Kernel based Multi-Class Classification of Satellite Images with RVM Classifier using Wavelet Transform

S.Sindhu
PG student
Department of ECE
Velammal College of Engineering & Technology, Madurai

S.Vasuki, Ph.D
Head and Professor
Department of ECE
Velammal College of Engineering & Technology, Madurai

ABSTRACT
Multispectral satellite images are more efficient and a suitable method of obtaining information about land, because it can captures an image at specific frequency across the spectrum. This spectral image can allow extraction of further information about ground survey than the other traditional image. Classification of multispectral image consists of image processing and classification method. Here, an efficient technique is proposed for classifying the multispectral images using fuzzy incorporated hierarchical clustering with RVM classifier. In the proposed technique, first the multispectral satellite image is subjected to set of pre-processing steps, which are used to transform an image into suitable form that is easier for segmentation and classification. Subsequently, the pre-processed image is segmented using fuzzy incorporated hierarchical clustering. Then, the proper kernel function is selected for RVM clustered output. Finally the multispectral image is classified into multiple sectors based on the training data. The classification is used in the application of land degradation studies, environmental damage, resource management and other environmental application.

Keywords
Classification, RVM, Multispectral satellite image ,clustering.

1. INTRODUCTION
Multispectral satellite image produces detail information about environmental factors. It is more economical than other traditional method of ground survey and aerial photography [1]. These details are utilized in a number of applications like making of mapping that was used in military and civil use, estimating of environmental degradation, soil test and crop outcome increment.

Multispectral satellite image classification is done by image processing and classification methods. Multispectral image classification increases efficiency from an efficient method capable of arranging the spectral information in Multispectral data. Usually, image classification refers to creating a group of similar pixels that are belongs to same class. Many classification techniques are introduced in remotely sensing data like KNN, MLP,SVM etc. Comparing to RVM classifier SVM classifier requires large training data [16]. The RVM classifier has a lower computational complexity when compared to other classifiers [15]. In real time communication fields, the RVM classifier is better for constructing smaller equalizer without reducing the performance [14].

The main focus of research is to classify the images into land use and land cover region. Land covers characterize the features of land surfaces like vegetation area, soil, mud and crop. Land use is a statement of how people utilize the land like building, roof, etc. This land cover mapping is an important parameter for environmental and land use planning at local and national level. Here, a new technique is introduced to increase the classification efficiency with RVM. In the proposed techniques consists of four phases of pre-processing by wavelet transform, segmentation by fuzzy incorporated hierarchical clustering, training data selection for RVM, classification using trained RVM. The multispectral satellite image is applied to Gaussian filter. The image quality is enhanced by reducing noise component in an image. The better image can be obtained from Gaussian filtering. Then, it is applied to wavelet transform. It can provide detail information about multispectral data which are used to increase the classification accuracy. Then, fuzzy incorporated hierarchical clustering is used for segmentation of the image into clusters. The training data are subjected to RVM and final classification of the Multispectral data is obtained.

2. PROPOSED WORK
In the proposed technique, the multispectral image is classified by using RVM classifier with clustering. Initially pre-processing is done on the input multispectral image where the image is subjected to set of pre-processing steps such as Gaussian filtering and wavelet transform. So that image is transformed suitably for segmentation. The pre-processed image is segmented using fuzzy incorporated hierarchical clustering. Then the proper kernel function is selected for RVM classifier. Finally classification of Multispectral image is done which is based on the training data given to the classifier. The proposed technique is illustrated by Figure (1).

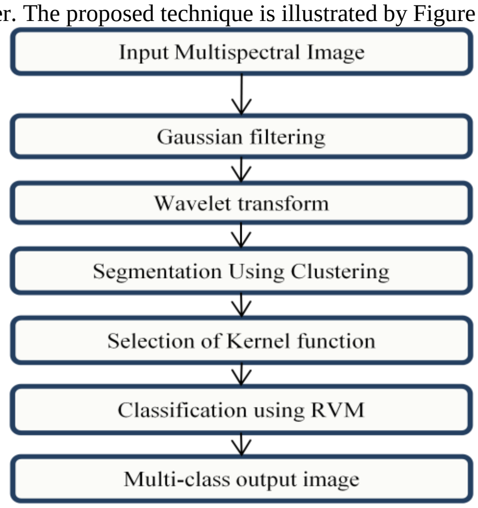

Fig.1. Block Diagram of the Proposed Work

3. GAUSSIAN FILTERING

A Gaussian function is the response of Gaussian filtering. Gaussian filtering is used to avoid the overshoot of step function while reducing the rise and fall time [4]. Because of this characteristics Gaussian filtering has a minimum group delay. Gaussian filter performing convolution between input and Gaussian factors. The result is called Weierstrass transform.

