International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-3, Issue-11)

Research Article | November 2014

Classification of Remote Sensing Data Using Support Vector Machine and Random Forests Classifier with Color, Textural and Shape Features

[1]PK Kumaresan, [2]SG Balakrishnan, [3]V Amirthalingam, [4]S Karthik, [5]S. SENTHIL
[1] Prof CSE, VMKV Engg College, Salem, Tamil Nadu, India
[2] Asso Prof CSE, VMKV Engg College, Salem, Tamil Nadu, India
[3] Asst Prof CSE, VMKV Engg College, Salem, Tamil Nadu, India
[4] Asso Prof IT, VMKV Engg College, Salem, Tamil Nadu, India
[5] Asst Prof IT, VMKV Engg College, Salem, Tamil Nadu, India

Abstract:

Computerized information extraction from remotely sensed imagery has been applied successfully over the last two decades. The data used in the processing have mostly been multispectral data, and the statistical pattern recognition methods are now widely known. Within the last decade, advances in space and computer technologies have made it possible to amass large amounts of data about the Earth and its environment. The data are more and more typically not only spectral data but also include, for example, forest maps, ground cover maps, radar data, and topographic information such as elevation and slope data. The accurate classification of remote sensing images is an important task for many practical applications, such as precision agriculture, monitoring and management of the environment, and security and defense issues. The advent and growing availability of hyper spectral imagery, which records hundreds of spectral bands, has opened new possibilities in image analysis and classification. In this work I presented a Support Vector Machine (SVM) and Random Forests (RF) based classification of remote sensing data using color, textural and shape features. The proposed work is experimented with the images from Google Maps to assist a normal person to classify the satellite / airborne images. The performance is evaluated with the recently proposed classifiers to state the power of the proposed systems

Key Terms: Support Vector Machine (SVM), Random Forests (RF)

I. INTRODUCTION

Remote sensing applications, such as land cover classification, provide a variety of important information for decision support and environmental monitoring systems. When complex environments are mapped or when very detailed analyses are performed, the spectral and spatial resolution requirements can be very high, e.g., in urban area mapping, the characterization of mineral composition, or in plant-type differentiation (Plaza et al., 2009; Goetz, 2009). In such situations, airborne hyper spectral sensors are—at the moment—probably the most valuable single data source. Data from these sensors provide detailed and spectrally continuous spatial information on the land surface, ranging from the visible to the short-wave infrared regions of the electromagnetic spectrum. They enable discriminating between spectrally similar land cover classes that occur at highly frequent spatial patterns. However, it is well known that increasing data dimensionality and high redundancy between features might cause problems during data analysis, e.g., in the context of supervised classification: The overall map accuracy can decrease when only a limited number of training samples are available.

Each data set has its own properties, defining its ability to deal with different natures of classes. The first main consideration is the complementary characteristics of the data. It has a consequence in the discrimination ability of such a feature, as will be seen in the experiments. The use of spectral information can be critical for classification of non-structured information in urban areas, e.g., vegetation and soil classes while the use of spatial information can be useful for classification of structured objects, e.g., road and building classes.

II. PROPOSED METHODOLOGY

However, the existing SVM based classifiers are accurate; they were not working on hyper spectral images at different zoom level and different set of features. The spectral images are either classified with pixel based information or with texture features. Here a the SVM and RF classifier proposed to work with both textural and shape features to improve the accuracy. Moreover, the application is built over with Google Maps, so the user can interactively choose the spectral images and the classification could be applied at any zoom level.

