Australian Journal of Basic and Applied Sciences, 9(2) February 2015, Pages: 292-303

AENSI Journals

Australian Journal of Basic and Applied Sciences

ISSN:1991-8178

Journal home page: www.ajbasweb.com

Fuzzy Topology Based Satellite Cloud Image Classification using Support Vector Machine

[1]G. Devika and [2]Dr. S. Parthasarathy

[1]Department of MCA, K.L.N. College of Information Technology, Madurai
[2]Department MCA, Thiagarajar College of Engineering, Madurai

ARTICLE INFO

Article history:
Received 25 October 2014
Received in revised form
26 November 2014
Accepted 29 December 2014
Available online 15 January 2015

Keywords:
SVM, Fuzzy Topology, Image Classification, Satellite Image, Machine Learning Approach

ABSTRACT

Image classification is playing a vital role in the field of remote sensing, image analysis, pattern recognition, etc. Satellite images provide vital support to various scientific knowledge and technical applications and moreover it is an essential segment of graphic communication in the media. Hence, classification of satellite images has become an active research area in the field of image processing. Conventional classification approaches are inappropriate to handle complex high resolution satellite data. Several authors have studied classification methods for satellite images and most of the methods provide insignificant accuracy and results. This paper presents a novel image classification technique using an efficient machine learning technique. Support Vector Machine has been widely used in the classification purpose because of its efficient generalized property. This research work uses an improved Kernel version of Support Vector Machine with Fast Training . In order to improve the efficiency of the SVM, SVM is trained using a novel algorithm. Fuzzy-topology is than integrated with SVM classification method for remotely sensed images based on the standard SVM.

To Cite This Article: G. Devika and Dr. S. Parthasarathy., Fuzzy Topology Based Satellite Cloud Image Classification using Support Vector Machine. ***Aust. J. Basic & Appl. Sci.,*** *9(2): 292-303, 2015*

INTRODUCTION

IMAGE classification is one of the essential tasks in remote sensing, image analysis and pattern recognition. In some cases, the classification itself might be the object of the investigation (Aykut Akgun, 2004). For instance, categorization of cloud use from remotely sensed data generates a map like image as the end product of the analysis (Campbell, 2002). As a result, the image classification forms an essential tool for inspection of the digital images. The term classifier implies loosely to a computer program that executes a particular process for image classification (Campbell, 2002). The analyst must choose a classification technique that will best complete a particular task. At any time, it is not feasible to confirm which classifier is most excellent for all cases, since the features of each image and the situations for each study differ very much. As a result, it is indispensable that each analyst recognize the alternative methods for image classification in order that they may be prepared to choose the most suitable classifier for the specific task.

At present, operators obtain cartographic features manually, for instance, buildings, roads and trees, by means of visual interpretation of satellite imagery and aerial photography (Alonso, M.C. and J.A. Malpica, 2010). Semi-automatic approaches in order to support cartographic technicians possibly will improve the process. General classification approaches for low-resolution satellite imagery are extremely limited to handle complex high resolution satellite data and necessitate new algorithms. Large numbers of authors have studied classification approaches for satellite images (Bernardini, A., 2008; Alonso, M.C., 2007).

Satellite image processing plays a very important role for research and advancement in “remote sensing”, GIS, “agriculture monitoring”, disaster administration and several other fields of investigation (Navdeep Kaur Johal, 2010). On the other hand, processing these satellite images needs a huge amount of computation time because of its complicated and extensive processing criteria. The most common difficulty in an image obtained from an imaging device is its imperfection. The obtained image can be incompatible, imperfect, vague or an absolutely muddled. These are the major barriers in real time decision making. Because, these barriers lies in providing an improved and consistent technique which can provide better performance for digital image analysis (even in cases like improbability in gray level, texture, contours, edges detection, association between two

Corresponding Author: G. Devika, Department of MCA, K.L.N. College of Information Technology, Madurai

segments of an image and all other noisy input conditions), with greatest efficiency and reduced manpower utilization (Lillesand, T.M., 2003).

Satellite images have numerous properties embedded in them like spatial, spectral and temporal properties etc. Through these properties feature extraction can be carried out. The difficulty in obtaining homogeneous regions from an image can be considered as the task of grouping the pixels in the intensity space (Sriparna Saha and Sanghamitra Bandyopadhyay, 2008). Especially, satellite images include land cover types, some of which cover considerably large areas, at the same time some cover comparatively much smaller areas. Automatically identifying regions or clusters of such extensively unreliable sizes presents a difficult task (Ben-Hur, A. and I. Guyon, 2003). At some point in future, new generation of high-resolution satellite sensors will get huge data quantity, so obtaining features from a remote sensed image will turn out to be even more complicated. Researchers have been constantly searching for novel approaches that can obtain maximum information from the remotely sensed image.

Classification of image is carried out in three modes:

- Supervised,
- Unsupervised
- Hybrid

A supervised classification needs the manual identification of recognized surface features within the imagery and then by means of a statistical package to find out the spectral signature of the recognized feature (Bruzzone, L., 2006). The "spectral fingerprints" of the recognized features are then exploited to categorize the remaining image. An unsupervised classification method exploits spatial statistics (e.g. the ISODATA algorithm) to categorize the image into a predetermined number of categories (classes). These classes are statistically important within the imagery, however might not represent real surface features of significance. Hybrid classification utilizes both approaches to make the process more effective and perfect.

This research work proposes a novel classification algorithm using efficient machine learning technique. The main aim of this research is to enhance the recognition of pixels with a degree of uncertainty in the classification through a semisupervised classification technique based on the standard SVM. The efficiency of the SVM is enhanced through better training approach and novel weighting algorithm. The research mainly focuses to attain the important boundary and the interior parts of the classification in the fuzzy topology space.

