International Journal of Electrical and Computer Engineering (IJECE)
Vol. 7, No. 3, June 2017, pp. 1180~1187
ISSN: 2088-8708, DOI: 10.11591/ijece.v7i3.pp1180-1187

Comparison of Accuracy Measures for RS Image Classification using SVM and ANN Classifiers

S.V.S.Prasad[1], T. Satya Savithri[2], Iyyanki V. Murali Krishna[3]
[1]Dept. of ECE, MLRIT, Hyderabad, India
[2]Dept. of ECE, JNTU, Hyderabad, India
[3]CSIT and Director (R&D), JNTU, Hyderabad, India

Article Info

Article history:

Received Dec 9, 2016
Revised Mar 27, 2017
Accepted Apr 11, 2017

Keyword:

Classification
Clustering
Multispectral
Satellite Image
Support Vector Machine

ABSTRACT

The accurate land use land cover (LULC) classifications from satellite imagery are prominent for land use planning, climatic change detection and eco-environment monitoring. This paper investigates the accuracy and reliability of Support Vector Machine (SVM) classifier for classifying multi-spectral image of Hyderabad and its surroundings area and also compare its performance with Artificial Neural Network (ANN) classifier. In this paper, a hybrid technique which we refer to as Fuzzy Incorporated Hierarchical clustering has been proposed for clustering the multispectral satellite images into LULC sectors. The experimental results show that overall accuracies of LULC classification of the Hyderabad and its surroundings area are approximately 93.159% for SVM and 89.925% for ANN. The corresponding kappa coefficient values are 0.893 and 0.843. The classified results show that the SVM yields a very promising performance than the ANN in LULC classification of high resolution Landsat-8 satellite images.

Corresponding Author:

S.V.S.Prasad,
Departement of Electrical and Computer Engineering,
MLR Institute of Technology,
Hyderabad, Telangana, India-500043.
Email: Prasad.sista@gmail.com

1. INTRODUCTION

In recent times, satellite images are delivering an enormous source of information for studying the spatial and temporal variability of environmental conditions. It can be utilized in a number of applications which consists of making of mapping products for military and civil purposes, exploration, nursing of land use land cover [1], assessment of environmental damage, radiation level check, soil test, growth directive, crop outcome increment and urban planning. These multispectral images can be used mainly in the course of classification and also in mapping of vegetation over extensive spatial scales. This is because multispectral image classifies land cover and land usage features such as vegetation, water, oil, forests and urbanization, it also delivers very good scope and mapping. This kind of classification technique replaces the traditional classification techniques [2]. As many environmental and socio-economic proposals are based on these classification results, researches and studies on these satellite image classifications have centralized the concentration of the scientific community. A classification technique classifies relevant components or classes of land cover sections over a particular area [3]. Multi spectral images primarily consists the data that is collected over a wide range of frequencies and it changes for different locations [4]. This overall peculiar nature of satellite image data can be related to the spectral features that correlate with the spatial features belonging to the same band and this is called as spatial correlation. Classification of land use and land cover using remote sensing imagery is tough because of compound landscapes and also the spectral and spatial resolution of the satellite images.

Satellite image classification can be primarily considered as a combination of both image clustering and classification methods [5]. At present there are many image classification techniques being introduced under different categories like pattern recognition and image analysis [6]. Under some circumstances, the process of classification might be considered as an entity of analysis, but in other cases it is considered as a centre stage in complicated cases like degradation of land rules, landscape modeling, process studies, environment monitoring and management of coastal zones.

To overcome various problems of image classification great literatures of supervised techniques are created [7]. Among them 'Maximum likelihood' classifier is one statistical technique that was used previously for the classification process. But at present, artificial intelligence technique is applied to satellite image classifications [8]. To increase the precision of classification different techniques have been presented [9], but scientists have put great efforts in finding out an efficient system and method for increasing the accuracy of classification.

Primary focus of this research is to classify multi spectral image into land use and land cover. Characteristic features of land surface can be related to land cover [10]. Things that we can consider as land cover can be natural, semi natural or complete manmade structures that are observable. Whereas land use can be considered as usability of land, socio-economic activities like agriculture [11]. These are the commonly used classes of usage. Classification of land cover is one of the major inputs when considered in planning at local, regional and national levels. Here, an efficient classification technique is proposed to classify remote sensed images into land cover and land use by using SVM classifier [12]. This technique comprises of four phases namely: pre-processing, clustering based segmentation, training data selection for SVM and classification using trained SVM. As, the satellite image cannot be fed directly into SVM for training and testing, initially pre-processing is to be done so that image can be converted preferably for the process of segmentation. Later, fuzzy incorporated hierarchical based clustering algorithm is used for segmenting the image into clusters. After that centroid of each cluster is subjected to trained SVM and ANN Classifiers respectively [13].