The 1D Gaussian filter is given by,

$$g((x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} \qquad (3.1)$$

The 1D Gaussian filter's impulse response that is given by,

$$g((x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-\sigma^2 u^2}{2}} \qquad (3.2)$$

Here in the pre-processing the unwanted noise components are reduced from the input image while passing through Gaussian filter. It is mainly used to enhance the image quality than the original image.

4. WAVELET TRANSFORM

Wavelet transform is based on wavelet theory. It was initially developed for signal transformation. In discrete wavelet transform (DWT) the signal is applied to series of low pass and high pass filter to analyze the low frequency and high frequency component [5]. The amount of detail information is measured from filtering operation and scale is changed by up sampling and down sampling. The DWT analyze the signal at different frequency and different resolution by decomposing signal into approximation and detail information. The DWT is developed by Mallet (1989) for image decomposition. The DWT decompose an image into one approximation and three detail information sub bands. The image is subjected into bank of low pass and high pass filter. It filters an image in all direction. The filtered images are down sampled at every pixel, it producing four sub band images of original image. Figure2.(a) illustrates the process of first level decomposition. For a one level decomposition wavelet transform, the approximation sub image maintain the maximum portion of the original image, while the detail sub image represents the differences between approximation sub image and original image in different direction. In second level decomposition the filtering process is applied to the approximation sub image, resulting is four additional sub bands in approximation. These details are given in Figure (2).

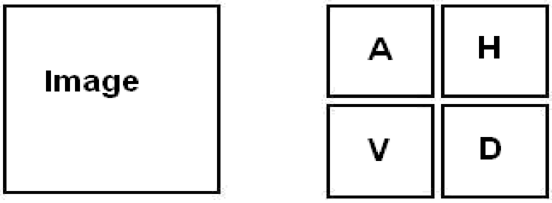

Fig.2.(a). Input image and sub images position in output image

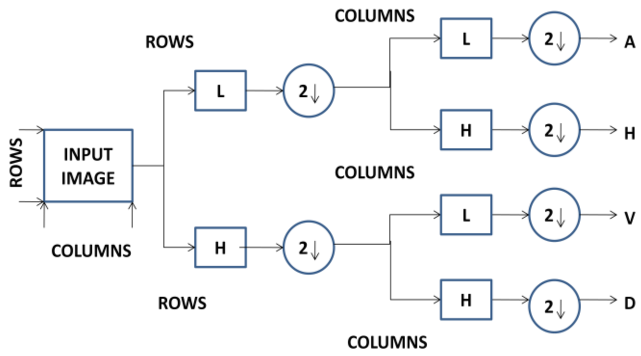

Fig.2. (b). Convolution Of Image.

The wavelet decomposition is otherwise called as filter convolution. In the first level decomposition, the image can be obtained by convolving the image with appropriate filter. The wavelet transform provides set of pixel information in each level of decomposition. From the approximation and detail information, the fine information is obtained for classification. This transform leads to reduce the misclassification of process.

The multispectral satellite images is used as input image which is to be subjected to the Gaussian filtering for noise reduction and applied to the wavelet transform. The three (RED, GREEN and BLUE) planes are extracted from the Gaussian filtered output. The DWT is applied into the separate planes. So the four sub bands are obtained for separate planes. The approximation sub band is enough for the future process. Hence, the approximation sub bands of three planes are concatenated to get the approximated output image. This approximated output image is applied to the segmentation using Fuzzy incorporated hierarchical clustering. The kernel function is chosen for RVM from the clustering output. The RVM classifier is classified the image into multiple region based on the training data.

5. SEGMENTATION USING PROPOSED CLUSTERING

After performing the pre-processing steps, an image which is suitable for segmentation is obtained. The processed image consists of thousand of pixels and to classify the image based on of pixels is time consuming. Hence, the input image is clustered only based on the centroid pixel value. This is due to the fact, that all pixels are differed only a small amount from centroid pixel value. Hence, the centroid pixel value represents all pixels in the cluster. This centroid value clustering will reduce the input to RVM classifier and reduce the complexity of design. In the proposed clustering, the hierarchical clustering is combined with the Fuzzy C means. This will increase the accuracy of Fuzzy C means and reduce the time complexity of Hierarchical clustering.

