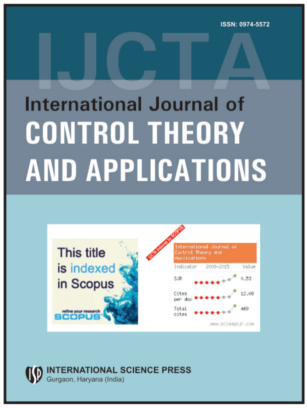


International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 9 • Number 44 • 2016

Rough Set based SVM Technique for Spatial Image Classification

D.N. Vasundhara[a] and M. Seetha[b]

[a]*Department of CSE, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana State, India. Email: vasundhara_d@vnrvjiet.in*
[b]*Department of CSE, G Narayanamma Institute of Technology and Science, Hyderabad, Telangana State, India. Email: smaddala2000@yahoo.com*

Abstract: There exist many traditional spatial image classification techniques which are developed over past years and exists in literature. Today, expert systems and machine learning methods are getting popularity in this area because of the effective classification.

In this paper, Rough set based Support vector machine classification method (RS-SVM) is proposed. In this technique, Rough set (RS) is used as a feature selection mathematical tool which eliminates the redundant features. Further, this reduced dimensionality data set is given to Support vector machine (SVM) classifier. This process improves the classification accuracy and performance. We have performed experiments using standard geospatial images for above-proposed method for classification.

Keywords: Feature Extraction, Classification, Rough sets, Artificial neural networks, Support vector machines.

1. INTRODUCTION

Image classification is concerned with identifying different types of images on the earth's surface, satellite collected image and spatial image and has been used for many purposes, such as urban planning and management, forestry, environmental monitoring, agriculture, etc.[1] Methodologies for classification of images can be divided into the two broad categories of supervised and unsupervised learning strategies. Supervised classification method takes an object, it is typically a vector as an input and outputs a desired value. In unsupervised classification method, a large number of unknown pixels can be examined. It divides all pixels into a number of classes within image. The need of input in the unsupervised classification method is the number of classes.

In the image classification, the images consist of allocating a particular class membership to each pixel present in the image, in this instance based exclusively on its spectral signature.[2] Many machine learning techniques are applied to the classification task in order to achieve improved classification performance and accuracy. Rough set theory evolved as a novel approach to manage ambiguity that has been used for the discovery of data dependencies, importance of features, patterns in sample data, feature space dimensionality reduction, and the classification of images[3][4].

Compared with other classification techniques such as discriminant analysis, random forest methods and artificial neural networks, Support vector machine (SVM) has been shown more advantageous in handling classification tasks with excellent generalization performance. SVM which is a relatively new machine learning technique, was first proposed by Vapnik in 1995 [5], which seeks to minimize the upper bound of the generalization error based on the structural risk minimization (SRM) principal that is known to have high generalization performance. Another key feature of SVM is that training SVMs is equivalent to solving a linear constrained quadratic programming problem.

A. Related Work

There exist many techniques in literature for data classification procedure, which have been proposed in past years. In 1996, Quinlan et. al., [21] proposed the classification technique using decision tree method. In 2003, Abonyi et. al., [6] proposed a supervised pattern recognition technique, which uses fuzzy classifier for the purpose of fuzzy clustering. In 1992, Boser et. al., [7] presented the training algorithmic procedure for the optimal margin classifier. In 2003, Frohlich et. al., [9] presented Feature selection procedure for support vector machines by means of genetic algorithms. In 2002, Goodman et. al., [10] proposed artificial immune system classification of multiple-class problems. Joachims et. al., [12] proposed some efficient methods for text categorization using support vector machines. Osuna et. al., [15] performed the support vector machine (SVM) concept in the process of face detection. In 1982 (Pawlak, 1982) proposed Rough Set (RS) concept [16][17][18] as a mathematical model to represent knowledge and to treat uncertainty. In 2010, U. Stanczyk [26] presented the rough set based features analysis for ANN classifier. Later, Vasundhara et. al., [27] proposed Rough-Set and Artificial Neural Networks Based mechanism for image classification. Rough set phenomenon can be applied in various domains. Specially it's applications are so wide in machine learning areas and knowledge based decision support systems.