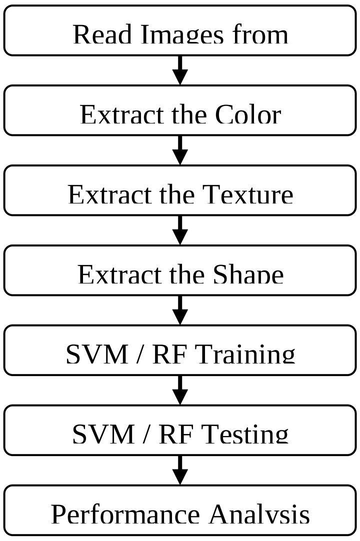

Fig.1.1 The Proposed Framework

III. SVM CLASSIFICATION APPROACH

3.1 SVM Mathematical Formulation

Let us consider a supervised binary classification problem. Let us assume that the training set consists of N vectors from the -dimensional feature space$x_i \in R^d$ $(i = 1,2, \dots. N)$. A target $y_i \in \{-1, +1\}$ is associated to each vector. Let us assume that the two classes are linearly separable. This means that it is possible to find at least one hyperplane (linear surface) defined by a vector $W \in R^d$ (normal to the hyperplane) and a bias $b \in R$ that can separate the two classes without errors. The membership decision rule can be based on the function sgn[f(x)], where [f(x)] is the discriminant function associated with the hyper plane and defined as $\boldsymbol{f(x) = w \cdot x + b}$ (1) In order to find such a hyper plane, one should estimate w and b so that $\boldsymbol{y_i(w \cdot x_i + b)} > 0$(2)The SVM approach consists in finding the optimal hyper plane that maximizes the distance between the closest training sample and the separating hyper plane. It is possible to express this distance as equal to 1/||W|| with a simple rescaling of the hyper plane parameters w and b such that$\mathbf{min}_{i=1,2,\dots,N}\, \boldsymbol{y_i(w \cdot x_i + b)} \geq \mathbf{1}$. **(1)**

The geometrical margin between the two classes is given by the quantity 2/||w|| . The concept of margin is central in the SVM approach, since it is a measure of its generalization capability. Accordingly, it turns out that the optimal hyper plane can be determined as the solution of the following convex quadratic programming problem:$\begin{cases} \text{minimize: } \frac{1}{2}\|w\|^2 \\ \text{subject to : } y_i(w \cdot x_i + b) \geq 1, \quad i = 1,2, \dots, N \end{cases}$

This classical linearly constrained optimization problem can be translated (using a Lagrangian formulation) into the following dual problem:$\begin{cases} \text{maximize: } \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \\ \text{subject to : } \sum_{i=1}^{N} \alpha_i y_i = 0 \text{ and } \alpha_i \geq 0, \ i = 1,2, \dots, N \end{cases}$ The Lagrange multipliers α_i's $(i = 1,2, \dots, N)$ expressed can be estimated using quadratic programming (QP) methods. The discriminate function associated with the optimal hyper plane becomes an equation depending both on the Lagrange multipliers and on the training samples.

3.2 SVMs in Hyper spectral Feature Spaces

Unlike traditional learning techniques, SVMs do not depend explicitly on the dimensionality of input spaces. They solve classical statistical problems such as pattern recognition, regression, and density estimation in high-dimensional spaces (Benediktsson et al., 2010a). A kernel transformation into a higher dimensional space, where it is expected to find a linear separation that maximizes the margin between the two classes.

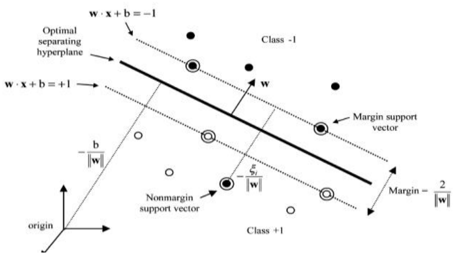

Fig. 3.2.1. Optimal separating hyper plane in SVMs White and black circles refer to the classes “+1” and “-1,” respectively. Support vectors are indicated by an extra circle.

In order to appreciate the potentialities of SVMs in high-dimensional spaces, it is useful to recall the statistical and geometrical properties of the data in such spaces. First, in a hyper spectral space, normally distributed samples (a reasonable assumption for optically remotely sensed data) tend to fall toward the tails of the density function with virtually no samples falling in the central region.

3.3 SVM Multiclass Strategies

As stated in the previous section, SVMs are intrinsically binary classifiers. However, the classification of hyper spectral remote sensing data usually involves the simultaneous discrimination of numerous information classes.Let $\Omega = (\omega_1, \omega_2, \ldots, \omega_T)$ be the set of T possible labels (information classes) associated with the d-dimensional hyper spectral image of the study area. In the multiclass case, the problem is to associate to each d-dimensional sample x the label of the set that optimizes a predefined classification criterion. In order to carry out this task, the general approach adopted in strategies based on binary classifiers consists of: 1) defining an ensemble of binary classifiers; and 2) combining them according to some decision rules.

