International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-6, Issue-6)

Research Article **June 2017**

Hybrid Algorithm for Deforestation Detection Using Satellite Data by Using Support Vector Machine (SVM) Algorithm

J Suresh Babu*
Department of Computer Science,
Vikrama Simhapuri University,
Andhra Pradesh, India

Prof. T. Sudha
Department of Computer Science,
SPMVV, Tirupathi,
Andhra Pradesh, India

Abstract:

There are many applications of Remote sensing Satellite images like (astronomy, military, forecasting, and geographical information). Using satellite remote sensing data sets we developed the mapping forest area cover change. This kind of multiple improvement and identifying methods have a Training Data Automation algorithm which is used for advanced vector machines procedure. This TDM technique capable of automatically generating exact image enhanced patches. The obtained high resolute training data allow in producing the dependable forest cover change products with the help of SVM. This process was tested in study areas selected from major forest areas across the globe. In each area, a forest cover change map was produced using a pair of real time Land sat images acquired around 1999 and 2015.

Keywords: ***Satellite Image, Image enhancement, Techniques for Image Enhancement, TDM, SVM.***

I. INTRODUCTION

Many environmental applications all over world require land use land cover information about the Earth's surface. The technologies related to remote sensing are capability of seeing the Earth's surface with different spatial, spectral and temporal resolutions and classifications.These kind of technologies find many advantages with respect to time and cost in comparison with land surveying procedures. In perspective of every one of these elements, remote detecting has a critical information source to concentrate land utilize/land cover data. With the advancement of remote detecting innovation, remotely detected information have been generally used to order arrive cover, allowing to refresh maps all the more as often as possible and on a close constant premise.

The classifying of remotely sensed images is a very important information for the estimation and determination of land use data/land cover data information. Digital image identification uses spectral data represented by the digital intensity values (i.e. pixels) in one or many spectral bands and attempts to identify each and every individual pixel on the basis of this spectral information. This kind of classification is termed spatial satellite classification/identification. In either case, the objective is to assign all pixels in the image to particular classes or themes (e.g. water, forest, corn, wheat). The output classified image is composed of a existing of pixels, each of which belongs to a different theme, and is essentially a thematic mapping of the original image. Many classification algorithms have been developed from the first Land sat image was acquired in 1972. Among all those most popular and repeatedly used is the maximum likelihood (ML)classifier.

It is based on parametric procedure that assumes the signature of the class in a normal distribution. In fact this assumption is valid generally, but it is not valid for classes which possess several subclasses or classes having verity of spectral features.

II. TEST SITE AND DATA

This study area has chosen for the research that covers round about 772 km2 area which is in the Gebze district of Kocaeli province, Turkey. It is in between longitudes 408450 0800E and 418020 3800E and between latitudes 298190 5600N and 298450 1400N.

At the border of Istanbul Gebze is situated which is said as an industrial town, the city with most in Turkey, on the northern shore of the Marmara Sea. Land sat ETM+ (1997) and Terra ASTER (2002) images which covers the study area were used to distinguish land cover and types of land used. The registered images were to the UTM projection system where first-order enhancement is applied using ERDAS/ Imagine software package. Ground control Co-ordinates points were obtained from 1:25,000 scale topographic maps.

All distributed control points at root-mean-square error about 0.5 pixels.A closest neighbor sampling method was used to distinguish the values of the new pixel. Verity of forestry maps and aerial photographs were taken in 1996, 1999, and 2003 were used in creating ground reference data maps for 1997 and 2002 pictures. Along with that to collect information about ground reference the field surveys were done using GPRS. After a large field survey, the decision has made that the study area importantly covered six land cover classes, they are water, pasture, coniferous trees.It should be noted that the urban group contain settlement, buildings, and roads whilst the bare soil class includes soil and rock covered lands. Under the land of sustainable improvement, and stimulated by the joint throughout world Land Use/Cover

Change project of the world wide Geosphere - Biosphere Program and the worldwide Human magnitude Program on Global ecological Change (Turner et al, 1995), detect, monitoring, understanding, modeling, and projections of land-use change from the global to the regional scale have done many investigate welfare.