B. Motivation and Contribution

Feature selection is a major issue in building the classification systems, which identifies the significant features and eliminate the irrelevant ones in order to build a good learning model. So a technique that can reduce dimensionality without any prior knowledge just using the information contained within the data set and preserving the meaning of the original features is strongly desirable. Rough Set theory can be utilized as such a tool to discover data dependencies and reduce the number of attributes (features) in a data set. RS theory plays a vital role to further improve the classification accuracy of the SVM model. Here RS-SVM model is proposed. RS-SVM process consists of two-stages. In the first stage, RS is applied as an attribute features reduction tool to extract the optimal features. This provides elimination of unnecessary data. In second stage, the optimal feature subset is used as the inputs to SVM classifier with good generalization and classification performance.

2. PRELIMINARIES

A. Rough Set Theory

The performance of any image classification technique majorly depends on the feature selection method. One of such method is rough set theory.

Rough Set Theory (RST) was proposed by Zdzislaw Pawlak in 1982, as a mathematical model to interpret knowledge and to handle uncertainty. An essential concept in RST is the reduct. A reduct is the minimal set of attributes (features) that can represent an object with the same accuracy and precision as it was represented by the original set of attributes. Exclusion of redundant attributes can support in the identification of strong, effective and non-redundant classification rules.

A decision table (also called Data Table) is a system of the form S = (U, C $\cup$ {*d*}) where, C: set of conditional attributes; *d*: decision attribute. In rough set theory, two terms- Lower Approximation & Upper Approximation, come under the Approximation of sets.

Let S = (U, R) is an approximation space and X be a concept in that space, then, the lower approximation is defined as

$$RX = \{x \in U \mid [x] \subseteq X\}$$

The upper approximation is defined as:

$$RX = \{x \in U \mid [x] \cap X \neq \Phi\}$$

where, $[x]$ is called an equivalence class which contains an element *e*.

Let, S = (U, C $\cup$ {*d*}) is a decision system, consisting the universe of objects, then a subset R of conditional attributes (C) is a reduct if, $POS_R(d) = POS_C(d)$. Computing the reduct is known to be an NP-hard problem, and processing of the reduct for large databases requires high computational processing. The intersection of all reducts is called as Core.

$$CORE\ (C) = \cap\ RED\ (C)$$

The reduct is generated from the discernibility matrix. Objects are distinguishable if for some attributes, they are having different attribute values, then they are called discernible. And this property is known as discernibility. Fuzzy Sets involve the membership among the elements from the same class, while RS bothers the relationship among groups of elements in distinct classes.

However, the theory of RS has no competence with the Fuzzy Sets Theory but rather complements it. In fact, Rough Set theory and Fuzzy Set theory are two individualistic strategies to handle the imprecise knowledge. The knowledge gaining bottleneck is a notable problem that limits the building of smart and intelligent monitoring rule systems. The generation of great knowledge bases for this job is tough. This problem is especially common where authorities are not easily available. Machine learning methods (especially rule induction methods) can be of excellent interest to this area by contributing strategies to automatically derive the useful information, given adequate amount of historical data.

Feature selection or also called Reduct generation (in Rough Set theory) intends to determine a minimal feature subset which can represent the same knowledge as it was described by the original features. Rough set theory (RST) has been employed as such a mathematical tool with much success. RST facilitates the discovery of data dependencies and the reduction of the number of attributes or features, enclosed in a dataset using the data only, requiring no additional information like probabilistic occurrence of data or other statistical information. he performance of any image classification technique majorly depends on the feature selection method. One of such method is rough set theory.

B. Support Vector Machines-SVM

Support vector machines (SVM) is originally developed by Boser et. al.(1992) and Vapnik (1995) [5], which is based on the Vapnik Chervonenkis (VC) theory and structural risk minimization (SRM) principle. SVM [23] is one of the best recognized techniques in pattern recognition and image classification purpose. It has been designed to divide a set of training images into two different classes in the d-dimensional feature space. SVM models the optimal separating hyper planes based on the kernel function. All the images, for which the feature vector lies on one side of the plane, belong to the class – 1 and the others belong to class +1.

SVM is capable of providing good performance and approximate accuracy in classification [24][25]. The parameters which effects performance of SVM algorithm is kernel functions, Gaussian kernel, standard deviation of the data and also depends on weights of the corresponding slack variable. Number of training examples also impact on performance of SVM classifier.

3. PROPOSED ALGORITHM: RS-SVM

In this section, we present our proposed Rough set based Support vector machine classification method (RS-SVM). In this technique, Rough set (RS) is used as a feature selection mathematical tool which eliminates the redundant features. Further, this reduced dimensionality data set is given to Support vector machine (SVM) classifier.