Literature Survey:

The data acquired by remote sensing systems facilitate obtaining thematic maps of the earth's surface, via the registered image classification. This means the recognition and classification of all pixels into land cover classes. Conventionally, techniques based on statistical constraints have been extensively utilized, even though they show some demerits. On the other hand, few authors point out that those techniques depending on artificial intelligence possibly will be a better option. Thus, fuzzy classifier, which depends on fuzzy logic, includes supplementary information in the classification method through based-rule systems. Gordo *et al*., (2011) exploited Genetic Algorithm (GA) to choose the optimal and minimum set of fuzzy rules to categorize remotely sensed images. Input data of GA has been acquired through the training space decided by two uncorrelated spectral bands (2D scatter diagrams), which has been unevenly segmented by five linguistic terms provided in each band. This method has been implemented to Landsat-TM images and it has confirmed that this set of rules provides better accuracy.

Classification of multispectral remotely sensed data with textural characteristics is examined with a special concentration on uncertainty analysis in the obtained land-cover maps. Numerous efforts have already been directed into the research of acceptable accuracy-assessment approaches in image classification; however a common approach is not yet commonly adopted. Giacco *et al*., (2010) examined the association between hard accuracy and the vagueness on the produced answers, introducing two measures depending on maximum probability and α quadratic entropy. Their impact differs based on the type of classifier. The author handled with two different classification techniques, depending on Support Vector Machines (SVMs) and Kohonen's Self-Organizing Maps (SOMs), both properly enhanced to provide soft answers. Once the multiclass probability answer vector is presented for each pixel in the image, the author examined the behavior of the overall classification accuracy as a function of the uncertainty linked with each vector, provided a hard-labeled test set.

The efficiency of Multi Layer Perceptron (MLP) networks as a tool for the classification of remotely sensed images has been previously proven in history. On the other hand, several studies consider images characterized by high spatial resolution (around 15-30 m) at the same time a comprehensive analysis of the performance of this type of classifier on very high resolution images (around 1-2 m) such as those offered by the Quickbird satellite is still missing. Furthermore, the classification difficulty is normally understood as the classification of a single image whereas the capabilities of a single network of performing automatic classification and feature extraction over a collection of archived images has not been investigates until now. Del Frate *et al*., (2007) examined the generalization capabilities of this type of approaches with the intention of using them as a tool for completely automatic classification of collections of satellite images, either at extremely high or at high-resolution, in addition assessing the performance of MLP for the classification of very high resolution images.

Methodology:

The accuracy of SVM classification is affected by whether the training data can offer a representative description of each class or not, etc. Generally, the number of "pure" training pixels is maximum, then higher the classification accuracy which can be achieved. On the other hand, because of low image resolution, difficulty of ground substances, variety of disturbance, etc., several mixed pixels present in a remotely sensed image. Additionally, the spatial autocorrelation of pixels is not frequently considered in the classification; commonly, the classification result map includes much "salt and pepper" noise. As a result, performance of the SVM classifier based on accuracy is affected. With the purpose of making SVM achieve a higher accuracy, in this paper, the novel Improved Kernel SVM with Fast Training which is then integrated with Fuzzy Topology method by induced threshold fuzzy topology, which is integrated into the standard SVM, is proposed. In , the spectral space classification in the fuzzy topological space is described. In order to do this, the best possible threshold is required to decompose the classes into interior, boundary and exterior parts. Accordingly, fuzzy boundary pixels, which include several misclassified ones, are able to be reclassified, offering enhanced classification accuracy.

1.1. SVM Classifiers:

Support vector machines are a class of learning approaches which depends on the principle of Structural Risk Minimization (SRM) (Vladimir N. Vapnik, 1998; Vladimir N. Vapnik, 1995). SVM has been extensively utilized in several machine learning fields, for instance, classification, regression estimation and kernel PCA, in support of its better generalization capability.

SVM classifier uses inner product as metric. In case of dependent relationships among pattern's attributes, this data will be contained via extra dimensions, and this can be done through a mapping $\Phi: X \rightarrow H;\ \boldsymbol{x} \rightarrow \Phi(\boldsymbol{x})$, where *H* represents the feature space. The inner product similarity is formulated through $< \Phi(\boldsymbol{x}), \Phi(\boldsymbol{y}) >$ for pattern **x** and **y**. In SVM literature, the above course is realized through the kernel function.

$$k(x,y) = \langle \Phi(x), \Phi(y) \rangle \tag{1}$$

A Gaussian RBF kernel is formulated as

$$k(x,y) = exp\left(-\frac{||x-y||^2}{2\sigma^2}\right)\langle \Phi(x), \Phi(y) \rangle \tag{2}$$

Given a dataset $D = \{(x_i, y_i)\}_{i=1}^{l}$ of labeled patterns, in which $y_i \in \{-1, 1\}$, SVM classification (Girosi, F., 1997) is to build a hyperplane $w^T x + b = 0$ to categorize the patterns into two classes. This hyperplane is established by increasing the margin between two classes, i.e., optimizes such a primal problem:

$$\min J(w,b,\varepsilon) = \frac{1}{2}||w||^2 + C\sum_{i=1}^{l} \varepsilon_i$$

$$\text{s.t. } 0 \le \alpha_i \le C, i = 1, \dots, l$$

$$\sum_{i=1}^{l} a_i y_i = 0 \tag{3}$$

The coefficients a_i are the solution of the dual problem, and then the decision function is given as

$$f(x) = sign\left(\sum_{i=1}^{l} a_i y_i \langle x, x_i \rangle + b\right) \tag{4}$$

The weight vector w is given with support vectors by determining the derivatives of J according to the primal variables.

$$w = \sum_{i=1}^{l} a_i y_i\, x_i \tag{5}$$

By implementing kernel substituting tricks, based on eqn (1) and (4), the nonlinear decision function is obtained as:

$$f(x) = sign\left(\sum_{i=1}^{l} a_i y_i k(x, x_i) + b\right) \tag{6}$$

The main purpose of this work is to vary the width σ to eradicate the irregularity of the coexisting over-fitting and under-fitting in SVM learning. Changing the window size (i.e. width σ in Gaussian kernel) is the approach of kth nearest neighbor estimate (Loftsgaarden, D., C. Quesenberry, 1965) and the adaptive kernel estimate.