In the past 20 years, sizeable vocation has been done with consider to land-use-change modeling. Number of models have been developed to help ecologists, urban planners, sociologists, administrators, and policy makers understand better the complexity of land-use-change patterns and to evaluate the impact of land-use change on the environment.

Study area

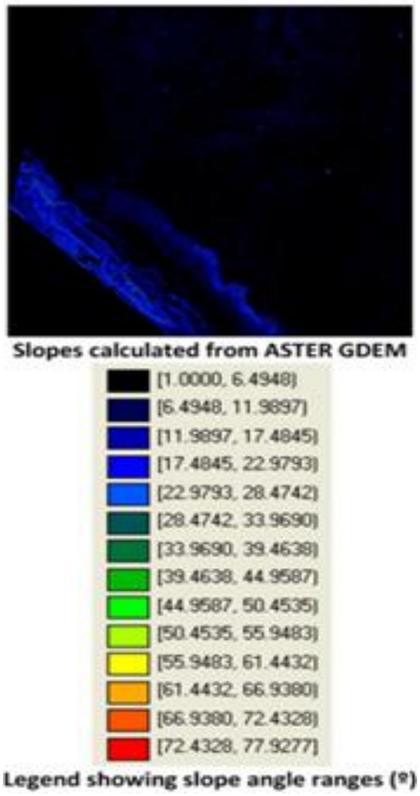
The learning region is situated in the section of Bolivia. We selected this large area because its landscapes are highly heterogeneous as a transition across three bio geographic areas:

(1) Montage tropical forests covering the foothills (over 400 m) of the Andes to the west.
(2) At the south lowland tropical forests and centre of the study area, and
(3) To the north and east wet savannah areas plain forest are located below 400 m and includes some deciduous species owing to a marked seasonality (dry seasons).

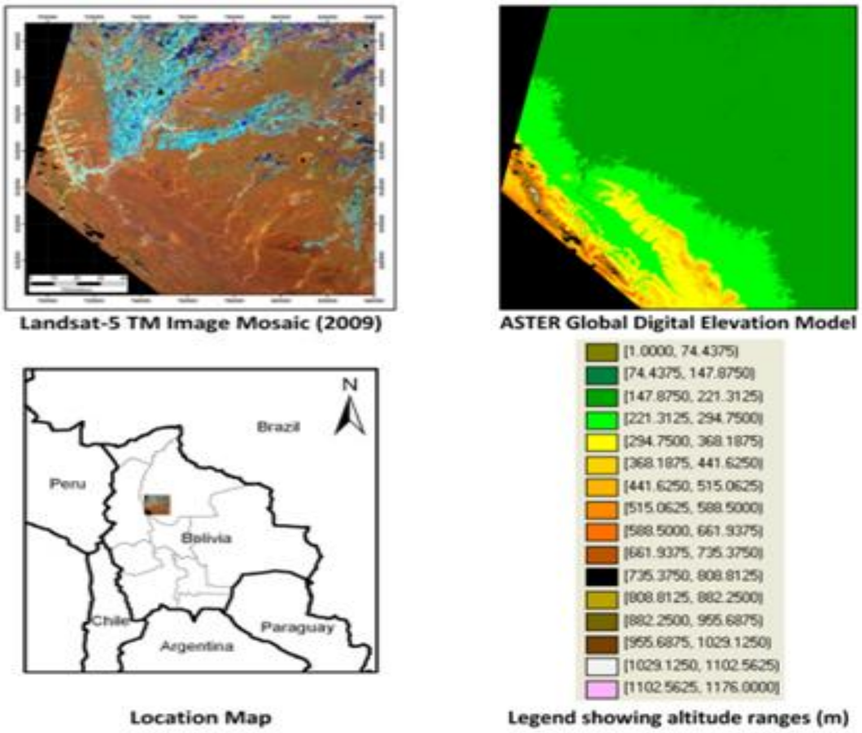
In wet savannas plants is controlled by small variations the earth distance from the ground and relief, which in turn are shaped by periodic flooding and river dynamics. Savannah areas consist of swamps, lagoons, and marshes with proper vegetation in the lowest areas; pastures and semi-natural grasslands in areas less prone to be flooded; and patches of forests and scrublands on mounds that do not get flooded seasonally. The vegetation formations of the study area are shaped even by the land use type and intensity of its different inhabitants, who range from Andean indigenous population in montage forests to local peasants, cattle ranchers, and various native and colonist indigenous peoples in savannah areas and lowland forests.