This paper is discussed under five different headings. In section 2 Support Vector machine technique is elaborated. Section 3 presents the proposed technique and study area. Section 4 presents the experimental results and discussion. Section 5 gives the conclusion.

2. RESEARCH METHOD

2.1. Support Vector Machine (SVM)

SVM is one of the statistical based classifier. SVM sections mainly relate to the decision surface that magnifies the boundary among the classes. These surfaces are termed as optimal hyper plane and the data points that are close to the hyper plane are termed as support vectors. When we are considering the training set, support vectors are very important. Deviations of SVM are stated below.

1) By utilizing nonlinear kernels, SVM can be altered to a non linear classifier.
2) By grouping large number of binary SVM classifiers, multi classifier can be acquired.

Pair wise classification strategy is generally utilized for the purpose of multiclass classification. The output of SVM classification is the decision values for each pixel having a place with every individual class and it is mostly utilized for probability estimates.

When a binary class is considered, this classifier will possess a hyper plane that gradually decreases the distance from its nearby pixels belonging to each class with respect to the hyper plane. This classification can be defined as follows: Consider there are N training samples and they can be represented by set of pairs $\{(x_i, y_i), i = 1,2,3......, N)\}$ where x_i the class label of value is ± 1 and y_i ∈ n where it shows the vector comprising of n elements. Basically classifier function is given by $f(y;\alpha) \rightarrow x$ where α, is the measure of classifier element. Figure 1 shows Representation of maximum-margin hyper plane and margins for SVM trained with samples belonging to two classes.

There is an optimum separating hyper plane that was found out by SVM algorithm in such a way that: 1) Samples consisting of labels ±1 are situated on either side of the hyper plane. 2) The distance of nearest vector which is termed as optimal margin which are closer to the hyper plane is called as a support vector.

Usually the equation for a hyper plane is given by $w.y + b = 0$ where (w,b) are the parameter factors of hyper plane. The inequality $w.y + b > 0$ basically denotes vectors that are not on the hyper plane, so the classifier is usually represented as $f(y;\alpha) = sgm(w.y + b).$ Support vectors basically lie on the two hyper planes, which are parallel to optimal hyper plane, of equation $w.y + b = \pm 1$. Maximization of margin with these equations of two support vector hyper planes mainly contributes to the following constrained optimization problem $\min \frac{1}{2} \| w \|^2$ with $x_i (w.y + b) \geq 1, i = 1,2,......., M$.

For the purpose of classifying the satellite image into land use and land cover effectively, we use concepts like Hierarchical clustering and Fuzzy C-Means algorithms in our Hybrid technique.

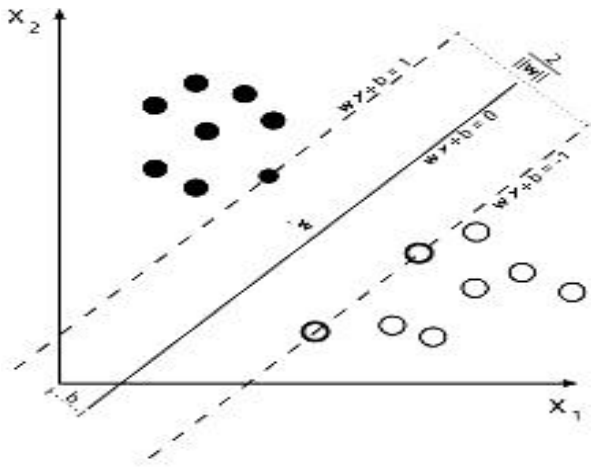


Figure 1. SVM classification for two different classes with margin levels

2.2. Hybrid Technique (Fuzzy Incorporated Hierarchical Clustering)

Here Hybrid technique will improve the drawbacks of both hierarchical and FCM clustering methods. This method is called as Fuzzy incorporated hierarchical clustering (Hybrid Technique). By using this technique accuracy level increases when we compare with the normal FCM.