Motivation:

- While ANN can suffer from multiple local minima, so the solution to an SVM is global and unique.
- SVMs have a simple geometric interpretation and give a sparse solution.
- The computational complexity of SVMs does not depend on the dimensionality of the input space.
- ANN uses the principle of empirical risk minimization (ERM), while SVMs are based on the principle of structural risk minimization (SRM).
- SVMs are less prone to over fitting.
- One advantageous key feature of using SVM is that training the SVM is somewhat like solving a linear constrained quadratic programming problem.

The RS-SVM system operates as follows:

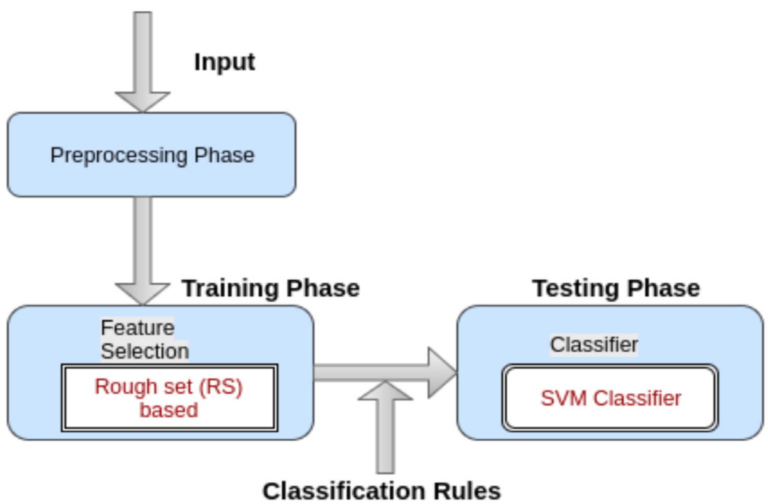


Figure 1: RS-SVM Procedure

A. Algorithm

Feature selection is an important step while building classification systems. With this technique, the dimensionality is reduced. Here, an expert system called RS-SVM is proposed, which consists of two steps. In first step rough set based attribute reduction algorithm is applied for the purpose of removing redundant features (attributes) from the data set. After this stage, unnecessary data, which further serves as improved classification speed. In the second step, the optimal feature subset are used as the inputs to SVM classifier with better generalization performance.

B. RS-SVM Algorithm Steps are as Follows

Feature Selection using Rough Sets

1. Input data-set is in the form of spatial images. As we know, images consist of pixels. In the pre-processing phase, basic image information for each pixel (e.g. RGB values, pixel intensity values, gray code level value etc.) will be obtained. These will be called attribute values for each pixel.

2. Draw the pixel, attribute-value information table. An Information System is a pair I = {P, A} where- P{P_1, P_2, ..., P_n}: Non-empty finite set of each individual pixel.

   A{A_1, A_2, ..., A_n}: Non-empty finite set of attributes.

   The process in which universal set is classified into modules or equivalence classes based on some attribute values, is called Granularization.

4. Draw the Discernibility matrix for the above pixel, attribute-value information. *A discernibility matrix of an information Table 1 = (P, A) is a symmetric | P | × | P | matrix in which the entries are defined as:*

$$c_{ij} = \{a \in A \mid a(x_i) \neq a(x_j)\} \; i, j = 1, ..., | P |$$

   Each c_{ij} consists of those attributes that differ between objects *i* and *j*.

4. For each row of the discernibility matrix, compute discernibility function as:

$$f_1, f_2, ..., f_{|P|} = \wedge \{\vee c_{ij} \mid 1 \leq j \leq i \leq | P |, c_{ij} \neq \emptyset\}$$

5. Compute resultant discernibility function as:

$$F = \wedge \{f_i, 1 \leq i \leq | P |\}$$

6. Attributes shrinking algorithm while computing (F, $\{f_i, 1 \leq i \leq | P |\}$)

   The algorithm steps are as below:

   - while ((F, $\{f_i, 1 \leq i \leq | P |\}$) are not in minimized form)
   - remove the super-sets by applying Absorption Rule in (F, $\{f_i, 1 \leq i \leq | P |\}$)
   - replace strong shrinkable attributes (F, $\{f_i, 1 \leq i \leq | P |\}$) strong shrinkability implements where the clause attributes are either together present or missing in all clauses. In that case, those attributes may be substituted by a single temporary attribute.
   - *s* <- most_frequent attribute (F, $\{f_i, 1 \leq i \leq | P |\}$)
   - apply Expansion Rule (s, (F, $\{f_i, 1 \leq i \leq | P |\}$))
   - substitute the strong shrinkable classes (F, $\{f_i, 1 \leq i \leq | P |\}$)
   - RED <- calculate the reducts (F, $\{f_i, 1 \leq i \leq | P |\}$)
   - return (RED)

7. The set of all prime implicants in discernibility function is obtained. In RST point of view these all are called as "Reducts". In general, it is the process of feature selection in rough set theory.

