TRUS Image Classification for Prostate cancer using Computational Intelligence

R. Manavalan
Department of Computer Science and Applications,
K.S.R. College of Arts and Sciences,
Tiruchengode– 637 209, India
manavalan_r@rediffmail.com

K. Thangavel
Department of Computer Science
Periyar University
Salem-636 011, India
drktvelu@yahoo.com

Abstract

In medical image analysis image classification is key task for diagnosing the disease of the patient. The classification results assist the doctors to treat the patient according to the severances of the diseases. TRUS imaging is one of the most important medical imaging technologies for scanning the prostate to detect the dissimilarity in the images. It is hard to extort the region of interest from the original TRUS prostate image, since the problems with images are low of intensity contrast and inherent with speckle noise. The M3-Filter is applied to remove the speckle noise. Then, the region of interest is extorted using DBSCAN clustering with morphological operators. The twenty two features are extracted from Gray Level Co-occurrence Matrix (GLCM) which is constructed using extracted ROI. Further, QR-ACO feature selection algorithm is adapted to select the optimum features from the constructed features set. This paper proposes Complex-valued Support Vector Machine (C-SVM) for the classification of TRUS prostate cancer images. And also investigates proposed approaches with SVM according to prostate cancer based on the underlying texture contained within the region of interest. Receiver Operating Characteristic (ROC) analysis is used to evaluate the performance of the proposed classifiers. Experimental results demonstrate that the proposed approach gives the best performance compare to SVM.

Keywords: Prostate, TRUS, DBSCAN, M3-Filter, QR-ACO, SVM, C-SVM

1. Introduction

Prostate cancer is the second leading cause of death among men, accounting for nearly one out of every eleven cancers in the developing countries [1]. The early detection and diagnosis of prostate cancer is the key to decrease death rate and to provide treatment at the right time [2]. The early detection of prostate cancer is highly recommended, since it is only curable at an early stage. Nowadays, it is detected and diagnosed using the physical examination, imaging, and biopsy collectively.

Ultrasound imaging techniques is most prominent tool for prostate cancer diagnosis. The low quality and noise characteristics of the TRUS images make very much challenging task for the early and accurate detection of tumor [3]. The regions of interest (ROI) in the images are terribly hard to found. ROI is identified with the support of expert in radiology which is time consuming process and operator dependent. Depending on these results, the radiologist scrutinizes the TRUS image and a biopsy operation may be suggested [3].

The biopsy is one sort of invasive surgical operation and it affects the patients both psychologically and physically. To avoid unnecessary biopsy, many researchers have been explored Computer Aided Diagnosis (CAD) system which provides more objective evidences and stable high diagnostic rates. The objective of proposed system is to minimize the number of misclassifications and improve the diagnosis performance and accuracy by classifying normal from abnormal prostate cancer for given TRUS images.

The features extracted from the co-occurrence matrices of the ultrasound images are more valuable to improve the ability of ultrasound to distinguish benign from malignant, not the fractal analysis [4]. Each image is usually represented in terms of feature vectors, which is obtained using feature extraction algorithms.

The hybrid approach quick reduct -ant colony optimization (QR-ACO) feature selection is used to discover the best optimal feature set from the constructed feature set. Since constructed features set may contain correlated and irrelevant information, which may steer the dimensionality curse problem and may decline the accuracy of the classification algorithms [5].

ISSN: 2349 - 6363

In this paper, Complex-valued Support Vector Machine (C-SVM) is proposed for the task of classification. The performance of C-SVM is thoroughly studied in contrast with SVM for prostate cancer classification. 10-fold cross validation is used to measure the classification evaluation on the selected feature sets. Finally, the performance of each classifier is also evaluated using Receiver Operating Characteristic (ROC) analysis. An overview of the proposed system is presented below.

1.1. Overview of the System

The proposed system consists of five stages such as acquisition of TRUS image of prostate, preprocessing and segmentation, feature extraction, feature selection and classification. It is developed for automatic detection of prostate tumor in Transrectal Ultrasound (TRUS) image. This paper emphasis feature selection and classification. A block diagram of the proposed systems is given in Fig 1.

The rest of the paper is organized as follows: section 2 deals about the problem definition for classification. Image preprocessing of original TRUS medical image of prostate and the DBSACN clustering with morphological operators for locating the Region of Interest (ROI) from preprocessed image are presented in section 3. The feature extraction through GLCM and QR-ACO feature selection method for minimizing features set are described in section 4. The proposed classification procedure C-SVM is detailed in section 5. The experimental analysis and discussion are presented in the section 6. Conclusion and research directions are provided in section 7.7.