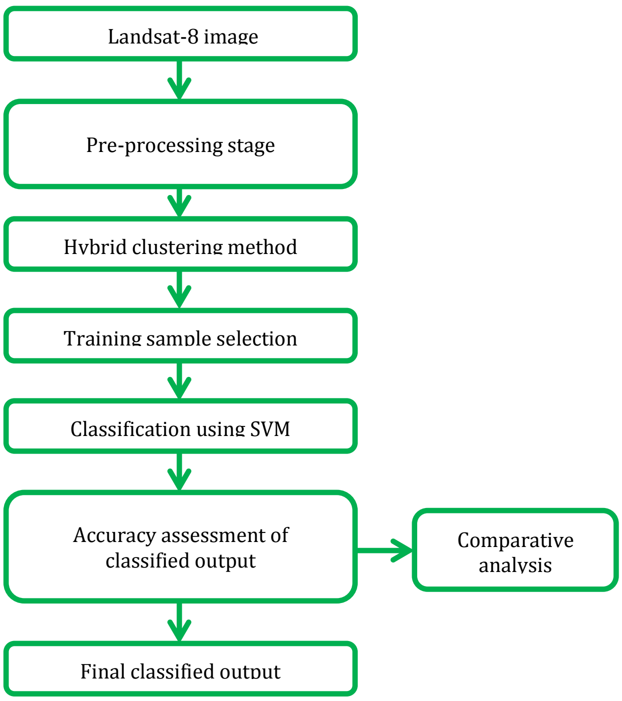


Figure 2. Methodology flow chart

As the satellite images cannot be fed directly into the SVM classifier for the purpose of training and testing. The input Landsat-8 image should undergo pre processing procedure so that image can be changed for further processing. In the second step of pre processing the given image is considered from RGB to Lab colour space which can be further segmented by using hybrid clustering technique. Figure 2 shows the block diagram for the proposed technique is shown below.

The resulting image that is obtained after pre processing consists of thousands of pixels and to classify based on singe pixel is a great task and also it is time consuming.. Hence, we group the pixels into clusters and for each cluster centroid is selected. This centroid will represent all pixels in a given cluster. Every pixel in a cluster will have nearly same pixel value and it differs only by the centroid value just by small amount. By classifying the centroid will indirectly classify all the pixels in a cluster. This in return decreases the inputs to the classifier system which mainly reduces the complexity and also time.

Now, we discuss about training data selection that is given to the SVM classifier for the purpose of classification. Classification is done effectively by making use of featured colours in the multispectral image. Every data element present on earth has a specific colour by which it is described. In-order to classify the given multispectral image using the SVM classifier, we use the colours of these earthly elements. Specific colours in the image represents for 'land use' and few for 'land cover'. After identification colour details are given to the SVM classifier for further classification purpose. Figure 3 shows the Chart describing colours with which elements of the earth are represented.

The Figure 3 mainly describes different colours and it's representation in satellite image. It also shows the classification of land use and land cover. Data elements that come under land use are roofs, concrete buildings and those that come under land cover mainly include vegetation, mud, soil and crops.

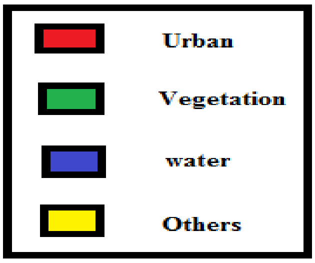


Figure 3. Chart describing colours with which elements of the earth are represented.

3. RESULTS AND ANALYSIS

3.1. Study Area

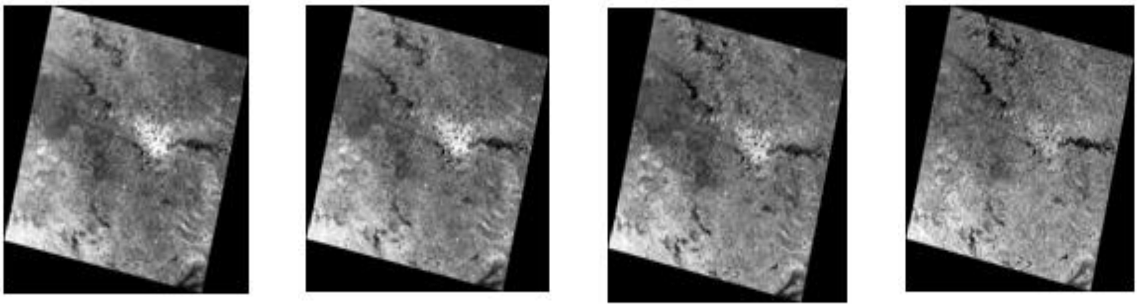
Here in this section we discuss about the multispectral satellite images. These images are captured from the Landsat-8 satellite. This observation satellite which is an American Earth satellite mainly consists of nine spectral bands. Among these bands, band 1 to band 7 have a spatial resolution of 30 meters and the resolution for band 8 is 15 meters and this band is panchromatic. Ultra-blue which is band 1 is mainly used for aerosol and coastal studies. Table 1 Shows all band Specifications of LandSat-8 Satellite.

