AN INTEGRATED UNSUPERVISED CHANGE DETECTION IN MULTISPECTRAL IMAGES USING SVM AND NEURAL NETWORKS

M. Aghamohamadnia, Y. Jouybari Moghaddam

Division of Remote Sensing Eng., Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran

KEY WORDS: Change Detection, SVM, Neural Networks, Multispectral, Image Fusion

ABSTRACT:

Importance:
Satellite Remote sensed data and GIS for land cover, land use and its changes is a key to many diverse applications such as Environment, Forestry, Hydrology, Agriculture and Geology. Natural Resource Management, Planning and Monitoring programs depend on accurate information about the land cover and its changes along the time in a region. Methods for monitoring changes varies from intense sampling from field with plot inventories to vast analysis of remotely sensed data which has proven to be more cost effective for large areas, small site assessment and analysis. A number of change detection methods have been proposed in the past, among which are efficient based on model techniques such as the Generalized Likelihood Ratio (GLR) stochastic test.

Introducing of the problem along with data used:
Here we regard the case where not accurate nor can tractable model be found, by a model-free method, namely as the Kernel change detection (KCD). The reliability of support vector machines for classifying multispectral data in the remote sensing has been proven in various researches. Depending on the soft margin single-class Support Vector Machine (SVM); a un-similarity measure in feature space between those sets, without estimating densities as an intermediary step is imported.

Method used:
A method based on SVM and artificial neural network (ANN) was developed to detect newly countryside urbanized areas illustrated in satellite sensor data. This method uses a pair of Landsat Thematic Mapper (TM) images taken from a region and acquired on different dates as input and supervises the SVM system to classify the image data into 'from to to' classes and also implies a parameter adjustment using the ANN method. Furthermore, The Levenburg-Marquardt algorithm was used to accelerate the ANN’s convergence. The high-dimensional feature vectors of multi spectral data often impose a high computational cost as well as the risk of “over fitting” when classification is performed. Therefore it is necessary to reduce the dimensionality through ways like feature selection. In this paper, we present a classification of multi spectral data using Support Vector Machine (SVM), a state-of-art classifier that has found success in a variety of areas along with the neural network (NN), which seeks to solve optimization problems using the methods of neuron connection, specifically implying hidden layers which was used to optimize both the feature subset, i.e. band subset, of multi spectral data and SVM kernel parameters simultaneously.

Analysis and conclusion:
An empirical result from a case study shows that the ANN-based method requires not so much training time but can be somehow 20-30% much more accurate than post-classification comparisons. The outputs imply that the practical value of ANN-based change detection is efficient and optimized. Comparison of the optimized results and the non-optimized results implies that the Neuro-SVM approach could significantly reduce the computation time and cost while raising the classification accuracy. The number of bands used for classification was reduced, while the classification and detection accuracy increased considerably. In addition and conclusion, the proposed Neuro-SVM can optimize feature subsets change detection and SVM kernel parameters at the same time, therefore can be applied in feature selection of the hyper spectral data.
1. INTRODUCTION

According to the vast number of available techniques, remote sensing images have the most feasible and efficient tool for land cover change detection. In common, changes are detected by comparing multiple images of the same terrain acquired at different times. However, although they are less accurate compared to modern pattern recognition methods, conventional change detection approaches such as differencing, rationing and spectral vector analysis have been used in various applications. Therefore, recently, interests have been focused on machine learning and artificial intelligence (AI) techniques such as Artificial Neural Networks (ANNs), Ant Colony Optimization (ACO) and Support Vector Machines (SVM) since they perform more nonlinear and outperform conventional methods in complicated cases [1], [2]. In this work, a straightforward and computationally interesting approach, called Support Vector Machines (SVMs), is presented along with ANN and conventional method all together integrated into a framework. These methods separately have been successful in pattern recognition [3]. The efficiency of SVMs lies in their strong connection to the statistical learning theory, where they implement the structural risk minimization for solving two class classification problems [4]. This issue states a better solution, in terms of generalization theory, can be found by minimizing an upper bound of the generalization error. This bound is formed by the sum of training errors and a term depending on the discrimination capacity of the learning machine given in its Vapnik Chervonenkis (VC) dimension.

Besides this, we have used the Multi-Layer Feed Forward Back Propagation Neural Network (MLFFBPNN) in part with conventional methods to facilitate the requirements and training data for input of SVM system. In remote sensing, SVMs are specially used in classifying hyper-spectral images [5], [6], as well as for modelling spectral mixtures [7]. This paper describes the use of SVMs for land cover change detection purposes. We are interested in extracting changes and our work is restricted to binary classifying SVMs implanted to specify either all land cover transitions or some kinds of change. However, when using SVMs the user faces many possible choices of kernel functions, its parameters and more importantly the training set quality for learning which are commonly yielding different results.