Materials and methods

LUC classifications were done on Land sat satellite mosaics composed of 2 scenes. We chose Land sat data because we needed to cover large extent and because Land sat is questionably the worlds for the most part commonly worn satellite to undertake ecological studies, including LUC classifications (Cohen and Goward, 2004), which renders our results more comparable to persons from former study. In calculation, to carry topographic and illumination corrections out we used the ASTER Global Digital Elevation Model (GDEM) possesses 30 m horizontal and 20 m vertical accuracies at 95% confidence and therefore it is more accurate than the older DEM provided by the Shuttle Radar Topography Mission (SRTM). The 2 25/08/2001 Landsat-7 Enhanced Thematic Mapped (ETM+) scenes were acquired through the United States Geological Survey (USGS). Their geometric exactness was assessed through the GPS tracks and ground control points we had collected in the field and we found misalignments of ~0.5 pixels in both cases. The two 17/04/2009 Landsat-5 TM scenes were acquired from the Brazilian National Institute for Space Research (INPE) and compulsory numerical and topographic correction, which were done with Miramon software using the procedure developed by Paland.



Legend showing slope angle ranges (º)



Classification algorithms

We used a non-parametric classifiers (k-closest neighbor KNN, and 4 different support vector machines SVM: linear, polynomial, radial basis function and sigmoid) parametric classifier (maximum likelihood ML), and a hybrid classifier (unsupervised supervised, contained in Miramon software MMHC). We do not explain here how the ML, KNN, and SVM algorithms work because detailed descriptions bound in remote sensing and pattern recognition textbooks.

MMHC involves the use of non-supervised ISO DATA algorithm to retrieve spectral classes, and a subsequent supervised classification done on the ISO DATA results using training areas to obtain thematic classes (Serra et al., 2003). MMHC has been used successfully to classify Mediterranean environments (Serra et al., 2003) and has also used in dry tropical areas of Nicaragua to classify vegetation (Garc a-Milln and More, 2008). According to our knowledge, this is the first time MMHC has used to identify tropical forests and savannas.

Training data

We carefully examined spectral signatures and field data across the images to select the training sets from the two images. In most cases, training data contains of small polygons, though there were few instances in which single pixels were chosen in narrow areas (e.g. rivers, roads, sand banks). Much care was taken to scatter training areas across each image to ensure they were representative of the entire image, and to retrieve as many training samples for each LUC class as needed to fulfill previously suggested conditions to establish an appropriate minimum sample size (also and Mather, 2009).

To improve the comparison of results between the classification of both dates we made use of the same training areas as high as possible (i.e., when no change had occurred). The Jeffries Matusita changed divergence index was used to assess the training data for both dates separability. We confirmed that separability for water was in fact high, bare soil/urban and old-growth forest, but for the other LUC classes it was very low.

Textural data

From the gray level co-occurrence matrix we extracted textural measures (GLCM), which is frequently employed to extract textural information from remote sensing images (Haralick et al., 1973). Specifically, we calculated 8 textural indices from 6 Land sat reflectance bands (15, 7) using moving windows of 3X3 and 7X7 pixels. We then used the 6 textural bands for each index and calculated the moving window size, along with the 6 Land sat spectral bands, classification of maximum likelihood was taken place. The potential usefulness of every index was assessed based on the overall accuracies obtained for every classification and users accuracies attained for the two forested classes and the producers and. All the calculations for both images were performed.