Hierarchical clustering has two types. 1. Agglomerative 2. Divisive. However, it has disadvantage of high complexity and minor variation in dataset also greatly changes the hierarchical dendrogram structure [6]. The traditional FCM has a disadvantages of it does not yield the accurate result. That is, every time Fuzzy clustered the same data but producing different results. These two drawbacks are overcome by the proposed Fuzzy incorporated hierarchical clustering [7]. The clustering process is explained below:

1. Let the image have M number of pixels. Initially all the pixels are act as a different cluster and hence it forms M different cluster. Let the M cluster of an image is represented as C_i, where 0< i<M. Calculate pixel difference matrix λ.

$$\lambda = \begin{bmatrix} \partial_{11} & \partial_{12} & \partial_{13} & \partial_{14} & .. & \partial_{1M} \\ \partial_{21} & \partial_{22} & \partial_{23} & \partial_{24} & .. & \partial_{2M} \\ \partial_{31} & \partial_{32} & \partial_{33} & \partial_{34} & .. & \partial_{3M} \\ \partial_{41} & \partial_{42} & \partial_{43} & \partial_{44} & .. & \partial_{4M} \\ .. & .. & .. & .. & .. & .. \\ \partial_{M1} & \partial_{M2} & \partial_{M3} & \partial_{M4} & .. & \partial_{MM} \end{bmatrix}$$

Where ∂_{ij} = pixel difference value between i th and j th cluster.

2. From the pixel difference matrix find out two clusters which have a minimum differences and merge together to form a new cluster $C_{ij.}$

3. The centroid value O_{ij} is calculated for new cluster.

$$\mathbf{O_{IJ}} = \frac{\mathbf{C_I + C_J}}{\mathbf{2}} \qquad (5.1)$$

4. Subsequently, find out the original centroid pixel value by using incorporated Fuzzy C means algorithm. Find membership value and then calculate the original centroid value using membership function.

$$\boldsymbol{\mu_{ij}} = \frac{\mathbf{1}}{\sum_{k=1}^{c}\left(\frac{\|P_i - O_{ij}\|}{\|P_i - o_k\|}\right)^{\frac{2}{m-1}}} \qquad \mathbf{(5.2)}$$

Where,
O_{ij} - approximated centroid pixel value
O_k - centroid pixels of other clusters excluding the newly formed cluster.
m – Positive integer

Modified centroid pixel value for newly formed cluster is given by,

$$\mathbf{C_{ij}\ is\ given\ by\ O_j} = \frac{\sum_{i=1}^{N} \mu_{ij}^{m} x_i}{\sum_{i=1}^{N} \mu_{ij}^{m}} \qquad (5.3)$$

5. The above result is the formation of two individual clusters which have greatest similarities. Hence it decreases the total number of clusters by one, after every iteration.

6. The pixel difference matrix dimension is also reduced by replacing the value of C_i and C_j value by Cij. It will lead to decrease the matrix dimension of M X M to (M - K) X (M – K) after K loops.

7. Go back to step 2 until desired number of cluster is obtained.

6. KERNEL FUNCTION

In digital image processing, kernel or convolution matrix is a small square matrix used for sharpening and smoothing operation [13]. Kernel methods map the input data to higher dimension. In this higher dimension the data can be easily separated and classified. Generally so many kernel functions are available in regression and classification fields. Some of the kernel functions are listed below:

- Linear Kernel
- Polynomial Kernel
- Gaussian Kernel
- Exponential Kernel
- Laplacian Kernel
- ANOVA Kernel
- Hyperbolic Tangent (Sigmoid) Kernel
- Rational Quadratic Kernel
- Multiquadric Kernel
- Inverse Multiquadric Kernel
- Circular Kernel
- Spherical Kernel
- Wave Kernel

Depending upon our application requirement any one kernel function is chosen for classification.

7. RVM CLASSIFIER

Now a day, so many techniques are developed for satellite image classification. The earlier day maximum likelihood classifier is used for classification. In recent times, artificial intelligence techniques are developed for remotely-sensed image classification application. Relevance vector machine (RVM) is a kernel based classification which is alternative for SVM classifier [23]. RVM is a identical form of SVM classifier, but RVM is a probabilistic classifier. The advantages of RVM classifier over SVM are listed below[24]:

(1) RVM classifier requires smaller amount of relevance vector than the SVM.
(2) The testing time is less than SVM classifier.
(3) The design complexity and cost is low for RVM when compared to SVM.