C. Classification using SVM Classifier

1. Consider a binary classification:

$$\{X_i, Y_i\},\ i = 1, \dots, l$$

$$Y_i \in \{-1, +1\},\ X_i \in R^d$$

Where, X_i: data points; Y_i: their corresponding levels. Note: features are represented by dimensions.

2. Data points are separated by a hyper plane:

$$W^T X + b = 0$$

   - At least one hyper plane is required to separate data points i.e. classification purpose.
   - Even though there can be more than one hyper planes.

3. Cases for classification task:
   - Case 1: Suppose a scenario with n hyper planes, select the hyper plane, which classifies set of data points (classes) in good manner.
   - Case 2: Suppose we have n hyper planes and all these hyper planes are separating the classes well. Then, decide the hyper plane by maximizing the margin vector between hyper plane and the nearest data points.
   - Case 3: If both two cases discussed above occurs simultaneously, give preference to case 1, over case 2, while classification.

4. Let, two hyper planes:

$$h1\text{: } W^T X^+ + b = 1$$

$$h2\text{: } W^T X^- + b = -1$$

where, h_1 is for nearest points on positive side. h_2 is for nearest points on negative side. W is an n-dimensional coefficient vector. b is offset from the origin. maximal width $W_m = 2/|W|$

5. To maximize margin, minimize $g(W) = (½)|W|^2$

   with condition, $\forall i,\ Y_i(W^T X_i + b) \geq 1$

4. EXPERIMENTAL RESULTS

In this section, we will present our experimental analysis which has been performed on real time spatial image data set. Details about the data set will be given in below subsections. We have mainly used R language environment for all our statistical computing and implementation purpose. Other software and hardware specifications are given below:

- Software Specifications

  OS - Ubuntu 16.04 LTS, 64 bit

  R Language Environment version 3.3.2

  R Studio IDE 1.0.44

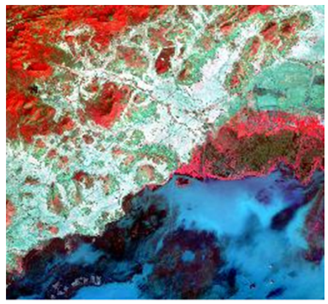
- Hardware Specifications

  RAM size - 4 GB

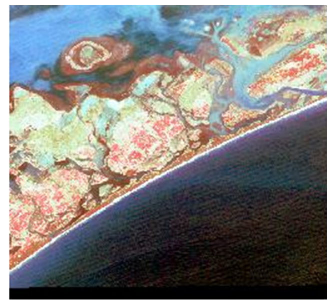
  Processor - Intel core i3 4030U CPU @1.90GHz *x* 4

A. Input and Setup

In our experiment we have taken some coastal region images of Andhra Pradesh from Bhuvan NRSC geospatial data[37]. Images are shown as below:

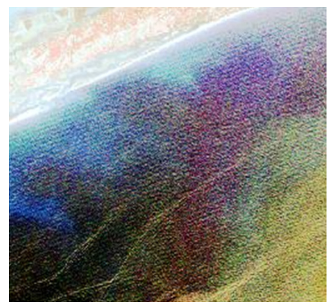


(a)



(b)

(c)



(d)

Figure 2

The specific technical details of above images are as:

Figure 2(a) Toposheet no: E45B05; Bounding box: 85.25E19.75N - 85.5E20.0N; Date of pass: 09mar13

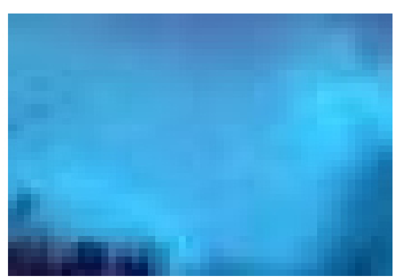
Figure 2(b) Toposheet no: E45B06; Bounding box: 85.25E19.5N - 85.5E19.75N; Date of pass: 09mar13

Figure 2(c) Toposheet no: E45B09; Bounding box: 85.5E19.75N - 85.75E20.0N; Date of pass: 09mar13

Figure 2(d) Toposheet no: E45B10; Bounding box: 85.5E19.5N - 85.75E19.75N; Date of pass: 09mar13