1.2. Weighting Kernel Width:

In order to decrease the coexisting over-fitting and under-fitting loss in support vector classification with the Gaussian RBF kernel, the kernel width is required to adjust, to some level, the feature space distribution. The scaling rule is that: in dense areas the width will happen to be narrow through some weights less than 1 and in sparse areas the width will be distended through some weights more than 1. Initially, several relationships are introduced.

(a) The relationship of σ and λ:

Equation (2) is rewritten in the new form of

$$k(x,y) = \exp\left(-\lambda||x-y||^2\right) \tag{7}$$

(b) The relationship of similarity and distance:

In feature space, in view of (1) and (7), the square distance of pattern x and y is given as,

$$d^2 = ||\Phi(x) - \Phi(y)||^2 = 2 - 2k(x,y) \tag{8}$$

It is revealed from equation (8), that the value of $k(\mathbf{x}, \mathbf{y})$ is inverse to d.

(c) Dense vs sparse in feature space:

When a pattern **x** drops in dense area its neighbors are close to it, and on the opposite in sparse area. Considered the x's *k*-NNs (*k* nearest neighbors) and obtained a value

$$sim_{knn(x)} = \frac{1}{k}\sum_i k(x,x_i), \qquad x_i \in k-nn \tag{9}$$

From (9) and the relationship *(b)*, it is observed that, *sim_knn*(**x**) is an index of density of **x**'s neighborhood. A large *sim_knn*(**x**) means a dense area where **x** lies. Selecting the **x**'s *k*-NNs is to select the *k* patterns which have kernel value *k*(**x**,.) not less than these of the rest part in training set.

(d) Scaling scheme:

In accordance with the weighting principle and the association *(a)* above, obtained a weighting scheme. Each pattern x has a λ*_weight*(**x**) which transforms the λ in (8). When the *sim_knn*(**x**) is huge (i.e. dense area), a huge λ*_weight*(**x**) will enhance the λ (i.e. diminish the width σ in (2)).

And on the converse a small *sim_knn*(**x**) indicates that a small λ*_weight*(**x**) is exploited. All the weights are confined to vary in an extremely small range around 1 to diminish distorting the metric induced by RBF kernel as possible. The modified Gaussian RBF kernel is given as

$$k(x,y) = \exp\left(-\lambda\_weight(x) \times \lambda\_weight(y) \times \lambda \times ||x-y||^2\right. \tag{10}$$

Now, the method to calculate the λ*_weight*(**x**) is explained. Consider a training set *S* with *l* patterns and the steps are given as follows.

1) Find out $sim\_knn(x_i), x_i \in S$, based on (9)
2) Calculate the mean of $sim\_knn(x_i)$, using the following equation

$$Mean\_sim\_knn = \frac{1}{l}\sum_i sim\_knn(x_i), x_i \in S$$

3) Calculate $\lambda\_weight\ (x)$ by using

$$\lambda\_weight\ (x) = 1 + \eta[sim\_knn(x) - mean\_sim\_knn]$$

In which $\eta \in [0,1]$ is a factor of weighting intensity. As a result, $\lambda\_weight\ (x)$ will vary within a very small range around 1.

1.3. Training SVM:

This paper also presents a four-step training technique for the purpose of increasing the efficiency of SVM. At first, SVM is initially trained by using all the training samples, by this means producing a number of support vectors. Then, the support vectors, which make the hyper-surface extremely convoluted, are removed from the training set. In third step, the SVM is re-trained only by the remaining samples in the training set. At last, the complexity of the trained SVM is further decreased by approximating the separation hyper-surface with a subset of the support vectors. When comparing to the initially trained SVM by all samples, the efficiency of the finally-trained SVM is extremely improved, without system degradation.

Based on the analysis discussed above, a four-step training algorithm for SVM is framed and it is given as below:

Step 1: Make use of all the training samples to train an initial SVM (Burges, C.J.C., 1998), resulting in l_1 support vectors $\{SV_i^{In}, i = 1, 2, \ldots, l_1\}$ and the equivalent decision function $d_1(x)$.

Step 2: Eliminate the support vectors from the training set, whose projections on the hyper-surface have the peak curvatures:

- 2a: For each support vector SV_i^{In}, calculate its projection on the hyper-surface, $p(SV_i^{In})$ along the gradient of decision function $d_1(x)$.
- 2b: For each support vector SV_i^{In}, calculate the generalized curvature of $p(SV_i^{In})$ on the hyper-surface, $c(SV_i^{In})$.
- 2c: Arrange SV_i^{In} in the decrease order of $c(SV_i^{In})$ and eliminate the top *n* percentage of support vectors from the training set.

Step 3: Exploit the remaining samples to re-train the SVM, resulting in l_2 support vectors $\{(SV_i^{Re}, i = 1, 2, \ldots, l_2\}$ and the equivalent decision function $d_2(x)$. On the whole, l_2 is typically less than l_1.

Step 4: Apply the l_2 pairs of data points $\{(SV_i^{Re}, d_2(SV_i^{Re})\}$ to finally train the SVRM, resulting in l_3 support vectors $\{(SV_i^{F1}, i = 1, 2, \ldots, l_3\}$ and the equivalent decision function $d_3(x)$. Mostly, l_3 is typically less than l_2.