Generally RVM classifier is designed for binary classification. The aim of RVM classifier is to find the posterior probability of membership for any one of two classes. Of 0 and 1 [25].

For a two class classification, consider the classes are X = (x_1,.......,x_n) having a classes of C = (c_1,.....c_n), where C_iЄ(-1,1).

Based on Bernoulli distribution, the likelihood is given by,

$$\mathbf{p}\left(\frac{\mathbf{c}}{\mathbf{w}}\right) = \prod_{i=1}^{n} \boldsymbol{\sigma}\left\{\left(\mathbf{Y(X_i)}\right)\right\}^{\mathbf{C_i}}\left[\mathbf{1} - \boldsymbol{\sigma}\left\{\left(\mathbf{Y(X_i)}\right)\right\}\right]^{\mathbf{1-c_i}} \qquad (7.1)$$

Where,

The logistic sigmoid function σ(y) is given by,

$$\boldsymbol{\sigma}\left(\mathbf{Y(X)}\right) = \frac{\mathbf{1}}{\mathbf{1 + exp(-Y(X))}} \qquad (7.2)$$

An iteration method is used to obtain P(C/W). Let α_i^* = Maximum posteriori estimate of the hyper parameter α_i W(m) can be obtained by maximizing the following objective

$$f(w_1, w_2, ., w_n) = \sum_{i=1}^{n} \log\, p\left(\frac{c_i}{w_i}\right) + \sum_{i=1}^{n} \log\, p\left(\frac{w_i}{\alpha_i^*}\right) \qquad (7.3)$$

W(m) = maximum posterior estimate of the weights

In this expression, the first summation term represents the likelihood of the class labels, and the second term represents the prior on the parameters $W_{i.}$
The gradient of f is calculated for results. the posterior is approximated by Gaussian approximation [26].

$$\mathbf{covariance}, \boldsymbol{\varepsilon} = -(\mathbf{H}|\mathbf{w_m})^{-1} \qquad (7.4)$$

$$\mathbf{mean}, \boldsymbol{\mu} = \sum \boldsymbol{\varphi}^{\mathbf{T}}\mathbf{Bc} \qquad (7.5)$$

Where,
H= Hessian of **f.**
B= Diagonal matrix.

7.1 Multiclass Approaches With RVM

RVM is mainly designed for binary classification. But the applications of binary classification are limited. The multiclass classification is introduced in order to increase the accuracy of land feature than the binary classification [9]. There are so many techniques were introduced to performing multiclass classification with RVM. The important four techniques are listed below:

- one vs. One
- one vs. rest
- Directed Acyclic Graph (DAG)
- Error Corrected Output Coding (ECOC)

The first two techniques are widely used multiclass approaches. Multiclass classification classifying the multispectral image into multiple regions rather than just land use and land cover [17] [18].
One-against-all classification, in which there is one binary RVM for each class to separate members of that class from members of other classes.
One against one, in which there is one binary RVM for each pair of classes to separate members of one class from members of the other.
In this is two methods one against one method is chosen for multiclass classification. It increases the classification accuracy than the one against one method [19].

7.2 Training Data Selection

In this technique, the training data selection from the clustered output image is discussed. This proposed technique is used to classify the multispectral image into multiple regions. This classification accuracy increased from the color features extraction in satellite images [21]. Each and every element in the earth is differentiated by its own color. In the multispectral image certain color stands for land use and certain color stands for land cover. These colors are identified from the satellite image and given to the input of RVM classifier.

7.3 Classification Using RVM

In the proposed method, the multispectral image is segmented by Fuzzy incorporated Hierarchical clustering. All clusters almost have a similar number of pixels value and the clusters are differ from the centroid value by small amount only. Hence, the centroid value represents all the pixels value in the cluster. This centroid valve is given to the RVM classifier hence it reduce the classifier complexity and also time incurred [11]. Consider the i^{th} cluster, which having n number of pixels, where each pixel having a value of P_k. Then the centroid value is calculated by

$$\mathbf{O_t} = \frac{\sum_{\mathbf{k=1}}^{\mathbf{n}} \mathbf{P_k}}{\mathbf{n}} \qquad (7.6)$$

If our image consists of N number of cluster, then find the centroid set O.
O = (O_1, O_2… ON). It will be given to the input of the RVM classifier. The centroid set is classified based on the training data given to the RVM classifier [22].