A. Parallel Approach

1) One-Against-All Strategy: The One-Against-All (OAA) strategy represents the earliest and most common multiclass approach used for SVMs. It involves a parallel architecture made up of SVMs, one for each class (Fig.3.3.1). Each SVM solves a *two*-class problem defined by one information class (e.g., $\omega_i \in \Omega$) against all the others, i.e.,

$$\begin{cases} \Omega_A = \omega_i \\ \Omega_B = \Omega - \omega_i \end{cases} \quad \textbf{(3)}$$

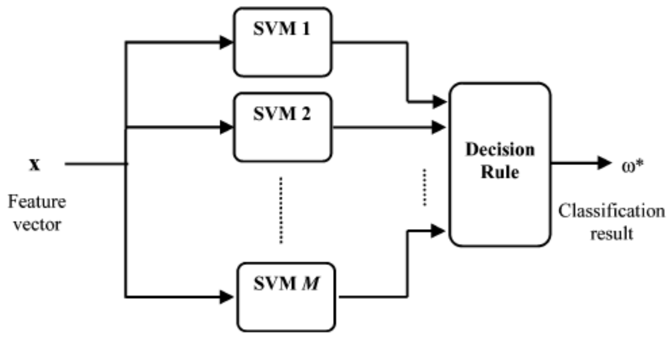

Fig.3.3.1. Block diagram of a parallel architecture

2) One-Against-One Strategy: The main problem of the OAA strategy is that the discrimination between an information class and all the others often leads to the estimation of complex discriminate functions. In addition, a problem with strongly unbalanced prior probabilities should be solved by each SVM.

IV. RANDOM FOREST CLASSIFIER

Random Forests (RF) are a variant of bagging proposed by Breiman (2001). It is a decision-tree-based ensemble classifier that can achieve classification accuracy comparable to boosting (Breiman, 2001). It was successfully applied to multispectral data, multi temporal SAR images, or multi-source data, where Land sat MSS and topographic data were used. In this work, exploit this property on multi-source data in order to measure the relevance of remote sending image features for classifying the maps.

4.1 Margin Definition

The margin concept of ensemble learning methods was first proposed by Schapire et al. (1998) to explain the success of boosting type algorithms. The concept was then generalized to analyze other types of ensemble classifiers (Breiman, 2001). The importance score for a feature *f* is then computed as the mean importance over all trees:

$$FI(f) = \frac{\sum_T FI^{(t)}(f)}{T}$$

where T is the number of trees. Suppose that the training samples consist of pairs of the form (x_i, l_j), where x_i is an instance and l_j its true label. The margin m_i of instance x_i is computed as follows (Tang et al., 2006):

$$m_i = m\arg in(x_i, l_i) = \frac{v_{i,l_j} - \sum_{c \neq l_j} v_{i,c}}{\sum_c v_{i,c}}$$

where v_i, l_j is the number of votes for the true class l_j, and $v_{i,c}$ is the number of votes for any class c with $c \neq l_j$. Hence, the margin is given by the difference between the fraction of classifiers voting correctly and incorrectly. It measures the strength of the vote. The margin ranges from −1 to +1.

The positive margin value of a sample indicates that this sample has been correctly classified, whereas a negative value means that the sample has been misclassified. The larger the margin, the more the confidence in the classification. A value close to 0 indicates a low confidence in the classification.

