ABSTRACT

Neural networks (NN) have been found to have good generalization properties and their use is becoming increasingly prevalent in the field of remote sensing and in particular for image classification [1]. However, the type of input to be considered for the algorithm in order to maximize the information available from the measurement is still an open issue. Using the mere spectral signature with no pre-processing is not an effective choice and many authors choose the option of the ratioing between bands to increase the signal to noise ratio level. Another point regards the use of textural features to improve the classification. Texture is undoubtedly one of the main approach to recognize the content of a scene and different texture feature extraction methods exist: statistical, geometrical (including structural), model-based, and signal processing. In any case the implementation of texture features in a classification algorithm involves taking decisions on how many and what specific features should be considered. More in general, minimizing the number of inputs of a neural network algorithm, avoiding significant loss of information, affects positively the NN mapping ability and computational efficiency. A network with fewer inputs has fewer adaptive parameters to be determined, which need a smaller training set to be properly constrained. This leads to a network with improved generalization properties providing smoother mappings. In addition, a network with fewer weights may be faster to train. All these benefits make the reduction in the dimension of the input data a normal procedure when designing NN, even for a relatively low dimensional input space.

In this paper we propose a new methodology facing with the aforementioned problems and providing a solution to them. The approach combines three processing aspects: the Tasseled Cap (TC) transformation, the Grey Level Co-Occurrence Matrix (GLCM), and neural networks. Tasseled Cap transformation, like other linear transformation such as Principal Components Transformation or Maximum Noise Fraction (MNF), specifically emphasizes the inherent data structures, and is intended to be an invariant transformation which can therefore be applied directly to any scene [2]. Despite of the other linear transformation TC transformations are not related with the statistics of the image, and not so extremely scene dependent. From this point of view, Tasseled cap transformations, are considered a more efficient compromise between data reduction and the preservation of relevant image transformation. The sum of the textural features to Tasseled Cap transformation can make an improvement of the classification. It was pointed out by Shanmugan et al [3]. that textural
features derived from GLCM are most useful for analyzing the contents of a variety of imagery in remote sensing while, according to Treitz et al. [4], statistical texture measures are more appropriate than structural in traditional land cover classification. Indeed the GLCM is widely accepted for classifying texture and several studies have used it for land cover classification with SAR data[5]. Finally the utilization of neural networks is twofold. Besides being considered for the classification task, neural networks have been exploited to select the best input combination for the classification algorithm. This is obtained via an extended pruning procedure. Normally a pruning procedure is applied to a trained net to get its structure optimization. According to this kind of procedure a network is examined to assess the relative importance of its weights, and the least important ones are deleted. Typically, this is followed by some further training of the pruned network, and the pruning and training may be repeated for several cycles. The most critical choice in the procedure implementation is how to decide which weights should be removed. To do this, we need some measure of the relative importance, or saliency, of different weights. Results that will be shown have been obtained applying the simple concept that small weights are less important than large weights and using the magnitude of a weight value as a measure of its importance. In order to select the most convenient inputs for the chosen context of application it is possible to prolong the pruning procedure to the input layer. According to this method, the first inputs to be removed (an input or hidden unit is removed when it has lost all its connections) should coincide with those containing less information [6].

The dataset considered in this work consists of about 20 LANDSAT TM/ETM+ images, acquired on urban areas belonging to very different countries in the world such as Australia, Austria, China, France, Germany, Italy, Mexico, Netherlands, U.S.A., U.K.. To prove the efficiency of this methodology with respect to the currently most advanced automatic approaches, the overall accuracies and the relative confusion matrices for different input data set were compared with the performance provided by the eCognition Software. A particular attention has been dedicated to the robustness of the algorithms, testing them on images not considered in the training phase.

REFERENCES