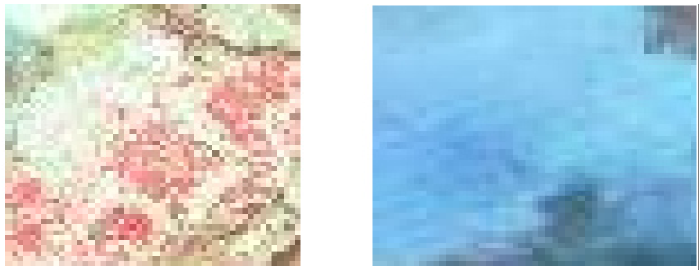
On these images, we have performed the multispectral image computing process. We have performed the RS based feature selection and later the classification based on SVM process, which is a supervised learning mechanism. We have performed our experiments in R environment. For each input image, we also need some images containing two jpeg images for each of the original input image. These images will separately consist of pixel samples in the form of patches, denoting coastal land and water pixels information. The pixel patches which we have taken in our experiment for classifier training phase are represented as below:

No_water patch



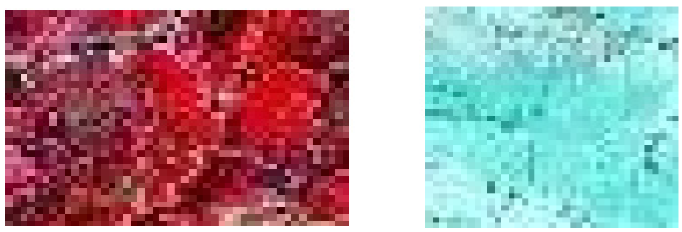
water patch

(a)



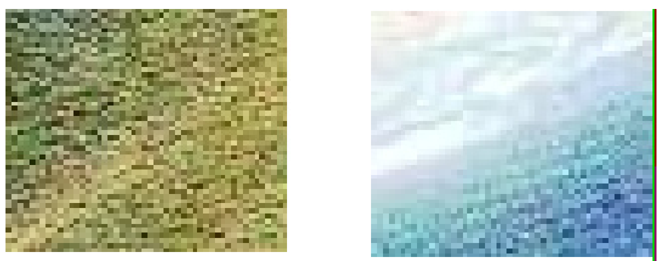
No_water patch water patch

(b)



No_water patch water patch

(c)



No_water patch water patch

(d)

Figure 3

Training and testing phases are performed for the image data set and later our classifier categorize each individual pixel into separate decision classes by finding the hyper plane in the n-dimensional space (here, n represents the number of features which we take).

B. Procedure

In our experimental process of RS-SVM based classification, initially we have different input images specified in above subsection and their pre-classified sample images containing pixel patches. These images are then loaded into R.

The experimental metadata output for input image represented by Figure 2(a) is as below:

dimensions :	256, 256, 65536, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 256, 0, 256 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	L3.NE45B05.09mar13.1, L3.NE45B05.09mar13.2, L3.NE45B05.09mar13.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(b) is as below:-

dimensions :	256, 256, 65536, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 256, 0, 256 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	L3.NE45B06.09mar13.1, L3.NE45B06.09mar13.2, L3.NE45B06.09mar13.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(c) is as below:-

dimensions :	256, 256, 65536, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 256, 0, 256 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	L3.NE45B09.09mar13.1, L3.NE45B09.09mar13.2, L3.NE45B09.09mar13.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(d) is as below:-

dimensions :	256, 256, 65536, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 256, 0, 256 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	L3.NE45B10.09mar13.1, L3.NE45B10.09mar13.2, L3.NE45B10.09mar13.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(a')-water patch is as below:-

dimensions :	38, 50, 1900, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 50, 0, 38 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	water052_train.1, water052_train.2, water052_train.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(a') - non-water patch is as below:-

dimensions :	114, 136, 15504, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 136, 0, 114 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	ground051_train.1, ground051_train.2, ground051_train.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(b')-water patch is as below:

dimensions :	44, 49, 2156, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 49, 0, 44 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	water062_train.1, water062_train.2, water062_train.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(b') - non-water patch is as below:-

dimensions :	50, 50, 2500, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 50, 0, 50 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	ground061_train.1, ground061_train.2, ground061_train.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(c')-water patch is as below:-

dimensions :	37, 37, 1369, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 37, 0, 37 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	water092_train.1, water092_train.2, water092_train.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(c') - non-water patch is as below:

dimensions :	34, 47, 1598, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 47, 0, 34 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	ground091_train.1, ground091_train.2, ground091_train.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(d')-water patch is as below:-

dimensions :	50, 50, 2500, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 50, 0, 50 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	water102_train.1, water102_train.2, water102_train.3
min values :	0, 0, 0
max values :	255, 255, 255