8. RESULTS AND DISCUSSION

The classification and segmentation of multispectral images is implemented by MATLAB. In this approach, the proposed method of classification is discussed.

8.1 Experimental Results

In this section, the output of the proposed technique is discussed. Here, multispectral satellite images as an input image which is to be subjected to the Gaussian filtering for noise reduction and applied to the wavelet transform. The three (RED, GREEN and BLUE) planes are extracted from the Gaussian filtered output. The DWT is applied into the separate planes. So the four sub bands for separate planes are obtained. An approximation sub band is enough for the future process. Hence, the approximation sub bands of three planes are concatenated to get the approximated output image. This approximated output image is applied to the segmentation using Fuzzy incorporated hierarchical clustering. The training data for RVM are taken from the clustering output. The RVM classifier is classified the image into multiple regions based on the training data. Figure (4) shows the input image. Along with Figure (5) shows the output of Clustering. The final classified output of RVM is given in figure (6).Figure (7) represents the final classified images of RVM classifier.

Fig.4. Input image.

Fig.5. Clustering Output

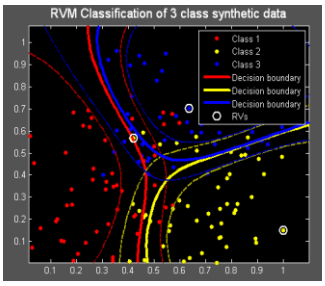

Fig.6. RVM Classifier output

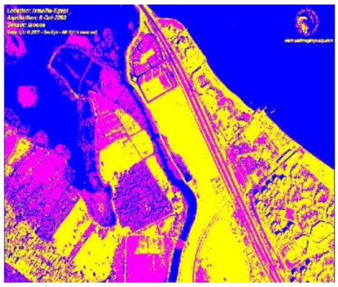

Fig.7. RVM Classifier output image

8.2 Performance Evaluation

In this section, the performance evaluation of proposed technique is discussed. Here, the performances of various satellite multispectral images are evaluated. The RVM output of different satellite images are shown in fig.7, 8, 9 and10. From the performance evaluation table, RVM classifier having higher accuracy and lower error rate for all type of input images. Sensitivity and specificity are statistical measures of the performance of a binary classifier. Sensitivity measures the proportion of actual land use pixels which are correctly identified. Specificity measures the proportion of land cover pixels which are correctly identified. RVM classifier having higher sensitivity and specificity value for all type of input images.

Table 1. Performance Evaluation table of RVM classifier

SATELLITE	ACCURACY (%)	ERROR RATE	SENSITIVITY	SPECIFICITY
IKONOS	**89.19**	**0.1081**	**0.9181**	**0.8649**
QUICK BIRD	**91.89**	**0.0811**	**0.9459**	**0.8919**
GEO EYE-1	**87.84**	**0.1216**	**0.8919**	**0.8649**
GEO EYE-2	**94.59**	**0.0547**	**0.9459**	**0.9459**