Figure 4 shows all bands of the input multispectral satellite image of Hyderabad district in the year 2014 which was taken from the Landsat-8 satellite. Considering all 8 bands of the input image, Hyderabad city and its surroundings region is cropped based on its latitude and longitude values by the help of Hyderabad city and its surroundings topo-sheet. All these 8 bands are stacked to acquire the complete information provided by them.

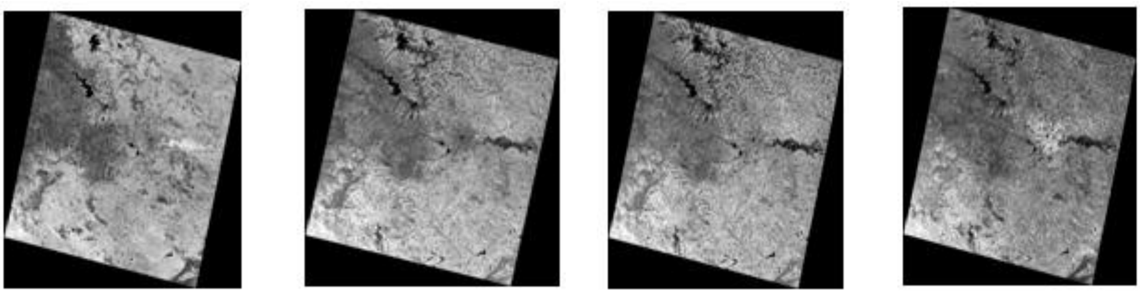
The methodology is tested on a Landsat-8 RS image which is shown in Figure 5 which is high resolution Gray-scale and false coloured RS images of Hyderabad city and its surroundigs which are obtained by combining all 8 bands of Landsat- 8.

Table 1. Landsat-8 Band Specifications

Satellite name	Bands	Wavelength (micrometers)	Resolution (meters)
Landsat-8	Band 1 Coastal aerosol	0.43 - 0.45	30
	Band 2 Blue	0.45 - 0.51	30
	Band 3 Green	0.53 - 0.59	30
	Band 4 Red	0.64 - 0.67	30
	Band 5 Near Infrared (NIR)	0.85 - 0.88	30
	Band 6 SWIR 1	1.57 - 1.65	30
	Band 7 SWIR 2	2.11 - 2.29	30
	Band 8 Panchromatic	0.50 - 0.68	15

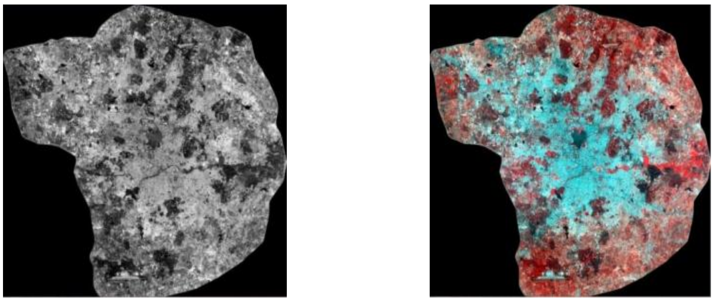


Band-1 Band-2 Band-3 Band-4



Band-5 Band-6 Band-7 Band-8

Figure 4. Input satellite image (Hyderabad district: year - 2014)



(a) Gray-scale (b) RGB

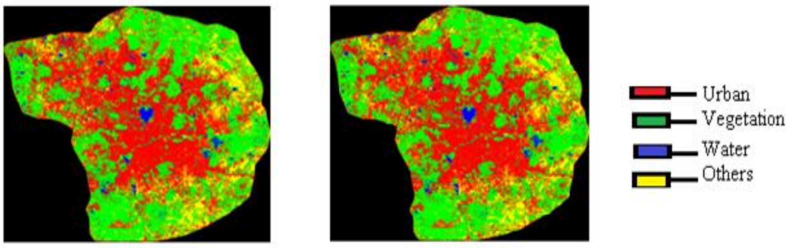
Figure 5 High resolution Gray-scale/ false colored RS image of Hyderabad city and its surroundigs area

3.2. Classification Results:

The study area was categorized into four land types, i.e. vegetation, Urban, water, and others (dry land, rock etc...) using the same region of Interest (ROI) training samples and covering a majority of the land cover features. Later, the training samples are considered for clustering Hyderabad city and its surroundigs area into classes using Fuzzy Incorporated Hierarchical clustering technique.