Hence, the core contribution of this paper is to facilitate the basic requirements of SVM system and prepare an unsupervised algorithm for change detection purposes. This in fact, implements a combination and integrated framework for SVMs in order to enhance change detection accuracy.

The rest of the paper is organized as follows: Section 2 briefly reviews the basic theory of SVM, MLFFBPNN and incorporated conventional method. Section 3 emphasises the integrated SVM - ANN change detection system, as well as the integration strategy and details. Section 4 conducts an experiment to illustrate the performance of SVM - ANN system. Finally, after a brief suggestion of possible future works in Section 5, some conclusions are given in Section 6.

2. THE BASICS AND PRINCIPLES OF SVM AND ANN

SVMs are appropriate for digital image classification. The SVM based approach to classification aims to find the most optimal discrepancy hyperplane between classes by focusing on the training data that lie at the edge of the class distributions, the support vectors, with the other training cases effectively and mostly ignored [7], [8]. Thus, not only it is an optimal fitted hyperplane but also the method may result to a high accuracy with small but carefully and specifically designed training set [9]. Given that the cost of training data acquisition is often noted as a concern in remote sensing [10], [11], [12], the ability to use small training sets could be an advantageous and attractive feature for some users. The idea of using small train data comes from a central issue of the SVM method to classification, that only the training sets which are both on the class boundaries and between the class distributions in feature space are necessary for accurate separation.

Beside these, neural networks have attracted the international research community due to its impressive properties such as learning capability, ability to generalize and adaptability. The process of training a NN is generally focused in neighbouring the individual weights between each of the adjacent neurons. At the beginning of the learning process a dataset, which is named as training set, is presented to the inputs to determine the correct outputs. Training of NN is an iterative process that continues until achieving the output value that is close to the desired output by adjusting the network weights accordingly. One of the popular algorithms for training in the NNs is the back-propagation method, which is a gradient descent approach in order to minimize the mean-squared error between the training and the actual outputs for the particular inputs of training set to the networks. But BP has some shortcomings: the first is the need of a differentiable neuron transfer function and the second one is the high possibility of converging into local minima. NNs generate complex error surfaces with multiple local minima and BPs tend to become trapped in local solution that is not global. As mentioned before, it is possible to implement change detection process both using the SVM and ANN algorithms separately but they may show weakness with results. SVM performs effectively with proper and marginal support vectors and NN models the problem space accurately with high amount of training data and hidden neuron layers which is in fact may be impractical, time consuming and not economic. These issues are addressed in fore coming section.

3. THE EXPLANATION OF METHOD AND THE ALGORITHM

As cited in previous section, there are shortcomings with each approach and it is impossible to construe information of change detection in an optimized state. Here we seek to benefit the helping properties of NN and incorporate them into the SVM system. The algorithm is unsupervised and needs to be trained automatically. The first step in the algorithm is using of a conventional method such as rationing and preparing a reliable training data set to be used in the NN for change detection process. This data set is chosen according to specific criterion and threshold. This threshold has to be chosen in a way that extracts and cuts out highly reliable data set as training samples. It would be more appropriate to separate the training samples using a statistical threshold related to whole data set, instead of constant one. Since the number of pixels to be examined in ratio analysis is relatively high, and radical changes are behaving as outliers or very low probable phenomena, using of a normal distribution for modelling of ratio output seems appropriate. So, the distribution is fitted to change ratios between two dates. After this stage, it is optimum to mark the first derivation of mean (σ) as the cut probability to extract semi-radical and radical changes to formation of change training set. By this selection, it is possible to prepare a relatively dependable data set for training. But it is important to mention that neither this selected set nor any other samples from the ratio analysis can be used in SVM system for an optimum result to be acquired. The reason for this is the difference of the nature of rationing or other conventional analysis compared to SVM or ANN. In this perspective, the degree of nonlinearity of rationing and other
conventional methods are way less than SVM and NN. This difference in complexity obligates us to choose a proper analyser system for input providing stage of SVM. Next step is incorporating this separated data set into NN. By this rationing change detection, it is possible to model the change space in a cost efficient and with rapid implementation so to produce rather proper and appropriate training data for a relatively more complex analysis of NN. The NN architecture is designed multi-layer with back propagation weight updating policy. The architecture design is illustrated in Fig.1 in which the first layer neurons consists of both dates images in all available spectral bands and the last layer is composed of two neurons addressing change or no change states of each input pixel. There exist two hidden layers of constant number of neurons depending on the required complexity according to scene non-linearity.