SVM was initially designed for binary classification. However, most remote sensing applications involve dealing with several classes. Numerous approaches, including one-against-one, one-against- all, directed acyclic graph strategies, and multiclass SVM, have been proposed for multiclass classification problems. The one-against-one technique has been shown to execute better than the one-against-all technique, and the one-against-one technique is typically executed easily. As a result, in this paper, the one-against-one approach (Knerr, S., 1990) is adopted.

1.4. Posterior Probability Estimate for Image Classification:

Several techniques (Hsu, C.W. and C.J. Lin, 2002), including Bayes, Voting, Pairwise coupling, etc., have been proposed to make available the posterior probability output of multiclass SVM applications. In this paper, the shortcomings of the conventional voting method are analyzed, and an improved voting method for posterior probability output is proposed depending on the traditional voting method.

1) Traditional Voting Method:

If there are *N* classes in the classification to be classified, then $N(N-1)/2$ binary sub-classifiers will be built. A pixel with the highest probability is classified into the matching class. For the *i*th class, the posterior probability is given as follows

$$P(i|X) = \frac{\sum_{j=1, j\neq i}^{N} P_{ij}(i|X)}{\sum_{k=1}^{N}\sum_{j=1, j\neq k}^{N} P_{kj}(K|X)} \quad i = 1,2,\ldots,N \tag{11}$$

where $P(i|X)$ represents the posterior probability of sample *X* falling in the *i*th class and $P_{ij}(i|X)$ is the probability of sample *X* falling in the *i*th class by the SVM_{ij} classifier generated for the *i*th and *j*th classes. The value of $\hat{P}_{ij}(i|X)$ calculated for $P_{ij}(i|X)$ is 0 or 1.

Considering the definition of $P_{ij}(i|X)$, it becomes,

$$P_{ij}(i|X) + P_{ij}(j|X) = 1, \quad i = 1,2,\ldots,N \tag{12}$$

As a result,

$$\sum_{k=1}^{N}\sum_{j=1, j\neq k}^{N} P_{kj}(K|X) = \frac{N(N-1)}{2} \tag{13}$$

By integrating (11) and (13)

$$P(i|X) = \frac{2}{N(N-1)}\sum_{j=1, j\neq i}^{N} P_{ij}(i|X), \quad i = 1,2,\ldots N \tag{14}$$

Australian Journal of Basic and Applied Sciences, 9(2) February 2015, Pages: 292-303

Since the value of $\hat{P}_{ij}(i|X)$ is 0 or 1 in the conventional voting method, $P(i|X)$ is an estimated probability value which pays no attention to the probability of another class (the *j*th class) in the classifier SVM_{ij}. Therefore, an improved technique should be developed to provide a more accurate posterior probability.

2) Improved Voting Method:

As above mentioned, $P_{ij}(i|X)$ represents the probability of sample *X* coming under the *i*th class using the classifier SVM_{ij}. From a different perspective, $P_{ij}(i|X)$ is also considered as the posterior probability of sample *X* coming under the *i*th class considering the *j*th class. In accordance with the Bayesian probability formula, the posterior probability of sample *X* falling in the *i*th class can be obtained.

$$P(i|X) = \sum_{j=1, j\neq i}^{N} \{P_{ij}(i|X)P(j)\}, \quad i = 1,2, \ldots\ldots, N \tag{15}$$

where $P(j)$, the scale of importance of SVM_{ij}, which represents the conditional probability of the *j*th class, is an indefinite value.

At this point, $P(j)$ can be estimated by correlation between the posterior probability $P(j|X)$ and $P(j)$. Noticeably, an association exists between the significance of SVM*ij* and $P(j|X)$; if $P(j|X)$ is larger, the significance of SVM_{ij} is more when sample *X* is classified as the *i*th class based on SVM_{ij}. Based on the definition of $P(j)$, a noticeable similarity exists between $P(j)$ and $P(j|X)$; therefore, $P(j|X)$ can be exploited to approximately express $P(j)$ as

$$P(j) = kP(j|X) \tag{16}$$

where *k* is a balance factor.

Integrating equations (15) and (16),

$$\sum_{i=1}^{N} P(i|X) = k \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} P_{ij}(i|X)P(j|X), i = 1,2, \ldots N \tag{17}$$

where the left part of (18) is the full-probability event whose value is 1 and the right part of (13) can be obtained as

$$\begin{aligned} k\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} P_{ij}(i|X)P(j|X) &= k\sum_{j=1}^{N} P(j|X) \sum_{i=1, i\neq j}^{N} P_{ij}(i|X) \\ &\approx k\sum_{j=1}^{N} P(j|X)\frac{(N-1)}{2} \\ &= \frac{k(N-1)}{2} = 1 \end{aligned} \tag{18}$$

In proportion to equation (18), k=2/(N − 1). The final posterior probability of sample X falling in the ith class can be obtained as

$$P(i|X) = \frac{2}{N-1}\sum_{j=1, j\neq j}^{N}\{P_{ij}(i|X)\ P(j|X)\}\ \ i = 1,2, \ldots\ldots, N \tag{19}$$

Equation (19) can be considered as a set of multiple linear equations, where the unknown variables $P(i|X), i = 1, 2, \ldots, N,$ can be solved through $P_{ij}(i|X)$, *i*, *j* = 1, 2, . . . , *N*. Furthermore, $\sum_{i=1}^{N} P(i|X) = 1;$ thus, equation (19) is also given as follows:

$$\begin{bmatrix}\theta^T \\ \Omega\end{bmatrix} P = \begin{bmatrix}1 \\ o\end{bmatrix} \tag{20}$$

Where

$$\theta = [1 \quad 1 \quad \ldots. \quad 1]^T$$

$$\Omega = \begin{bmatrix} N-1 & -2P_{21}(2|X) & \ldots & -2P_{N1}(N|X) \\ -2P_{12}(1|X) & \mathrm{N}-1 & \ldots & -2P_{N2}(N|X) \\ \ldots & \ldots & \ldots & \ldots \\ -2P_{1N}(1|X) & -2P_{2N}(2|X) & \ldots & \mathrm{N}-1 \end{bmatrix}$$

$$O = [0 \quad 0 \quad \ldots. \quad 0]^T$$

and $P = [P(1|X)\ P(2|X) \dots P(N|X)]^T$ is the final posterior probability vector of sample X falling in the ith class. Equation (16) is an over determined equation which can be solved by the least squares technique. Depending on the $P_{ij}(i|X)$ obtained from all the classifier SVM_{ij}, the probability voting value $P(i|X)$ of sample X falling in the ith class can be obtained if $P(i|X)\ (i = 1, 2, \dots, N)$ is the maximum value and sample X will be categorized as the ith class.