The classified images of GHMC using ANN and SVM classifiers are shown in Figure 6. In Figure 6, we can see different colors each color represents different class. In classification of Hyderabad city and its surroundigs area we considered four classes such as urban, vegetation, water and others. In Figure 6, red color indicates urban area, Green color indicates vegetation, Blue color indicates water and Yellow color indicates others.

In section 2, Figure 3 shows the chart describing colors with which elements of the earth are represented.



(a) (b)

Figure 6 (a) Classified image using ANN classifier (b) Classified image using SVM classifier

3.3. Accuracy Assessment

In this section, we present the performance evaluation results of the proposed technique. Here we have evaluated and compared the results with various clustering algorithms and also with various classifiers. The first part in this section, deals with the comparison with various clustering algorithms. In the later part, evaluation and comparison is made by comparing various classifiers. A detailed analysis, followed by the evaluation graph is made in each part. Analysis made and results obtained clearly demonstrate the efficiency of the proposed approach in classifying the multispectral image into land use and land cover regions.

The overall accuracy is the percentage of correctly classified pixels. Kappa coefficient is another criterion classically used in remote sensing classification to measure the degree of agreement and takes into account the correct classification that may have been obtained by chance by weighting the measured accuracies.

3.3.1. Evaluation using different Clustering Algorithm

For performance evaluation, the proposed technique (proposed clustering + SVM) is evaluated with the traditional clustering algorithm (proposed clustering + ANN). The accuracy value is computed by dividing the total number of similar pixels identified as land use to the number of pixels in the land use region. A summary of accuracy and kappa coefficient of ANN and SVM classifiers can be found in Table 2&3 respectively.

Table 2. Confusion matrix for ANN classification on Hyderabad city and its surroundigs area

CLASS	Vegetation	Urban	ater	Others	Total
Vegetation	504294	2374	1178	20442	528288
Urban	21778	525317	567	11856	559518
Water	299	0	33088	0	33387
Others	16179	15677	0	163351	195207
Total	542550	543368	34833	195649	1316400

Overall Accuracy: 89.925%
Kappa: 0.843

Table 3. Confusion matrix for SVM classification on Hyderabad city and its surroundigs area

CLASS	Vegetation	Urban	ater	Others	Total
Vegetation	491342	10124	2001	24821	528288
Urban	21480	507616	17566	12856	559518
Water	2047	0	31340	0	33387
Others	23989	19786	0	151432	195207
Total	538858	537526	50907	189109	1316400

Overall Accuracy: 93.159%
Kappa: 0.893

Here, Table 2 and 3 show the confusion matrices of the different methods i.e. ANN and SVM in land use classification. Among the two classified outputs the interesting observation was that the SVM got good classification results than ANN with an overall accuracy of 93.159% for SVM and 89.925% for ANN. The kappa coefficient values are 0.893 and 0.843 for SVM and ANN respectively. The obtained results are compared based on their accuracies of individual classes and it is indicated in Table 4.

Table 4: Comparison of Classification Accuracies

	Overall Identified Pixels (%)	
	ANN	SVM
Vegetation	93.00%	95.458%
Urban	90.724%	93.887%
Water	93.88%	99.10%
Others	77.575%	83.68%
Overall Accuracy	89.925%	93.159%
Kappa	0.843	0.893

From above table it is observed that the SVM classifier has detected all classes such as Vegetation, Urban, water and others comparatively better than ANN classifier.

4. CONCLUSION

In this paper, we explored the accuracy and reliability of SVM classifier for classifying multi-spectral image of Hyderabad city and its surroundigs area, India and also compared its performance with ANN classifier. Here in SVM, we have used an efficient image clustering technique named Fuzzy

Incorporated Hierarchical clustering to cluster the multispectral remote sensed satellite images. And their classification accuracies were compared using the Confusion matrix analysis.