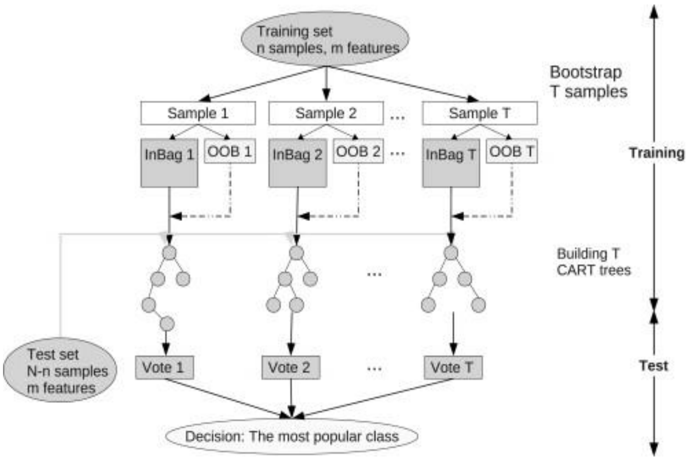

Figure 4.1.1 General Procedure of Random Forests Classifier

V. EXPERIMENTAL RESULTS

This section describes the experimental validation of our developed system by using satellite/airborne images obtained from Google Maps across different locations. Initially the image features are extracted and fed to the classifiers (SVM / RF) for the learning. Then the classifiers are tested with the samples. The expected outcomes are given as:

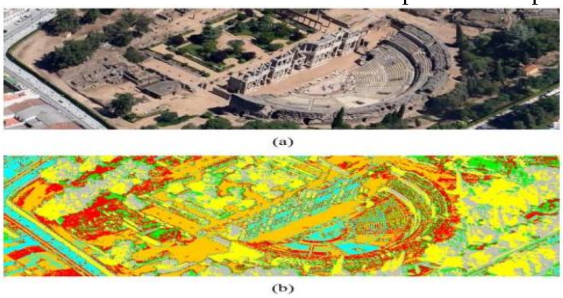

Fig. 5.1 (a) Satellite Image, (b) Classified Image

VI. CONCLUSION

This paper has described a new desktop application for supervised classification of remotely sensed images. The system has been developed using the MATLAB®. Our experimental results, conducted by comparing the obtained classification results with the existing methods like k-means and ISODATA algorithms, reveal that the proposed tool provides classification maps of high similarity with regards to those provided by the existing algorithms, but with the possibility to perform classification of any image portion available in Google Maps™ engine in supervised fashion. In future developments, plan to incorporate additional feature extraction techniques such as attribute morphological profiles, and also like to extend the developed tool with the incorporation of content-based image retrieval functionalities.

REFERENCES

[1] Ferrán, Á., Bernabé, S., Rodríguez, P. G., & Plaza, A (2013). A Web-Based System for Classification of Remote Sensing Data, IEEE Journal of Selected Topics in Applied Earth Observations And Remote Sensing, 6(4), 1934 – 1948.

[2] Jun G, and Ghosh J (2011), “Spatially adaptive classification of land cover with remote sensing data”, IEEE Transactions on Geoscience and Remote Sensing, vol. 49(7), pp. 2662-2673.

[3] Li W, Guo Q, and Elkan C (2011). “A positive and unlabeled learning algorithm for one-class classification of remote-sensing data”, IEEE Transactions on Geoscience and Remote Sensing, vol. 49(2), pp. 717-725.

[4] Tarabalka, Y, Chanussot, J, and Benediktsson, J. A. (2010a), “Segmentation and classification of hyperspectral images using watershed transformation”, Pattern Recognition, vol. 43(7), pp. 2367-2379.

[5] Plaza A, Benediktsson J. A, Boardman J. W, Brazile J, Bruzzone, L, Camps-Valls, G, and Trianni G (2009), “Recent advances in techniques for hyperspectral image processing”, Remote Sensing of Environment, vol. 113, pp. 110-122.

[6] Tarabalka Y, Benediktsson, J. A, and Chanussot J (2009), “Spectral–spatial classification of hyperspectral imagery based on partitional clustering techniques”, IEEE Transactions on Geoscience and Remote Sensing, vol. 47(8), pp. 2973-2987.

[7] Breiman L, “Random forests”, Machine Learning, vol. 45(1), pp. 5–32 Chun Y. D, Kim N. C, and Jang I. H (2008), “Content-based image retrieval using multiresolution color and texture features, IEEE Transactions on Multimedia, vol. 10(6), pp. 1073–1084.