1.5. Image Class and Fuzzy Set:

In conventional classification techniques, a pixel is classified to be only one class. One pixel may be enclosed by more than one class in a remotely sensed image. A fuzzy set in a fuzzy topological space idea is introduced into the classification, in which one pixel might contain multiple classes. Additionally, the percentage of each class in a pixel is characterized by the equivalent membership value, i.e., a pixel is considered as having a multiple membership value determined in proportion to the area proportion of each class within the mixed pixel. Fuzzy-topology theory can be implemented to conventional classifications for additional clarification of unclear class boundaries. Fuzzy topology is an expansion of ordinary topology. Based on (Chang, C.L.,1968), consider X be a non-empty ordinary set and I be an I-lattice. δ is recognized as an I-fuzzy topology if $\delta \subset Ix$ on X. (I^x, δ) is recognized as an I-fuzzy topological space (I-fts) if δ satisfies the following conditions.

1) An empty set and the entire set belong to δ.
2) When A and B belong to δ, then $A \wedge B \in \delta$.
3) Let $\{A_i : \in\ J\} \subset \delta$, where J is an index set; then $\vee_{i \in j} A_i \in \delta$.

For a fuzzy set A, the joint of all the open subsets in A is defined as the interior of A and is represented by A^0. The closure of A is represented by $\bar{A}$, which is the assemble of all the closed subsets containing A. In addition, $\partial A = \bar{A} \wedge \overline{A^c}$ is the boundary of A.

A fuzzy topological space structure, $\{(X_\alpha, \tau_\alpha, \tau^{1-\alpha}) : \alpha \in (0, 1)\}$, is stimulated by two threshold values (Liu, K.F. and W.Z. Shi, 2006), that is, α and $1 - \alpha$. Furthermore, these two threshold values (α and $1 - \alpha$) can be believed as the interior and closure operators, correspondingly. The fuzzy topological space has several prime properties under mapping.

In this paper, the dual threshold fuzzy topological space is employed for the decomposition of the class into interior, boundary and exterior parts using the threshold α depending on the SVM-estimated probability. Subsequently, the interior part pixels are then categorized as the particular classes, while the boundary part pixels have to be re categorized. Spatial connectivity theory in fuzzy topology is then employed to integrate the interior and boundary parts (Shi, W.Z. and K.F. Liu, 2007).

In this scenario, each class in the image can be considered as a fuzzy set in a fuzzy topological space. If there are N predefined classes, then each class has its interior and boundary. The association between classes might disjoint, contain, overlap, touch, etc.

1.6. Optimal Threshold Value:

Typically the association between two or more distributions can be concluded by the inter-correlation coefficient,

$$r_{xy} = \frac{\sum_{k=1}^{n}(x_k - \mu_x)^T(y_k - \mu_y)}{\sqrt{\sum_{k=1}^{n}|(x_k - \mu_x)|^2 \sum_{k=1}^{n}|(y_k - \mu_y)|^2}} \quad (21)$$

in which r_{xy} is an index that determines the magnitude and the direction of the association between the two distributions. It finds out the correlation between two or more distributions, its value ranges in [0, 1], and the larger r_{xy} is, the higher the inter-correlation is. In some case, where r_{xy} is positive, the inter-correlation is a positive correlation. In some other case, where r_{xy} is negative, the inter-correlation is a negative correlation.

Thresholding is a method that divides a class into regions. A threshold α separates a region into three segments: interior, boundary, and exterior.

Each class c_i in the image can be considered as a fuzzy set in a fuzzy topological space, represented by P_i. As a result, for each $\alpha \in [0, 1)$ and a class c_i the interior of c_i is $(P_{c_i})_\alpha$. For two classes, c_i and c_j the inter-correlation coefficient of these two classes is defined as (21). The thresholding value of these two classes can then be given as

$$\alpha = r_{ij} \text{ or } \alpha = |r_{ij}|^2 \text{ or } \alpha = \sqrt{|rij|)} \quad (22)$$

The inter-correlation thresholding coefficient for these two classes within the spectral space of two bands.

For the complete predefined N classes, the threshold value is then given as

$$\alpha = \max_{i,j} |r_{ij}| \tag{23}$$

A fuzzy topological space is then acquired from the value $\alpha = \max_{i,j} |rij|$ (Liu, K.F. and W.Z. Shi, 2006), and each class is now divided into the three segments: an interior, a boundary and an exterior. The interior of class c_i is represented as $(P_{c_i})_\alpha$, and the boundary of c_i is represented as $\partial P c_i = (Pc_i)^{1-\alpha} \wedge (Pc_i)^c_\alpha$. The interior segment of the class represents the significant part of this class; the boundary is the insignificant segment of this class.