The goal here in this step of NN analysis after training and modelling is an overall estimation or parameterization of weights of last layer of neurons. As illustrated in the Fig.1, there are weights twice the number of last hidden layer neurons and while half of them are summed into the first neuron of the output layer, the other half is gathered in the other neuron. The main attention in this step is evaluation of these two sets of weights of each pixel, in order to reach a decision if this pixel is a marginal and close to two class (Change/No Change) border sample or it is a far from border and confident one. For this purpose, it is convenient for analysis to give a new form to cumulate of weights of both output neurons as below:

\[ C_q = \frac{\sum_{q=1}^{M} w_q}{\text{Thr}} \]

In which, \( C \) stands for the cumulate of weights parameter of each output layer neuron, \( w_q \) is weight originated from \( i^{th} \) neuron of last hidden layer and headed to neuron 1 of the output layer, \( \text{Thr} \) is threshold value of activation function of the output layer and \( M \) is number of neurons in the last hidden layer. Every member of change or no change class show 1-0 or 0-1 results in the output layer as shown in Fig.2.

In this stage it is important to cross a proper criterion to separate off-border and close ones to border, which in this level the normalized \( C \) of each output neuron are evaluated by this criterion. It can be described as below:

\[ C_i \leq \sigma_i \quad \text{for } i=1 \text{ or } 2 \]

In which, \( \sigma_i \) is derivation of mean of \( C \) values of all pixel. So the samples reaching above term are nominated to be in the list of SVM training data.

At the third step, by introducing the SVM with the selected data set which is not contained with confident and definite pixels of each classes of Change or No Change, it is possible to aim for a more reliable and less time consuming non-linear classifier.

4. THE DATA USED AND RESULTS

The method described above must be applied on a dataset in order to see its impact and results in a real case. The data used in this paper is from Landsat5 satellite, TM sensor in two dates of July 2001, Fig.3, June 2005, Fig.4 and was taken from the USGS [13] website. The region consists of urban and country areas and since it was hard to find an exact and archived sample of change detection map from related organizations, the truth data was produced using manual and visual change detection analysis by researchers themselves. The test images have been cropped out of original image with 380×1200 spatial dimension which are in spectral bands of 1-5 and 7 of TM sensor images. It also covers an approximately 11.4×36 km region. The polynomial model is used for geometrical registration of corresponding temporal images. By rationing between corresponding individual bands of both dates and calculating RMS (Root Mean Square) of results, it is possible to form a single image of rationing of the whole bands Fig.5. In this stage we extracted needing training data (6,245 pixels in this case) of next step using the normal distribution analysis described in previous section. In the second step the training data was put into process of MLFFBPNN and the next step training data resulted (217,056 pixels) Fig.6, as discussed in Section 3, NN consisted of two hidden layers each with 20 neurons and an input layer of 12 neurons (6+6) and a 2 neuron output layer.

Finally, this training data set was given to SVM system with the selected data list which is not contained with confident and definite pixels of each classes of Change or No Change, it is possible to aim for a more reliable and less time consuming non-linear classifier.
Figure 3. First date original image in visible view

Figure 4. Second date original image in visible view

Figure 5. Single image of rationing of the whole bands.

Figure 6. Training data of MLFFBPNN to be used in the SVM system.

Figure 7. The SVM output.
5. FUTURE WORKS

In the future works it would be a suggestion to work on comparing the other useful methods to be substituted instead of the steps taken in this algorithm. It also worth considering improvements and evaluation of changes in each steps parameters.

6. CONCLUSION

There have been various approaches of determining changes in an area. Most of them are placed in the image processing areas at different levels such as pixel, object or image basis. Meanwhile some of them are concentrated on using of AI (Artificial Intelligence) and evolutionary algorithms which are beneficial in complex and process time consuming scenes. Unlike common attitude toward using multiple evolutionary algorithms into a hybrid framework in which one of them solves the system and the other one supports the solver parameters optimization, our algorithm performs solving directly and independently of each other in specific sequence of problem solving system. By this, it is possible to train an evolutionary method using other AI algorithms. Thus sequence of different methods and incorporating several techniques results in a more robust and intelligent system. We first used rationing between corresponding individual bands of both dates and calculated RMS (Root Mean Square) of results. By this, we extracted needed training data of next step using the normal distribution analysis described in previous section. In the second step the training data was put into process of MLFFBPNN and the next step training data resulted and eventually, this training data set was given to SVM system for final changes pixels.

7. REFERENCES