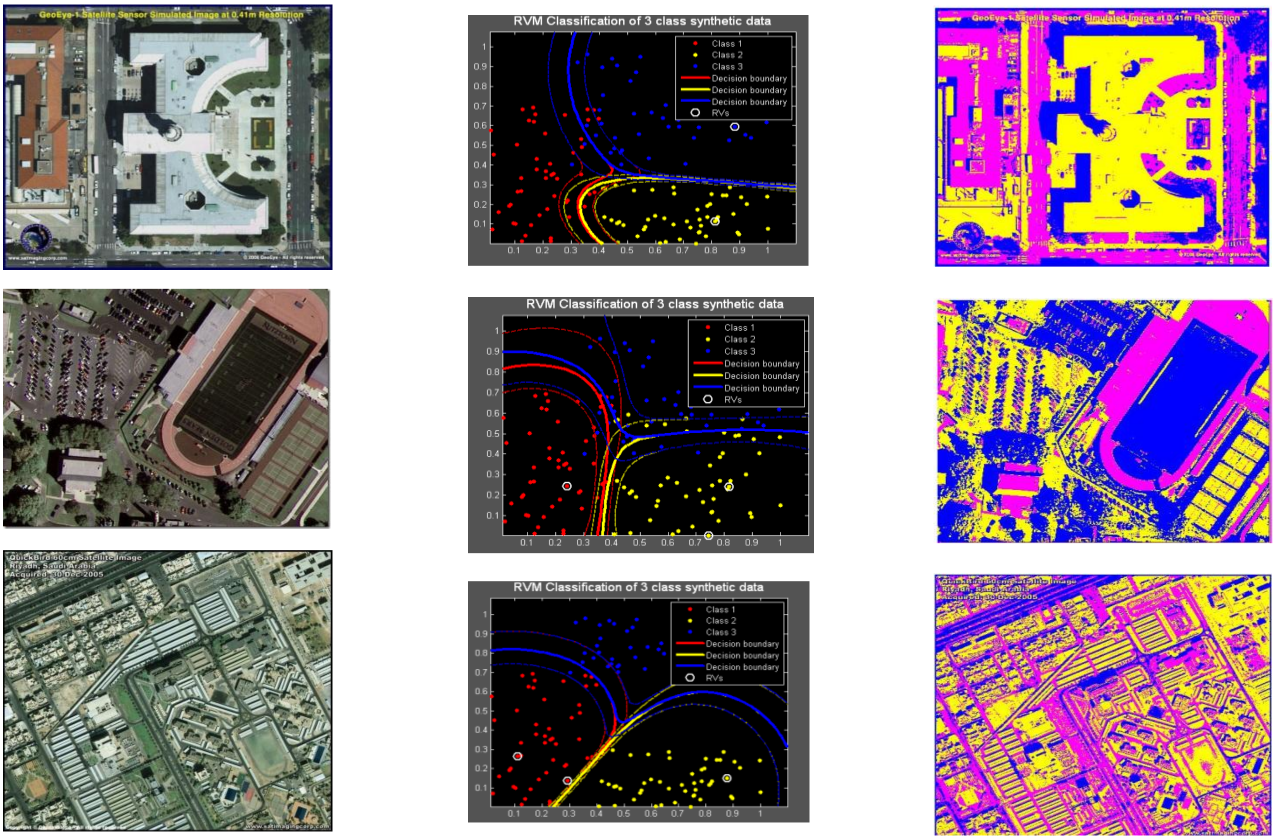

Fig.8. (a) (b) **(c)**

Figure (8).(a)shows the input images of **QUICK BIRD, GEO EYE-1 and GEO EYE-2** respectively**.** (b) Represents the RVM classifier output of **QUICK BIRD, GEO EYE-1 and GEO EYE-2** respectively. (c) Represents the final three class classified output images of RVM classifier for **QUICK BIRD, GEO EYE-1 and GEO EYE-2** respectively**.**

9. CONCLUSION

An efficient image classification technique is proposed with the help of RVM classifier. Here in our proposed technique is made of four phases namely pre-processing, image segmentation, training data selection and final classification using RVM classifier. First the multispectral satellite image is subjected to set of pre-processing steps, which are used to transform an image into suitable form that is easier for segmentation and classification. Subsequently, the pre-processed image is segmented using fuzzy incorporated hierarchical clustering. This result in the image is segmented into number of clusters. Then, the training data for RVM is chosen from clustered output. The chosen data is given to the input of trained RVM. Finally the multispectral image is classified into multiple regions based on the training data. The experimental results demonstrated the effectiveness of the proposed classification techniques. The analyse ensures that the classification has good accuracy in all type of multispectral satellite images.

10. REFERENCES

[1] K Perumal and R Bhaskaran , "SVM-Based Effective Land Use Classification System For Multispectral Remote Sensing Images", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 2, pp.95-107, 2009.

[2] M. Fernandez-Delgado, P. Carrion, E.Cernadas, J.F. Galvez, Pilar Sa-Otero (2004)-"Improved Classification of Pollen texture Images using SVM and MLP"

[3] Shaomei Yang and Qian Zhu (2008)-"Research on Comparison and Application of SVM and FNN Algorithm".

[4] Dong-Hyuk Shin, Rae-Hong Park, Senior Member, IEEE, Seungjoon Yang, Member, IEEE, and Jae-Han Jung, Member, IEEE (2005)-"Block-Based Noise Estimation Using Adaptive Gaussian".

[5] Chen, D. Chen, and D Blostein (2007)-"Wavelet-Based Classification of Remotely Sensed Images".

[6] Wei Yu, Jason Fritts and Fangting Sun (2000),"A Hierarchical Image Segmentation Algorithm".