1.7. Classification of Boundary Pixels of a Class:

The pixels on the image are divided as interior, boundary and exterior segments of a class. The pixels of the interior segments can be easily categorized as particular classes by evaluating the SVM membership value of the pixels on the image with the best possible threshold value α. On the other hand, it is more complicated to classify boundary pixels. Fuzzy topology can be employed to examine the structure of a neighborhood and the levels of spaces (Liu, Y.M. and M.K. Luo, 1997). When the class itself is assumed to be linked in the neighborhood sense, the interior and boundary can be reconstructed by using connection properties.

In a spatial analysis, the connectivity of spatial objects is significant and the majority of the spatial objects are supported (neighborhood) connected. As a result, the idea of supported connection for spatial objects can be exploited to reclassify the boundary pixels of the image.

In the application of fuzzy topology to land-cover classification, the neighborhood associations between classes are extremely essential. The theory of connectivity in (Liu, K.F. and W.Z. Shi, 2006) presents a tool to integrate the interior and the boundary pixel, which completely depends on a spatial object's connections in its spatial space.

1.8. Implementation of Fuzzy Topology based Improved Kernel SVM with Fast Training:

A new classification approach based on the FTSVM is presented in this paper for classification of remotely sensed images. Provided a multispectral remotely sensed data set, the multispectral remotely sensed data are initially classified depending on the SVM and the posterior probability for each pixel is calculated with the help of the improved voting method. Then, by means of the dual threshold fuzzy topological space for the segmentation of a class into interior, boundary and exterior parts with the help of the best possible threshold value α depending on posterior probabilities, the interior-part pixels of a certain class is classified as the particular class based on maximum likelihood, at the same time the boundary pixels of that class have to be reclassified. At last, the idea of the supported connection is employed to reclassify the boundary pixels.

The implementation of the FTSVM classification approach consists of the following five operational steps:

A. Selection of the Regions Of Interest (ROIs) or Samples:

In FTSVM, the training samples are indicated as the training set $\{x_i, y_i\}$ in which x_i represents the pixel values vector and y_i is the class label.

B. Estimation of the posterior probability of each pixel:

Initially, a Radial Basis Function (RBF) kernel for training and classification is used for the SVM analyses, and the parameters C and γ are set correspondingly using a "grid search" (Hsu, C.W., 2010) following a five-fold cross-validation approach. The training samples are trained using SVM, and the posterior probability of each pixel is then calculated using the improved voting method.

C. Computation of optimal threshold value;

With the intention of determining whether the segments of the pixels on the image are interiors of a class, boundaries of a class or exteriors of a class, the inter-correlation coefficients are computed based on (22); the threshold value is then calculated using (23).

D. Determination of the Interior and Boundary of a Class:

Determination of interior and boundary of a class for each pixel by using the best possible threshold value α, followed by the output of the interior of a certain class which is categorized to the particular class;

Using the optimal threshold value, for all pixel x_o and for every class c_i, consider $P_{k_o}(x_o) = \max_k P_k(x_o)$

1) When $P_{k_o}(x_o) > \alpha$, then pixel x_o belongs to the interior of class k_o. Subsequently, the pixel will be classified as belonging to that specific class and output.

2) When $P_{k_o}(x_o) \leq \alpha$, then pixel x_o belongs to the boundary of a particular class and those boundary pixels have to be further treated by Step 5.

Australian Journal of Basic and Applied Sciences, 9(2) February 2015, Pages: 292-303

E. Classification of the Boundary Pixels Based on Connectivity:

The theory of the supported connection can be exploited to reclassify the boundary. For each pixel of the boundary, by investigating its eight connected pixels and recording each connected pixel which belongs to the interior of a certain class, then this boundary pixel will belong to that class, previously allocated to the largest number of connected pixels.

Experimental Analysis:

The proposed classification approach depends on the extensively used SVM library LIBSVM in Matlab 7.8. A performance of the proposed approach was evaluated based on the accuracy and the training time. A satellite image was taken for the evaluation purpose and it is given as input into the classifier.

1.9. Performance Evaluation:

A. Accuracy of Classifier:

The performance of the proposed Fuzzy topology based is evaluated based on the accuracy of the classifier. The performance of Proposed Fuzzy topology based is compared with standard SVM classifier and Fuzzy Topology SVM classifier.

Table 1: Accuracy Comparison of the Classifiers.

Classifier	Accuracy
SVM	83%
FTSVM	87%
Proposed Fuzzy topology based	93.4%

Table 1 shows the comparison of the classification accuracy of the proposed approach with other approaches such as SVM and FTSVM. It is observed from the table that the proposed classifier provides 93.4% accuracy where as the standard SVM and FTSVM provides 83% and 87% accuracy respectively.

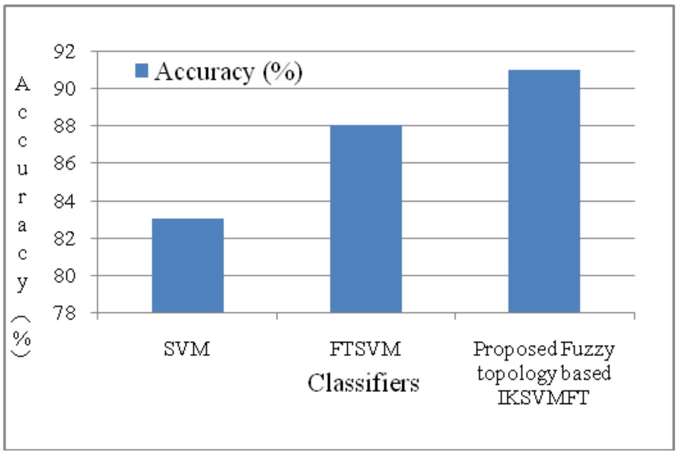

Fig. 1: Classification Accuracy Comparison.

B. Training Time:

Table 2: Training Time Comparison.

Classifier	Training Time (sec)
SVM	6.8
FTSVM	5.4
Proposed Fuzzy topology based	3.2

Table 2 shows the training time comparison of the proposed approach with the standard SVM and FTSVM classifiers. It is observed that the proposed approach outperforms the other classifiers in terms of training time.

