ISSN 2047-3338

Performance Evaluation on Classification of Remote Sensing Images

G. Sakthi and Dr. R. Nedunchezhian

Abstract– **In this paper, the performance of the classifier for classifying the Remote Sensing (RS) images is evaluated. The variety, type and extent of land cover throughout a region can be estimated through satellite remote sensing. This meets a need common to ecological applications. Land cover classification is undertaken through RS images acquired by multispectral sensors. In this study, Walsh Hadamard transform was used for feature transformation of the RS images and information gain for feature reduction. The features selected were classified using Naïve Bayes, k Nearest Neighbor, J48 decision tree and Support Vector Machine (SVM). The classifiers were evaluated using a Subset of Salinas dataset consisting of corn_senseced_green_wees and Lettuce_romaine_5wk.**

Index Terms– **Remote Sensing (RS) Images, Salinas Dataset and Support Vector Machine (SVM)**

I. INTRODUCTION

REMOTE Sensing (RS) receives and interprets information from a distance through the use of sensors which are away, without physical contact, with the object being observed. RS science includes aerial, satellite, and spacecraft observations of the solar system's planet's surfaces and atmospheres. But the Earth is usually the most frequent target. Land use/land cover change (LUCC) is in addition to being a local environmental issue is also a force driving climatic warming [1]. RS classification is an important procedure to extract/analyze LUCC through multispectral imagery data. Urban environment monitoring and climate change detection uses accurate land use/cover (LUC) classifications and biophysical estimations from remotely sensed data.

Hyperspectral remote sensors like airborne visible infrared imaging spectrometer (AVIRIS), compact airborne spectrographic imager (CASI), multispectral infrared and visible imaging spectrometer (MIVIS), and hyperspectral mapping (HyMapTM) system, are used for agricultural applications due to the expeditious development of remote sensing technology in recent times. Such sensors ensure quality images having high spatial/spectral resolutions needed for precise agricultural operations [2, 3] as such high resolution images can also be utilized for potential applications in environmental impact assessments [4, 5]. Due to both high spectral resolution and narrow band range of about 10 nm or lesser, hyperspectral remote sensing images result in each pixel having a complete spectrum within a scene. These when joined to high signal-to-noise ratio enable differentiation of differing vegetation types based on spectra of small ground surface patches [5].

Traditional ground scouting was uneconomic for efficient detection/monitoring of large cropping areas whereas RS provides a powerful tool to collect crop canopy data for analysis of geo-temporal and geo-spatial properties of crop canopies including late blight symptoms [6]. This paper evaluates classifier performance in classification of RS images. Evaluation was through use of Salinas dataset subset including corn_senseced_green_wees and Lettuce_romaine_5wk.

A. Related Works

Demir et al [7] investigated various batch mode active learning techniques for RS images classification with SVMs through generalizing to multiclass problems techniques meant for binary classifiers. The techniques investigated exploited varied query functions, based on a two criteria evaluation: uncertainty and diversity. A novel query function based on kernel clustering technique to assess samples diversity and a new strategy to select the most informative representative sample from individual clusters was proposed. Both investigated and proposed techniques were theoretically and experimentally compared to state-of-the-art methods use for RS applications. The proposed method leads to improved accuracy.

Munoz-Mari, et al [8] addressed supervised classification issues in RS images when incomplete training sets were available. The problem analysis is based on 2 principles, 1) description/recognition of specific land-cover class using single-class classifiers and 2) multiclass problems solution with single-class classification techniques. The proposed framework both analyzes and introduces various one-class classifiers and introduced to the remote sensing community the support vector domain description method (SVDD), a

G. Sakthi, is Research Scholar, Manonmaniam Sundaranar University, Tirunelveli, TamilNadu, India, (Email: sakthihit@gmail.com)

Dr. R. Neduchezhian, is Professor & Head, Ramakrishna Engineering College, Coimbatore, TamilNadu, India, (Email: rajuchezhian@gmail.com)

kernel-based method showing intrinsic regularization ability and robustness versus high-dimensional samples low numbers. SVDD is compared to other single-class methods both with problems that focused on recognizing a single specific land-cover class and also in multiclass problems. An easily scalable multiclass architecture was defined for the latter capable of dealing with incomplete training data. Experimental results from various data methods including synthetic, hyperspectral, and multisensor images, demonstrate the SVDD's effectiveness.