The overall accuracies of LULC classification of the Hyderabad and its surrounding area are approximately 93.159% and 89.925% and the kappa coefficients are 0.893 and 0.843 for the ANN and SVM classification methods respectively. Although these LULC classification results classified by the ANN and the SVM both achieve relatively high accuracy and is adequate enough for the study on urban land use/cover classification, a great number of uncertainties exist in the practice due to many influencing factors such as training ROI selections, atmospheric situations, underlying surface conditions and mixing pixel effects. However, results are indicating that SVM has the better final classification accuracy among the two methods.

REFERENCES

[1] Hua-Mei Chen, Varshney, P.K. and Arora, M.K, "Performance of Mutual Information Similarity Measure for Registration of Multi Temporal Remote Sensing Images", *IEEE Transactions on Geoscience and Remote Sensing*, Vol.41 No.11, pp. 2445 – 2454, 2003.

[2] Ming-Hseng Tseng, Sheng-Jhe Chen, Gwo- Haur Hwang, Ming-Yu Shen, "A Genetic Algorithm Rule-based Approach for Land-cover Classification", *Journal of Photogrammetry and Remote Sensing* ,Vol.63, No.2, (3), pp. 202-212, 2008.

[3] S.V.S Prasad, Dr.T.Satya Savitri, Dr.I.V.Murali Krishna, "Classification of Multispectral Satellite Image using Clustering with SVM Classifier", *International Journal of Computer Applications*, Vol.35, No.5, pp. 4399-6107,2011.

[4] Huang C., L. S. Davis, J. R. G. Townshend 2002, "An Assessments of Support Vector Machines for Land cover Classification," *Int. J. Remote sensing*, vol. 23, no. 4, pp. 725–749, 2002.

[5] Anderson, J.R., E.E. Hardy, J.T. Roach, R.E. Witmer, 1976, "A Land Use and Land Cover Classification System for Use With Remote Sensor Data", USGS Professional Paper No 964, U.S. Geological Survey, U.S. Government Printing office, Washington, D.C.

[6] 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.

[7] Jan Knorn, Andreas Rabe, Volker C. Radeloff, Tobias Kuemmerle, Jacek Kozak, Patrick Hostert, "Land cover Mapping of Large areas using Chain Classification of Neighboring Landsat Satellite Images", *Remote Sensing of Environment*, Vol. 118, pages 957-964 , 2009.

[8] S.V.S Prasad, Dr.T.Satya Savitri, Dr.I.V.Murali Krishna., "Techniques in Image Classification; A Survey ", *Global Journal of Researches in Engineering: Electrical and Electronics Engineering* "Vol. 15, No. 6 , pp. 2249-4596, 2015.

[9] M. A. Hearst ; California Univ., Berkeley, CA ; S. T. Dumais ; E. Osman ; J. Platt, "Support Vector Machines", I*EEE Intelligent Systems and their Applications* (Volume:13 , Issue: 4), ISSN :1094-7167.

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

[11] Shivali A. Kar ; Vishakha V. Kelkar, *"Classification of Multispectral Satellite Images"*, Advances in Technology and Engineering (ICATE), 2013 International Conference on 23-25 Jan. 2013, ISBN: 978-1-4673-5618-3.

[12] Ranzhe Jing; Sch. of Inf. Manage. & Eng., Shanghai Univ. of Finance & Econ., Shanghai, China ; Yong Zhang, *"A View of Support Vector Machines Algorithm on Classification Problems"*, Multimedia Communications (Mediacom), 2010 International Conference on 7-8 Aug. 2010, Print ISBN: 978-0-7695-4136-5.

[13] Campbell, J.B.,1981, "Spatial Correlation Effects upon Accuracy of Supervised Classification of land cover", *Photogrammetric Engineering and Remote Sensing*, 47, 355–363.

[14] M. A. Hearst ; California Univ., Berkeley, CA ; S. T. Dumais ; E. Osman ; J. Platt, "Support Vector Machines", *IEEE Intelligent Systems and their Applications* (Volume:13 , Issue: 4), ISSN :1094-7167.

[15] Hussain Abu Dalbouh; Fac. of Sci. & Technol, Univ. Sains Islam Malaysia (USIM), Nilai, Malaysia ;Norita Md. Norwawi, *"Improvement on Agglomerative Hierarchical Clustering Algorithm Based on Tree Data Structure with Bidirectional Approach"*, 2012 Third International Conference on Intelligent Systems Modelling and Simulation 8-10 Feb. 2012, ISSN :2166-0662.

[16] Wang F,1990, "Fuzzy Supervised Classification of Remote Sensing Images", *IEEE Transactions on Geo-science and Remote Sensing*, 28, 194–201.