1.10. Output Results:

Figure 2 shows the original satellite image taken for the experimental evaluation.

Australian Journal of Basic and Applied Sciences, 9(2) February 2015, Pages: 292-303

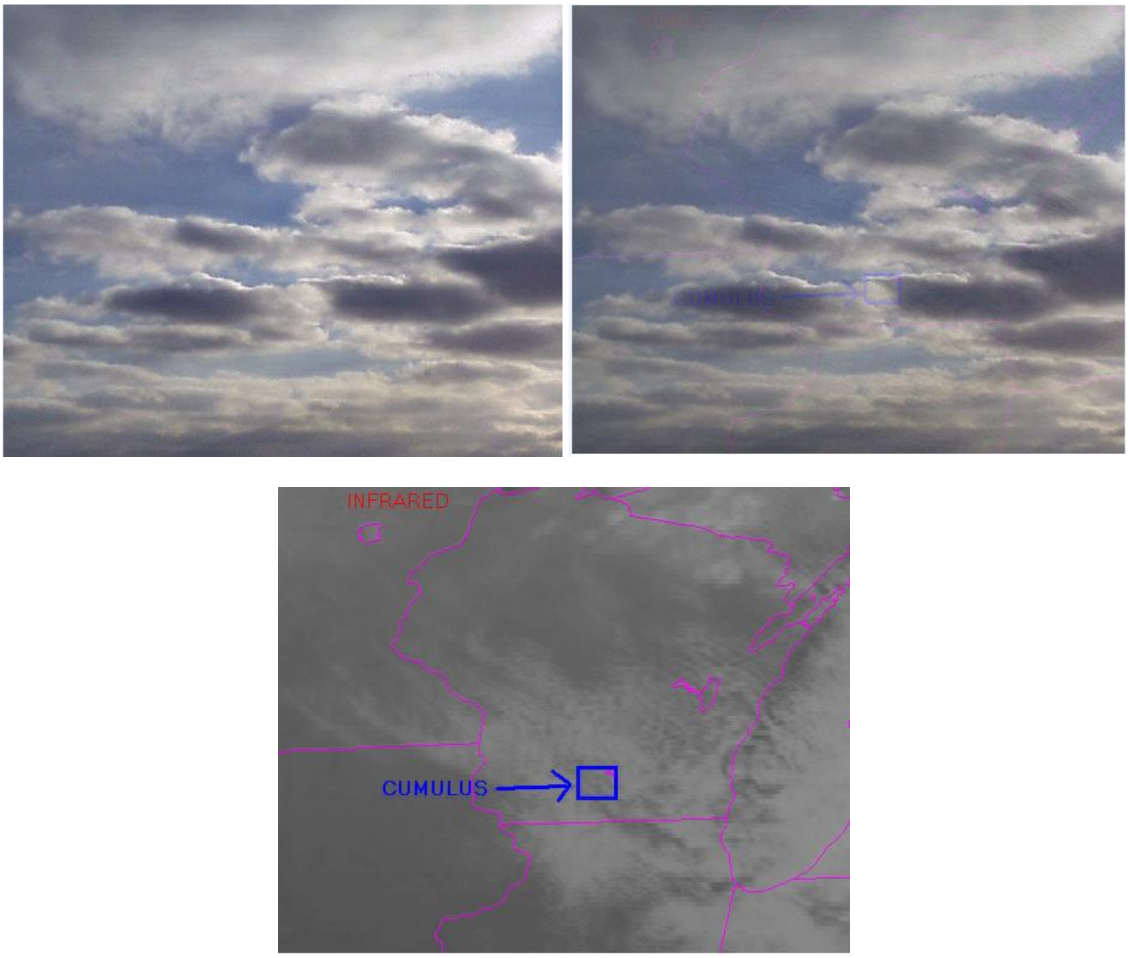

Fig. 2: Original Satellite Image.

Figure 3 shows the region growing output image of the proposed approach. Figure 4 shows the classified output image.

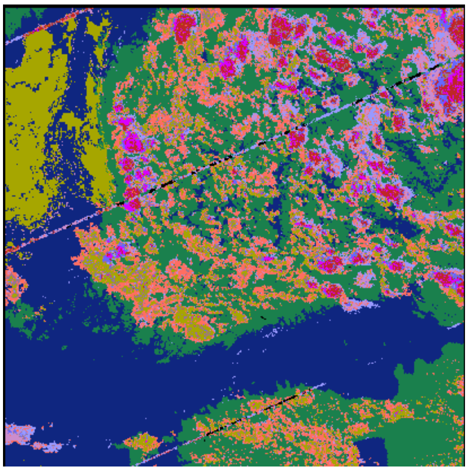

Fig. 3: Region Growing.

Australian Journal of Basic and Applied Sciences, 9(2) February 2015, Pages: 292-303

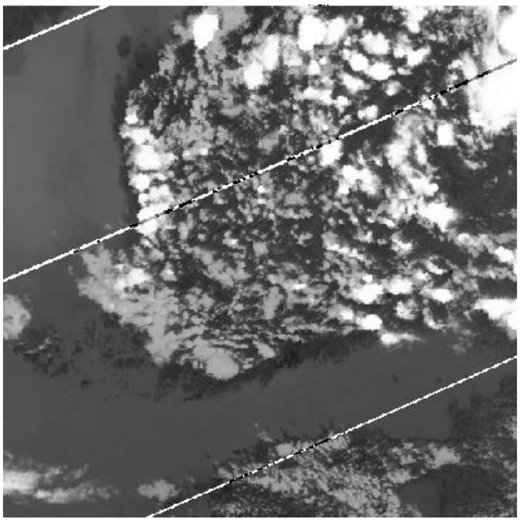

Fig. 4: Classified Output Image.