A simple, fast, and reliable active-learning technique to solve RS image classification issues through SVM classifiers was proposed by Patra et al. [9]. A characteristic of the proposed technique includes its robustness in handling biased (poor) initial training sets. The procedure considers the classifier's 1-D output space to identify uncertain samples whose labelling and inclusion in the training set involved high probability if classification results needed to improve. Use of a simple histogram-thresholding algorithm led to location of low-density region in 1-D SVM output space. The proposed method was compared with active-learning techniques proposed in RS literature to assess its effectiveness using multispectral/hyperspectral data. Experiment results are confirmed that it provides the best trade-off in robustness to biased initial training samples, computational complexity, classification accuracy, and new labelled samples required to converge.

II. METHODOLOGY

A. Dataset

A part of a 2001 AVIRIS data set taken over an agricultural test site in Salinas Valley, California is the hyperspectral scene chosen for the experiment. It includes 512 lines by 217 samples, with 154 spectral bands after removal of water absorption/noisy bands. Data includes vegetables, bare soils and vineyards with subcategories as Fig.1 shows, revealing the entire scene and a dataset sub-scene. Salinas A, the subscene outlined by a rectangle in Fig. 1 has 83x86 pixels dominated by directional classes. Ground figures taken during data acquisition are also seen. A most interesting Salinas dataset feature is its representation of a hyperspectral analysis scenario dominated by directional classes having the same spatial and spectral properties. An example is the Romaine lettuce at different weeks since planting with growth covering soil increasingly resulting in slightly distinct spectral signatures. A Salinas dataset's subset including corn_senseced_ green_wees and Lettuce_romaine_5wk was used for evaluation in this study.

B. Feature transformation using Walsh Hadamard Transform

Walsh Hadamard Transform (WHT) is an image processing tool with Unser [10] using Hadamard matrices along with local transforms like DCT and KLT to ensure texture measurements. To evaluate filter effectiveness in texture analysis varied, small size filters and a filter sliding scheme were applied to the spatial domain.

$$W(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} I(x,y)\left[(-1)^{\psi(u,v,x,y)}\right] \tag{1}$$

where I is image, N image size and ψ determines the transform's parametric kernel function [11] selected from a diverse possibilities set.

WHT's many computational advantages include it being a real transform only requiring addition and subtraction operations. Input signal requires only integer operations when it is a set of integer-valued data. Walsh transforms have a fast algorithm through substituting Fast Fourier transform's exponential kernel with the $-1^{\psi(.)}$ kernel of Walsh. The transform matrix also known as Hadamard matrix leads to many lowered memory requirement when saved in binary format. Compared to other transforms WHT is simpler to implement in hardware.

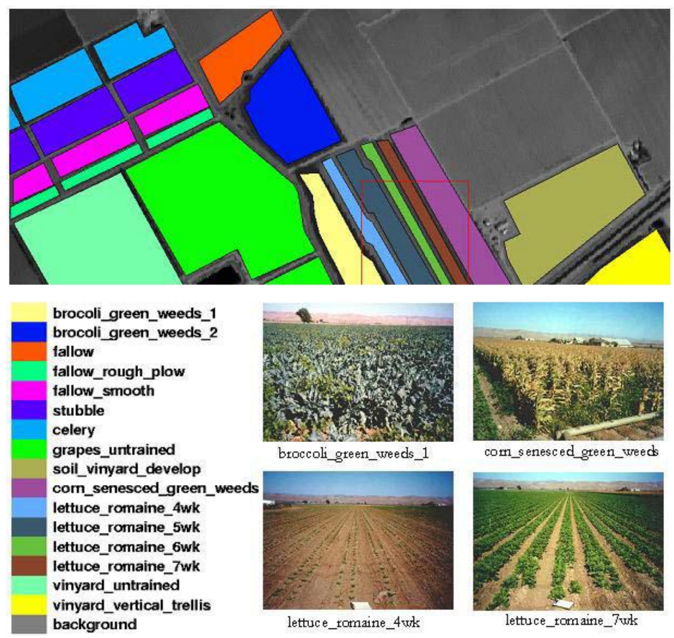

Figure 1: Data set collected over Salinas Valley in California

C. Feature selection using Information gain

Information gain selects feature vectors required for classification [12]. This study, chose information gain values based, top 100 features as classifier inputs. On the computed coefficient from WHT, the information gain can be computed based on the class attribute. The information gain (I(Y;X) for attribute X whose class attribute Y is given by the conditional entropy of Y given X (H (Y|X)) is