The experimental metadata output for input image represented by Figure 2(d') - non-water patch is as below:-

dimensions :	50, 50, 2500, 3 (nrow, ncol, ncell, nlayers)
resolution :	1, 1 (x, y)
extent :	0, 50, 0, 50 (xmin, xmax, ymin, ymax)
coord. ref :	NA
names :	ground101_train.1, ground101_train.2, ground101_train.3
min values :	0, 0, 0
max values :	255, 255, 255

Then on training data, feature selection process is performed. Then these pixel images are transformed into data frames format. Now they can be used as part of training data. Then corresponding factor variables are assigned. For the purpose of SVM classifier model fitting, random sampling of pixels is performed. Now with this setup cross validation process is performed, in which original data is randomly being partitioned in k-size sub-data samples. Then out of that one sub-data sample is utilized as validation data for testing predictive model and rest of (k – 1) sub-data samples are utilized as the training data. Predictions are fitted into classifier.

The resulting error matrix through random sampling by cross validation process is represented as below:-

Error matrix for 1st case input image experiment:

```
            true
pred        Nowater Water
  Nowater      2508     0
  Water           0   932
```

Error matrix for 2nd case input image experiment:

```
            true
pred        Nowater Water
  Nowater      2519     0
  Water           0   921
```

Error matrix for 3rd case input image experiment:

```
            true
pred        Nowater Water
  Nowater      2497     0
  Water           0   943
```

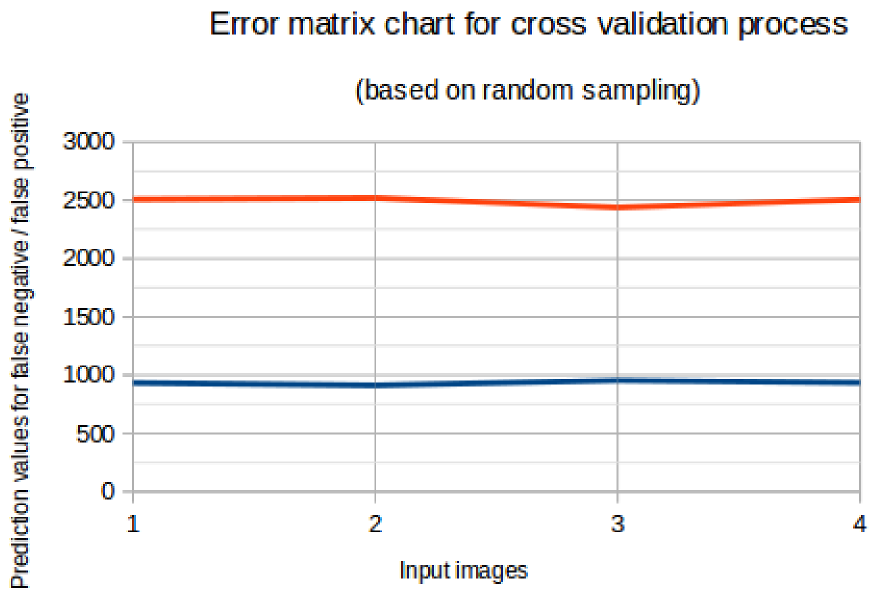
Error matrix for 4th case input image experiment:

```
            true
pred        Nowater Water
  Nowater      2506     0
  Water           0   934
```