Conclusion:

Satellite image classification has become an active area of research in the field of image processing. This research work focuses on efficient classification of satellite images using efficient machine learning technique. The efficiency of the SVM classifier is enhanced through the utilization of improved kernel with Fast training. Thus, an Improved Kernel SVM with Fast Training is used in this research work. Then fuzzy topology based is used in this research work for the betterment of the overall performance. Initially, the satellite image is considered as a fuzzy set in a fuzzy set in a fuzzy topological space. Then, posterior probability of each pixel is computed by enhancing the voting technique. Ultimately, the pixels of an interior can be significantly classified as the specific classes depending on maximum likelihood. Thus, the detection of classifications in uncertainty pixels is improved greatly.

Thus, this research work focused on the improvement of the SVM techniques by incorporating improved kernel with fast training algorithm for image classification.

REFERENCES

Alonso, M.C. and J.A. Malpica, 2010. "Satellite Imagery Classification With LIDAR Data", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, XXXVIII, Part 8.

Alonso, M.C., A. Sanz and J.A. Malpica, 2007. "Classification of High Resolution Satellite Images using Texture", In Bebis G. et al. (Eds.): Advances in Visual Computing ISVC07. Part II LNCS 4842, pp: 499-508.

Aykut Akgun, A. Husnu Eronat and Necdet Turk, 2004. "Comparing Different Satellite Image Classification Methods: An Application in Ayvalik District,Western Turkey", XXth ISPRS Congress Technical Commission, IV: 1091-1097.

Ben-Hur, A. and I. Guyon, 2003. "Detecting Stable Clusters using Principal Component Analysis in Methods in Molecular Biology", Humana press.

Bernardini, A., E.S. Malinverni, P. Zingaretti and A. Mancini, 2008. "Automatic Classification Methods of High-Resolution Satellite Images: the Principal Component Analysis Applied to the Sample Training Set", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing, XXXVII. Part B7: 701-706.

Bruzzone, L., M. Chi and M. Marconcini, 2006. "A novel transductive SVM for semisupervised classification of remote-sensing images," IEEE Trans. Geosci. Remote Sens, 44(11): 3363-3373.

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition, Data Mining Knowledge Discovery, 2(2): 121-167.

Campbell, 2002. Introduction to Remote Sensing", Corine Land Cover Technical Guide, European Commission, Luxemburg, pp: 21-53.

Chang, C.L.,1968. "Fuzzy topological spaces," J. Math. Anal. Appl., 24: 182-190.

Del Frate, F., F. Pacifici, G. Schiavon and C. Solimini, 2007. "Use of Neural Networks for Automatic Classification From High-Resolution Images", IEEE Transactions on Geoscience and Remote Sensing, 45(4): 800-809.

Australian Journal of Basic and Applied Sciences, 9(2) February 2015, Pages: 292-303

Giacco, F., C. Thiel, L. Pugliese, S. Scarpetta, M. Marinaro, 2010. “Uncertainty Analysis for the Classification of Multispectral Satellite Images Using SVMs and SOMs”, IEEE Transactions on Geoscience and Remote Sensing, 48(10): 3769-3779.

Girosi, F., 1997. “An equivalence between sparse approximation and Support Vector Machines”, A.I. Memo 1606, MIT Artificial Intelligence Laboratory.

Gordo, O., E. Martinez, C. Gonzalo and A. Arquero, 2011. “Classification of Satellite Images by means of Fuzzy Rules generated by a Genetic Algorithm”, IEEE (Revista IEEE America Latina)Latin America Transactions, 9(1): 743-748.

Hsu, C.W. and C.J. Lin, 2002. “A comparison of methods for multi-class support vector machines,” IEEE Trans. Neural Netw., 13(2): 415-425.

Hsu, C.W., C.C. Chang and C.J. Lin, 2010. A Practical Guide to Support Vector Classification.

Knerr, S., L. Personnaz and G. Dreyfus, 1990. “Single-layer learning revisited: A stepwise procedure for building and training a neural network,” in Neurocomputing: Algorithms, Architectures and Applications, J. Fogelman, Ed. New York: Springer-Verlag.

Lillesand, T.M., R.W. Kiefer and J.W. Chipman, 2003.“Remote Sensing and Image Interpretation”, Fifth Edition, Wiley & Sons Ltd., England, pp: 586-592.

Liu, K.F. and W.Z. Shi, 2006. “Computation of fuzzy topological relations of spatial objects based on induced fuzzy topology,” Int. J. Geogr. Inf. Syst., 20(8): 857-883.

Liu, Y.M. and M.K. Luo, 1997. Fuzzy Topology. Singapore: World Scientific.

Loftsgaarden, D., C. Quesenberry, 1965. “A nonparametric estimate of a multivariate density function,” Ann. Math. Stat., 36: 1049-1051.

Navdeep Kaur Johal, Samandeep Singh and Harish Kundra, 2010. “A Hybrid FPAB/BBO Algorithm for Satellite Image Classification”, International Journal of Computer Applications (0975 – 8887), 6(5).

Shi, W.Z. and K.F. Liu, 2007. “A fuzzy topology for computing the interior, boundary, and exterior of spatial objects quantitatively in GIS,” Comput. Geosci., 33(7): 898-915.

Sriparna Saha and Sanghamitra Bandyopadhyay, 2008. “Fuzzy Symmetry Based Real-Coded Genetic Clustering Technique for Automatic Pixel Classification in Remote Sensing Imagery”, Fundamenta Informaticae, 84(3): 471-492.

Vladimir N. Vapnik, 1995. The Nature of Statistical Learning Theory, New York: Springer-Verlag.

Vladimir N. Vapnik, 1998. Statistical Learning Theory. New York: Wiley.