Land-use-change modelling

Causal factors Land-use change is a huge process influenced by a various of natural and humanactivities. Aim of Land-use-change modeling to explore the dynamics and causal factors of land-use change and to inform policies affecting such change. In general, land-use change is influenced by a number of factors which may be social, economic, or spatial variables. Earlier studies convey that single set of factors cannot explain the changes in different places since each context is different from each other. More often than not, land-use studies put forward different causal forces to explain land-use trends in different places. Factor selection should be taken into consideration the context of the region and period to be modeled as well as the intention of the model. Typically, land-use changes are influenced by a few recurrent parameters that cannot be overlooked. Demographic factors (population size, population growth, and population density) are widely treated as major causal factors of land-use change (Verburg et al, 2002). Usually, if population increases the city will automatically grow. Consequently, in close proximity new residential areas will emerge to transportation facilities (roads, railways, and bus lines), and concurrently development in the commercial centres will also be seen. In the meantime, industrial buildings will develop in the vicinity of those which existed previously. On the whole, urban expansion will transform vacant or low-rent areas will turn into the built-up land foe accommodation as the increasing population with necessary resources for their work, living, and entertainment.

Moreover, the agglomeration of improved areas and the availability of exploitable sites will comparatively influence land-use-change patterns. Accessibility is very frequently treated as a very key driving factor for land-use change through its effect on ease of settlement and cost of transportation. SVM modeling framework is an SVM land-use-change using the C programming language the development of modeling framework taken place. The modeling framework was combined into Arc Map as an extension so it can make use of Arc Map's powerful spatial data processing and visualization capacity provides a key factors of the main components of the SVM land-use-change modeling framework.

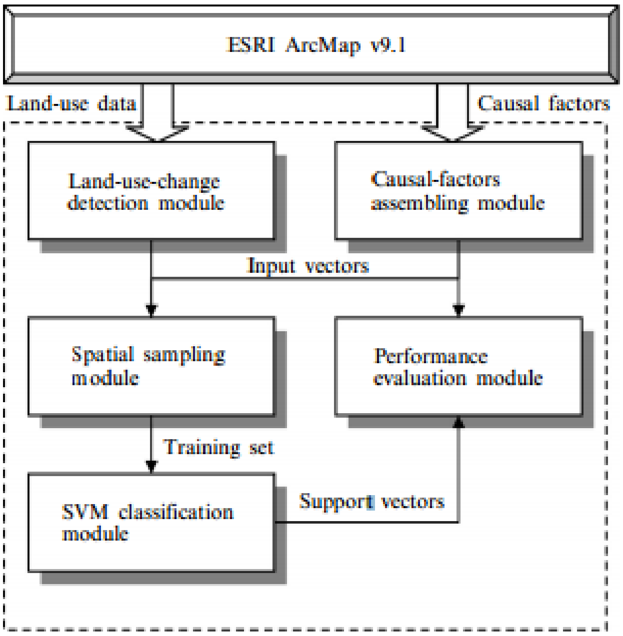


Figure 5. General structure of the SVM land-use-change modeling framework.

III. EXPERIMENT RESULTS

Stepwise variable selection with logistic regression was achieved the lowest estimates of conditional error frequency both for the spatial prediction of present landslides and the spatiotemporal prediction of the future landslides outside the area of training. This is applicable for all training data sets considered. For the largest training data sets best results are obtained, the overall optimum being an error rate of 0.24 for present and 0.32 for future landslides.

Without variable selection this method is followed by logistic regression in the case of spatial prediction, and by logistic regression with spatial autocorrelation structure in the case of future landslides. Average results achieved by SVM as regards estimates of conditional error rates, with better results for present landslides than for future ones comparatively. Regarding the conditional error estimates

Bagging and double bagging perform considerably worse than the other methods.