$$I(Y;X) = H(Y) - H(Y|X) \tag{2}$$

The conditional entropy of Y given X is:

$$H(Y \mid X) = -\sum_{j=1}^{j=l} P(X = x_j) H(Y \mid X = x_j) \tag{3}$$

D. Classifiers

Naive Bayes: Classification is fundamental issue to both machine learning and data mining. The aim of a learning algorithm in classification is to construct a classifier given a class labelled set of training examples. An example E is represented by an attribute values (x1, x2, ,⋯, xn) tuple where xi is attribute Xi value. Let C represent classification variable and let c be C's value. From the probability perspective, According to Naïve Bayes, from a probability perspective, the probability of an example E = (x1, x2,⋯, xn) being class c is:

$$p(c|E) = \frac{p(E|c)\,p(c)}{p(E)} \quad (4)$$

If it is assumed that all attributes are independent, given the value of a class variable; that is,

$$p(E|c) = p(x_1,..,x_n|c) = \prod_{i=1}^{n} p(x_i|c) \quad (5)$$

k Nearest Neighbor: k- Nearest neighbor classifier is based on the premise that vector space model is the same for similar documents. Indexed training documents are associated with corresponding labels. A test document after submission is treated like a query and retrieves from documents from a training set which are similar to test documents. The test document's class label is assigned on the basis of distribution of its k nearest neighbors. This is further refined through the addition of weights. Thus higher accuracy is obtained by tuning k.

J48: A decision tree is a predictive machine-learning model decides a new sample's target value (dependent variable) based on available data's differing attribute values. The tree's internal nodes denote different attributes with inter-nodal branches revealing attributes possible values in observed samples. Terminal nodes tell us of the dependent variable's [13] final value. A predicted attribute is known as dependent variable, as its value is based on/decided by other attributes values. The latter which aid prediction of dependent variable value are known as the independent variables in datasets.

Support vector machine (SVM): A linear machine, Support vector machine (SVM) constructs a hyperplane as a decision surface, based on a structural risk minimization method; the error rate being bound the total of training-error rate and a term dependent on Vapnik-Chervonenkis (VC) dimensions [14]. SVM ensures proper generalization performance on pattern classification. SVM algorithm principle is based on inner-product kernel between a "support vector" x_i and vector x drawn from input vector.

SVM uses larger space mapping to compute cross products easily with original space variables ensuring an easier computational load. Cross products in the larger space are defined with a kernel function $K(x,y)$ selected to suit the problem domain. Cross products having a space vector if constant can define hyperplanes [15]. Vectors defining hyperplanes are combinations with parameters α_i of feature vectors occurring in the data base. After selection of the hyperplane, points x in feature plane are defined by:

$$\alpha_i K(x_i,x) = \textit{constant } i \quad (6)$$

If $K(x,y)$ is small when y goes further from x, the closeness degree is provided by the sum measures of test point x's closeness to corresponding data base point x_i. This paper uses Radial Basis Function (RBF) kernel.

$$\text{RBF} = \text{Exp}\,(-\gamma|x_i-x_j|^2) \quad (7)$$

This method measures each test point's closeness to the data points from discriminated data sets. As mapped set points can be convoluted, it results in complex discrimination between sets that are not convex in original space.

SVM's perform well as a learning algorithm on various classification problems. They also ensure rapid classification from trained models and can accommodate input vectors of very high-dimensions.