[7] Geva,A.B(1999)-"Hierarchical unsupervised fuzzy clustering ".

[8] D. Lu, Q. Weng, "A survey of image classification methods and techniques for improving classification performance", International Journal of Remote Sensing, Vol. 28, No. 5, pp. 823-870, January 2007.

[9] Zhihua Zhang a, _, GangWangb, Dit-YanYeung b, GuangDai b, FrederickLochovsky, " A regularization framework for multiclass classification: A deterministic annealing approach ", Elsevire,pattern Recognization(2010).

[10] B.Sowmya and B Sheelarani , "Land cover classification using reformed fuzzy C-means", Sadhana, Vol. 36, No. 2, pp. 153–165, 2011.

[11] B. Yogameena, S. Veera Lakshmi, M. Archana and S. Raju Abhaikumar, "Human Behavior classification Using Multi-Class Relevance Vector Machine", Journal of Computer Science 6 (9): 1021-1026, 2010,ISSN 1549-3636© 2010 Science Publications.

[12] Huang B, Xie C, Tay R, Wu B, 2009, "Land-use-change modeling using unbalanced support-vector machines" , Environment and Planning B: Planning and Design , Vol.36, No.3, pp.398-416,2009.

[13] Mathieu Fauvel ¤¦, Jocelyn Chanussot ¤ and Jon Atli Benediktsson "evaluation of kernels for multiclass classification of hyperspectral remote sensing data, ¤Laboratoire des Images et des Signaux.

[14] S. Chen, S. R. Gunn, and C. J. Harris,' The Relevance Vector Machine Technique for Channel Equalization Application', IEEE transactions on neural networks, vol. 12, no. 6, november 2001.

[15] Fereidoun A. Mianji, "Robust Hyperspectral Classification Using Relevance Vector Machine", IEEE transactions on geoscience and remote sensing, vol. 49, no. 6, june 2011.

[16] Begüm Demir,," Hyperspectral Image Classification Using Relevance Vector Machines", IEEE geoscience and remote sensing letters, vol. 4, no. 4, october 2007.

[17] S. Chen, S. R. Gunn, and C. J. Harris,' The Relevance Vector Machine Technique for Channel Equalization Application', IEEE transactions on neural networks, vol. 12, no. 6, november 2001.

[18] Pak-Kin Wong, Qingsong Xu,' Rate-Dependent Hysteresis Modeling and Control of a Piezostage Using Online Support Vector Machine and Relevance Vector Machine', IEEE transactions on industrial electronics, vol. 59, no. 4, april 2012.

[19] Behnood Gholami,' Relevance Vector Machine Learning for Neonate Pain Intensity Assessment Using Digital Imaging, IEEE transactions on biomedical engineering, vol. 57, no. 6, june 2010.

[20] Mahesh Pal and Giles M. Foody,' Evaluation of SVM, RVM and SMLR for Accurate Image Classification With Limited Ground Data', IEEE journal of selected topics in applied earth observations and remote sensing, vol. 5, no. 5, october 2012.

[21] Andreas Ch. Braun, Uwe Weidner, and Stefan Hinz,' Classification in High-Dimensional Feature Spaces—Assessment Using SVM, IVM and RVM With Focus on Simulated EnMAP Data', IEEE journal of selected topics in applied earth observations and remote sensing, vol. 5, no. 2, april 2012.

[22] Mianji, F.A.; Ye Zhang; Babakhani, A.,' Nonlinear discriminant analysis and RVM for efficient classification of small land-cover patches ',IEEE Conference Publications Communications and Signal Processing (ICCSP), 2011 .

[23] Uehara, Hisashi; Watanabe, Hideyuki; Katagiri, Shigeru; Ohsaki, Miho,' Comparison between Minimum Classification Error training and Relevance Vector Machine ',IEEE Conference Publications TENCON 2012 - 2012.

[24] Babaeean, A.; Tashk, A.B.; Bandarabadi, M.; Rastegar, S.,' Target Tracking Using Wavelet Features and RVM Classifier', IEEE Conference Publications Natural Computation, 2008. ICNC '08.

[25] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," IEEE Trans. Geosci. Remote Sens., vol. 43, no. 6, pp. 1352–1362, Jun. 2005.

[26] M. E. Tipping, "The relevance vector machine," in Advances in Neural Information ProcessingSystems , vol. 12, S. A. Solla, T. K. Leen, and K.-R. Müller, Eds. Cambridge, MA: MIT Press, 2000.