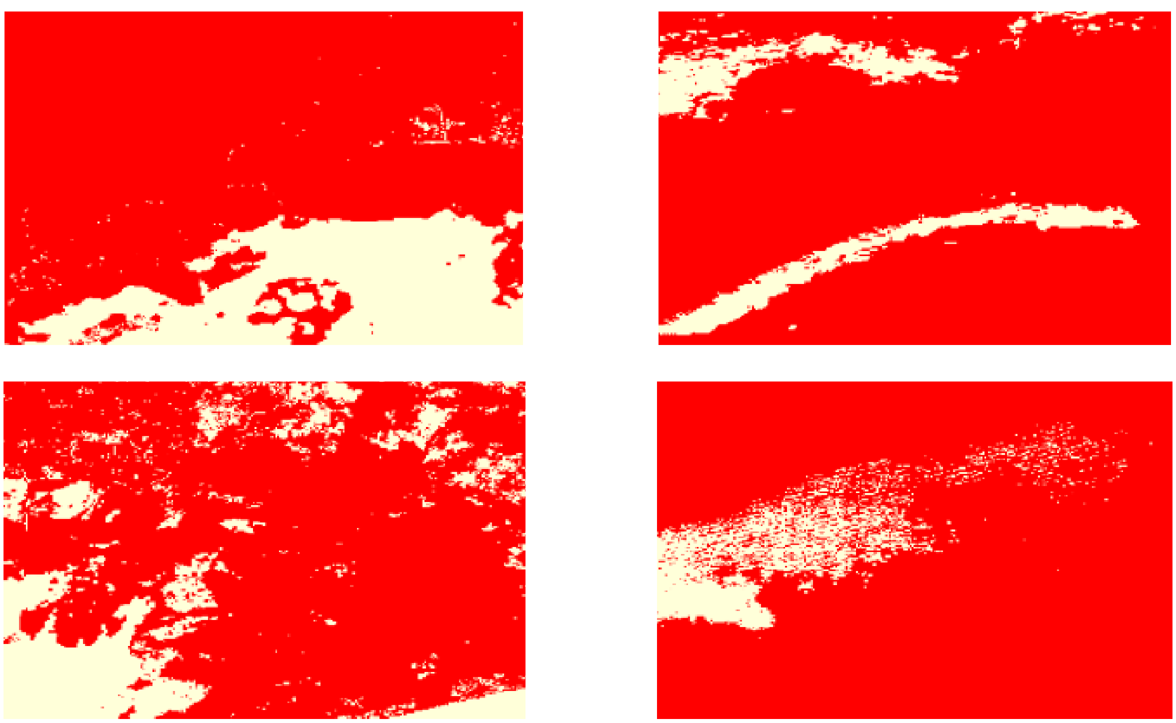
The overall scenario of error matrix through cross validation process and random sampling of pixels is shown as chart below:

Note (i): In above plotted graph, the input image speci_cations(taken as X- axis) are as below:-

1. L3.NE45B05.09mar13
2. L3.NE45B06.09mar13
3. L3.NE45B09.09mar13
4. L3.NE45B10.09mar13



Note (ii): In the graph, blue colour line represents - water class pixel predictions and orange colour line represents - No water class pixel predictions. Linear kernel classification and non-linear kernel classification is achieved. Further, SVM is fitted into whole data set then the fitted SVM is utilized for the prediction of decision classes for individual pixel in input image. Finally classifier detects the hyper-plane or support vector in n-dimensional space. So, here one class represents the region pixels associated to coastal land area and other pixels fall into ocean water region class. Finally the classified images are plotted as below:-



Execution time: We have performed the experiment on different system specification machines. The execution time of the experiment for overall process is presented as table below:-

Table I
Comparison Table for Execution Time

System Spec	*RAM Size*	*Execution Time*				*Average Execution Time(Sec)*
		1st image exp (Sec)	*2nd image exp (Sec)*	*3rd image exp (Sec)*	*4th image exp (Sec)*	
Core i3	4 GB	539.86	520.71	490.60	467.56	504.68
Core i5	4 GB	368.71	365.60	332.51	327.86	349.66
Core i7	8 GB	155.34	149.63	121.42	116.79	134.80

5. CONCLUSION

Today, expert systems and machine learning methods are getting much popularity in the classification area because of the effectiveness and accuracy of classification. In this paper, we have presented the Rough set based Support vector machine method for classification of images. In our proposed mechanism, Rough set (RS) is used as a feature selection mathematical tool, which eliminates the redundant features. Further, Support vector machine (SVM) is employed for image classification. Later, in this paper, experimental procedure and results for RS-SVM method are discussed.

REFERENCES

[1] C. I. Chang, Hyperspectral Data Exploitation: Theory and Applications. NewYork, NY, USA: Wiley, 2007.

[2] Li, D., Wang, J., Zhao, X., Liu, Y., Wang, D.(2014). Multiple kernel-based multi instance learning algorithm for image classification. Journal of Visual Communication and Image Representation, 25(5), 11121117.

[3] Pawlak, Z.(1982). Roughsets. International Journal of Computer and information Sciences, 11, 341-356.

[4] Petrosino, A., Salvi, G.(2006). Rough fuzzy set based scale space transforms and their use in image analysis. International Journal of Approximate Reasoning, 41, 212228.

[5] Cortes, C., Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273297.

[6] Abonyi, J., Szeifert, F. (2003). Supervised fuzzy clustering for the identification of fuzzy classifiers. Pattern Recognition Letters, 24(14), 21952207.