The error function used in our implementation is given by equation:

$$\tfrac{1}{2} w^T w + C\sum_{i-1}^{N} \xi_i + C\sum_{i-1}^{N} \xi^{*} i \quad (8)$$

which can be minimized to

$$w^T \phi(x_i) + b - y_i \le \varepsilon + \xi_i^{*}$$
$$y_i - w^T \phi(x_i) - b_i \le \varepsilon + \xi_i$$
$$\xi_i, \xi_i^{*} \ge 0, i = 1,..., N$$

where C is the capacity constant, w is the vector of coefficients, b a constant and ξ_i are parameters for handling non separable data (inputs). The index i label the N training cases.

III. RESULTS

The performance of the classifiers was evaluated using a subset of Salinas dataset consisting of corn_senseced_green_wees and Lettuce_romaine_5wk. The experiments are run 10 fold cross validation. The WHT was applied to the images for feature transformation. Feature selection is done through Information gain and selecting top 100 features as input for the classifiers.

The classification accuracy of images is evaluated using Naïve Bayes, K Nearest Neighbor, J48 decision tree and SVM with RBF kernel. The results are tabulated in Table 1. Figure 2 shows the classification accuracy and root mean squared error (RMSE) and Figure 3 shows the precision and recall.

Table 1: Summary of Results

Classifier	Classification Accuracy	RMSE	Precision	Recall	f Measure
Naïve Bayes	49.62%	0.7092	0.526	0.496	0.449
J48 Algorithm	82.81%	0.3966	0.828	0.828	0.828
K Nearest Neighbor	82.81%	0.4145	0.828	0.828	0.828
SVM with RBF kernel	89.54%	0.3234	0.895	0.895	0.895

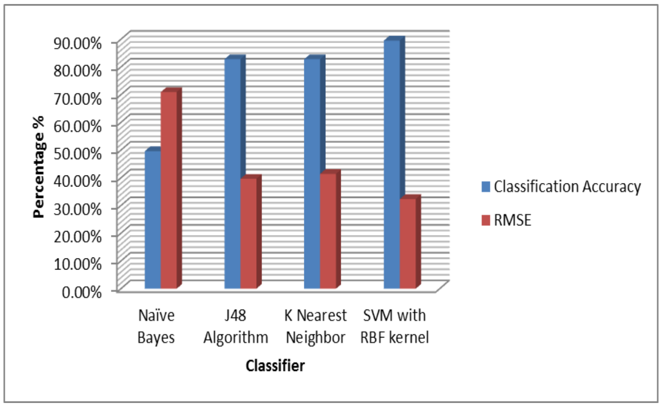

Figure 2: Classification Accuracy and RMSE

It is observed from the above tables and figures that the SVM with RBF kernel achieves the highest classification accuracy, precision and recall for the hyperspectral images. It also has the lowest RMSE.

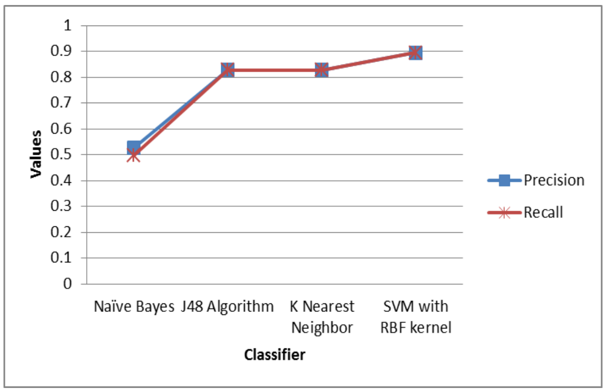

Figure 3: Precision and Recall

IV. CONCLUSION

In this study, Walsh Hadamard transform was used for feature transformation of the RS images and information gain for feature reduction. The features selected were classified using Naïve Bayes, k Nearest Neighbor, J48 decision tree and Support Vector Machine (SVM). The classifiers were evaluated using a Subset of Salinas dataset consisting of corn_senseced_green_wees and Lettuce_romaine_5wk. Experimental results demonstrate the superiority of the SVM for classification of the RS images. The SVM with RBF kernel achieves the highest classification accuracy of 89.54%.

REFERENCES

[1] E. Kalnay and M. Cai, "Impact of urbanization and land-use change on climate," Nature, vol. 423(6939), 29 May 2003, pp. 528-531, doi:10.1038/nature01675.