IV. SOFTWARE INTRODUCTION

MATLAB read A =read (filename, fat) reads a grayscale or color image from the file specified by the string filename, where the string fit specifies the format of the file. If the file is not in the current directory or in a directory in the MATLAB path, specify the full pathname of the location on your system.noise= noise adds zero-mean, Gaussian noise to an image I, where the local variance of the noise, var , is an function of the image intensity values in me.

J = noise adds salt and pepper noise to the image I, where d is the noise density. Filter The function filter2 performs two-dimensional correlation, conv2 performs two-dimensional convolution, and convence performs multidimensional convolution. Each of these filtering functions always converts the input to double, and the output is always double. Apply filter to an image using filter. Morphological operations Morphology is a broad set of image processing operations that process images based on shapes.

Morphological operations apply a structuring element to an input image, creating an output image of the same size. In a morphological operation, the value of each pixel in the output image is based on a comparison of the corresponding pixel in the input image with its neighbours. By choosing the size and shape of the neighborhood, you can construct a morphological operation that is sensitive to specific shapes in the input image.

V. CONCLUSION

This study provides a support vector machine method for urban land-use-change modelling. Experiment results determine that SVMs perform well in deriving the relationship between the causal factors and land-use change in our case study of Calgary land-use change. Clearly, For land-use-change modeling SVMs provide a new and effective option. Experiment results also suggests that the unbalanced SVMs can significantly improve the performance of the standard SVMs for an unbalanced dataset. By slightly sacrificing the overall prediction accuracy, the unbalanced SVMs can dramatically improve the prediction accuracy of the minor data and thus make a uniform performance for predicting both minor and major categories.

That is very attractive for land-use-change modeling since the changed land cells only account for a small part of the whole data set and the accuracy of predicting the changed land cells is relatively important in land-use-change modeling. Our case study illustrates that a well-optimized SVM model outperforms the commonly used spatial logistic regression model with respect to land-use-change modeling.

REFERENCES

[1] Balabanovic, M. and Shoham, Y. (1995). Learning information retrieval agents: Experiments with automated web browsing.In Working Notes of the AAAI Spring Symposium Series on Information Gathering from Distributed, Heterogeneous Environments. AAAI-Press.
[2] Boser, B., Guyon, M., and Vapnik, V. (1992). A training algorithm for optimal margin classier . In Conference on Computational Learning Theory (COLT), pages 144{152.
[3] Burges, C. and Sch□olkopf, B. (1997). Improving the accuracy and speed of support vector machines In Neural Information Processing Systems, volume 9.
[4] B.E. Boser, I.M. Guyon, and V.N. Vapnik.A training algorithm for optimal margin classier. In Proc.5th ACM Workshop on Computational Learning Theory, pages 144{152, Pittsburgh, PA, July 1992.
[5] C. Cortes and V. Vapnik. Support vector networks. Machine Learning, 20:1{25, 1995.
[6] B. Murtagh and M. Saunders. Large-scale linearly constrained optimization. Mathematical Programming, 14:41{72, 1978.
[7] E. Osuna, R. Freund, and F. Girosi. Support vector machines: Training and applications. A.I. Memo 1602, MIT A. I. Lab., 1997.
[8] M. Schmidt. Identifying speakers with support vectors networks. In Proceedings of Interface '96 Sydney, July 1996.
[9] B. Sch□olkopf, C. Burges, and V. Vapnik. Extracting support data for a given task. In U.M. Fayyad and R. Uthurusamy, editors, Proceedings of the First International Conference on Know ledge is discovery and Data Mining, Menlo Park, CA, 1995. AAAI Press.
[10] V. Vapnik. The Nature of Statistical Learning Theory. Springer, New York, 1995.
[11] X. Zhang. Non-linear predictive models for intra-day foreign exchange trading. Technical report, PHZ Partners, August 1993.

