the descending order of the reciprocals of the bands gives an order for selecting the appropriate feature subsets. In the experiments, three hyperspectral image data sets, the Indian Pine Site data set, the PAVIA data set, and the Salinas A data set, were used to demonstrate the selected feature subsets by the proposed nonlinear feature selection method are more appropriate for hyperspectral image classification. Only ten percent of samples were randomly selected to form the training dataset. All non-background samples were used to form the testing dataset. The support vector machine was applied to classify these testing samples based on selected feature subsets. According to the experiments on the Indian Pine Site data set with 220 bands, the highest accuracies by applying the proposed method, F-score, and HSIC are 0.8795, 0.8795, and 0.87404, respectively. However, the proposed method selects 158 features. F-score and HSIC select 168 features and 217 features, respectively. Moreover, the classification accuracies increase dramatically only using first few features. The classification accuracies with respect to feature subsets of 10 features, 20 features, 50 features, and 110 features are 0.69587, 0.7348, 0.79217, and 0.84164, respectively. Furthermore, only using half selected features (110 features) of the proposed method, the corresponding classification accuracy (0.84168) is approximate to the highest classification accuracy, 0.8795. For other two hyperspectral image data sets, the PAVIA data set and Salinas A data set, we can obtain the similar results. These results illustrate our proposed method can efficiently find feature subsets to improve hyperspectral image classification. One can apply the proposed method to determine the suitable feature subset first according to specific purposes. Then researchers can only use the corresponding sensors to obtain the hyperspectral image and classify the samples. This can not only improve the classification performance but also reduce the cost for obtaining hyperspectral images.

Keywords : hyperspectral image classification, nonlinear feature selection, kernel trick, support vector machine
Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development
Conference Location : Chicago, United States
Conference Dates : December 12-13, 2020