[2] Leblanc, S., Chen, M.J., Miller, R.J., Freemantle, J., 1999. Compact airborne spectrographic imager (CASI) used for mapping leave area index (LAI) of cropland. In: Proceedings of the Fourth International Airborne Remote Sensing Conference and Exibition/21st Canadian Symposium on Remote Sensing, vol. 1, Ottawa, Ont., Canada, pp. 297–303

[3] Nolin, A.W., Dozier, J., 2000. Hyperspectral method for remotely sensing the grain size of snow. Remote Sens. Environ. 74 (2), 207–216

[4] Barducci, A., Pippi, I., 1997. Environmental monitoring of the Venice alggon using MIVIS data. Int. Geosci. Remote Sens. Symp. (IGARSS) 2, 888–890.

[5] Bianchi, R., Cavalli, R.M., Fiumi, L., Marino, C.M., Pignatti, S., 1999. Airborne MIVIS hyperspectral imaging spectrometer over natural and anthropic areas. In: Proceedings of the Fourth International Airborne Remote Sensing Conference and Exibition/21st Canadian Symposium on Remote Sensing, vol. 1, Ottawa, Ont., Canada, pp. 337–344.

[6] Rush, C., 2002. Prediction, detection and quantification of plant diseases with remote sensing technologies. Precision Agriculture Initiative for Texas High Plains, 2002 Annual Comprehensive Report.

[7] Demir, B., Persello, C., &Bruzzone, L. (2011). Batch-mode active-learning methods for the interactive classification of remote sensing images. Geoscience and Remote Sensing, IEEE Transactions on, 49(3), 1014-1031.

[8] Munoz-Mari, J., Bovolo, F., Gómez-Chova, L., Bruzzone, L., & Camp-Valls, G. (2010). Semisupervised one-class support vector machines for classification of remote sensing data. Geoscience and Remote Sensing, IEEE Transactions on, 48(8), 3188-3197.

[9] Patra, S., &Bruzzone, L. (2011). A fast cluster-assumption based active-learning technique for classification of remote sensing images. Geoscience and Remote Sensing, IEEE Transactions on, 49(5), 1617-1626.

[10] M. Unser. Local linear transform for texture measurements. Signal Processing, 11:61–79, 1986.

[11] K. Beauchamp. Applications of Walsh and related functions. Academic Press, 1984.

[12] Guyon, I., &Elisseeff, A. (2003). An introduction to variable and feature selection. The Journal of Machine Learning Research, 3, 1157-1182.

[13] Zhang, Y., Zhao, Y., A Comparison of BBN, ADTree and MLP in separating Quasars from Large Survey Catalogues, ChJAA 7, 289-296, 2007.

[14] Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. Neural processing letters, 9(3), 293-300.

[15] Vapnik, V. N. (1999). An overview of statistical learning theory. Neural Networks, IEEE Transactions on, 10(5), 988-999.

G. Sakthi is currently working as an Assistant Professor in the Department of Computer Science and Engineering, Hindusthan Institute of Technology, Coimbatore, Tamil Nadu, India. He is currently pursuing his doctoral degree at Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India in the area of Image Processing. He obtained his B.E., degree in Computer Science and Engineering from Bharathidasan University, Tiruchirapalli, and M.E., degree in Computer Science and Engineering from Manonmaniam Sundaranar University, Tirnelveli. His research interests are Satellite Image Processing and Data Mining. He is a Life Member of Indian Society of Remote Sensing, Dehradun.

Dr. R. Nedunchezhian is currently working as a Professor and Head, Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore. He has more than 20 years of experience in research and teaching. Currently, he is guiding many Ph.D scholars of the Anna University, Coimbatore, and the Bharathiyar University. His research interests are knowledge discovery and data mining, distributed computing, and database security. He has published many research papers in national/international conferences and journals. He has edited a book entitled “Handbook of Research on Soft Computing Applications for Database Technologies: Techniques and Issues” which was published by IGI publications, USA in April, 2010. He is a Life member of Advanced Computing and Communication Society and ISTE.