[7] Boser, B. E., Guyon, I. M., et. al., (1992). A training algorithm for optimal margin classifiers. In Fifth annual workshop on computational learning theory. Pittsburgh: ACM.

[8] Chang, C. C., Lin, C. J. (2001). LIBSVM: A library for support vector machines.

[9] Frohlich, H., Chapelle, O., et. al., (2003). Feature selection for support vector machines by means of genetic algorithms. In Proceedings of the 15th IEEE international conference on tools with artificial intelligence, Sacramento, CA, USA, pp. 142148.

[10] Goodman, D., Boggess, L., et. al., (2002). Artificial immune system classification of multiple-class problems. Intelligent Engineering Systems Through Artificial Neural Networks, Fuzzy Logic, Evolutionary Programming Complex Systems and Artificial Life, 12, 179184.

[11] Hsu, C. W., Chang, C. C., et. al., (2003). A practical guide to support vector classification. Technical report, Department of Computer Science and Information Engineering, National Taiwan University, Taipei.

[12] Joachims, T., Nedellec, C., et. al., (1998). Text categorization with support vector machines: Learning with many relevant. In Proceedings of the 10th European conference on machine learning, pp. 137142.

[13] John, G. H., Kohavi, R., et. al., (1994). Irrelevant features and the subset selection problem. In Proceedings of the 11th international conference on machine learning.

[14] Kryszkiewicz, M., Rybinski, H., (1996). Attribute reduction versus property reduction. In Proceedings of the fourth European congress on intelligent techniques and soft computing, pp. 204208.

[15] Osuna, E., Freund, R., Girosi, F. (1997). Training support vector machines: Application to face detection. In Proceedings of computer vision and pattern recognition, Puerto Rico, pp. 130136.

[16] Pawlak, Z. (1982). Rough sets. International Journal of Parallel Programming, 11(5), 341356.

[17] Pawlak, Z. (1996). Why rough sets. In Proceedings of the IEEE international conference on fuzzy system.

[18] Pawlak, Z. (1997). Rough set approach to knowledge-based decision support. European Journal of Operational Research, 99(1), 4857.

[19] Vapnik, V. (1995). The nature of statistical learning theory. New York: Springer.

[20] Vapnik, V. (1998). Statistical learning theory. New York: Wiley.

[21] Quinlan, J. R. (1996). Improved use of continuous attributes in C 4.5. Journal of Artificial Intelligence Research, 4(77-90), 325.

[22] Scholkopf, B., Smola, A. J. (2002). Learning with kernels: Support vector machines, regularization, optimization, and beyond. The MIT Press.

[23] C.-F. Lin and S.-D. Wang, Fuzzy support vector machines, IEEE Transactions on Neural Networks, Vol. 13, No. 2, March 2002.

[24] Ahmed, Shohel Ali, SnigdhaDey, and Kandarpa Kumar Sarma. Image Texture Classification Using Artificial Neural Network (ANN). Emerging Trends and Ap-plications in Computer Science (NCETACS), 2011 2nd National Conference on IEEE, 2011.

[25] Chen, Degang, Qiang He, and Xizhao Wang. FRSVMs: Fuzzy rough set based support vector machines." Fuzzy Sets and Systems 161.4 (2010): 596-607.

[26] Urszula stanczyk, Rough Set-Based Analysis of Characteristic Features for ANN Classifier, Hybrid Artificial Intelligence Systems, pp.565-572, DOI: 10.1007=978-3-642-13769-3 69, June 2010.

[27] D. N. Vasundhara and M. Seetha, Rough-Set and Artificial Neural Networks Based Image Classification, IC3I (14 - 17) Dec, 2016

[28] https://cran.r-project.org/web/packages/e1071/index.html

[29] https://cran.r-project.org/web/packages/raster/index.html

[30] https://cran.r-project.org/web/packages/sp/index.html

[31] https://cran.r-project.org/web/packages/RoughSets/index.html

[32] http://stat.ethz.ch/R-manual/R-devel/library/datasets/html/datasets-package.html

[33] https://cran.r-project.org/web/packages/RGraphics/index.html

[34] https://stat.ethz.ch/R-manual/R-devel/library/grDevices/html/grDevices-package.html

[35] https://cran.r-project.org/web/packages/R.methodsS3/index.html

[36] https://stat.ethz.ch/R-manual/R-devel/library/stats/html/stats-package.html

[37] http://bhuvan.nrsc.gov.in/data/download/index.php.