2. Problem Definitions

The problem describe in this paper is that the TRUS image classification for diagnosing prostate tumor. Choosing an appropriate model for TRUS image classification is difficult task that model has to achieve better classification accuracy. Formally, it can be defined as:

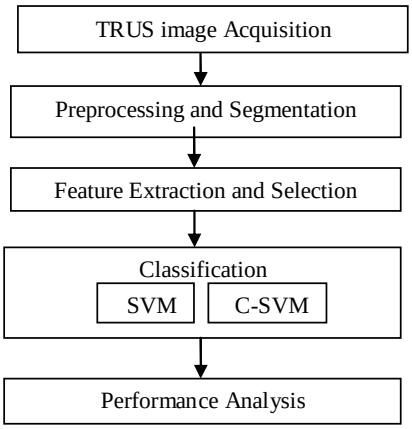

Fig. 1 Block Diagram of Proposed System

Objects are to be classified as belonging to one of a number of predefined classes {1, 2, … , K}. Each object associated with a class label Y = {1, 2,…, K} and a feature vector of p measurements: $X = (X_1, X_2 \ldots, X_p)$. The aim is to predict Y from X. A classifier partitions the features space into K disjoint subsets, $A_1, \ldots, A_K$, such that for a sample with feature data set $X = (X_1, X_2, \ldots, X_p)$ in A_k the predicted class is k Classifiers are from a training set. $L = (X_1, Y_1), \ldots, (X_n, Y_n)$. Classifier C built from a learning set L: C (. ; L) : X →{1, 2, …, K}. Predicted class for observation X: C(X; L) = k if X is in A_k..

3. Image Preprocessing and Segmentation

Ultrasound imaging is the best approach for prostate cancer diagnosis and prognosis. The accurate region of interest is impossible to produce in manual segmentation, since it is slow and heavily user dependent. To overcome this, automatic segmentation is developed which has a significant advantage over manual segmentation [6]. However, the problems with ultrasound are low intensity contrast and inherent speckle noise, which makes difficult in automatic segmentation of ultrasound images. The accurate segmentation of TRUS prostate images plays a key role in classification. The M3-Filter is applied to acquire despeckled image that preserve proteomic while removing unwanted noise [7]. Once noise is removed the top-hot filter is applied to form require edges. Since the edge information is needed for the proper segmentation. Then, ROI is extorted by using DBSCAN clustering with morphological operators from the thresholded image, which is obtained by using Local Adaptive thresholding method to reduce the complexity of segmentation process in large extent [8]. The sample output of step by step process of ROI extraction is presented in Fig 2.

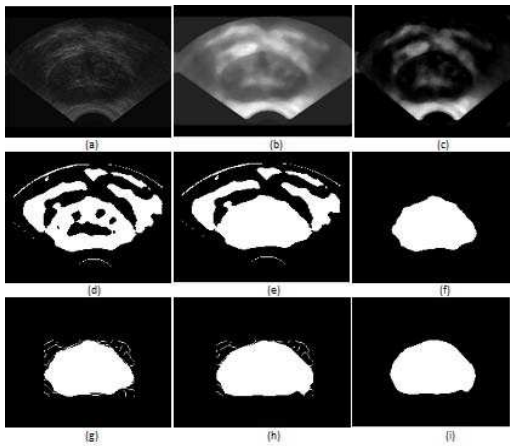

Fig. 2 (a) Original image, (b) Smoothed image using M3-Filter, (c) contrast enhanced image (d) result of thresholding (e) filled version of the thresholded image,(f) isolated object corresponding to prostate (g) result of applying canny edge detection on the enhanced image(h) result of DBSCAN Clustering that shows prostate pixels, border pixels and noise pixels, (i) resultant image after the noise pixel removed.

4. Feature Extraction and Selection

One of the significant characteristics of an image is Texture which is used to describe the local spatial variations in image brightness and also related to image properties such as coarseness, and regularity. Statistical texture features proved its high recognition ability in ultrasound images.

The features extracted from Gray Level Co-occurrence Matrix (GLCM) used for prostate cancer identification as well as some other applications [9]. In this paper, Region of Interest (ROI) is utilized to construct feature sets using GLCM [10]. The detailed description of GLCM method is described here under.

4.1. GLCM texture features

Texture information is characterized by the spatial arrangement of the pixel intensities. This can be specified by a 2-Dimensional spatial dependence matrix known as the Gray Level Co-occurrence Matrix (GLCM) [10]. It is a Statistical method uses second order statistics to model the relationships between pixels within the region by constructing Gray Level Co-occurrence Matrices. It is computed based on an estimation of the second-order joint conditional probability density function $p(i, j \mid d, \theta)$ for various directions $\theta = 0°$, 45°, 90°, 135° etc., and different distances, d = 1, 2, 3, 4 and 5. The function $p(i, j \mid d, \theta)$ is the probability that two pixels, which are located with an inter-sample distance d and a direction θ, have a gray level and a gray level i and j. The spatial relationship is defined in terms of distance d and direction θ.

The Gray Level Co-occurrence Matrix (GLCM) is generated for each ROI of the TRUS prostate image to extract the twenty two features as feature set by considering the direction $\theta = 45$ and distance $d = 2$ since which is more suitable for discriminating the prostate cancer from normal [10]. List of extracted features from the region of interest are shown in table 1.

Table 1: List of extracted Features

Feature Name	Description
f_1	Energy(ENE)
f_2	Entropy(ENT)
f_3	Homogeneity(IDM)
f_4	Inertia(CON)
f_5	Correlation(COR)
f_6	Variance(VAR)
f_7	Shade(SHA)
f_8	Prominence(PRO)
f_9	Sum Average(SA)
f_{10}	Sum Entropy(SE)
f_{11}	Sum Variance(SV)
f_{12}	Difference Average(DA)
f_{13}	Difference Entropy(DE)
f_{14}	Difference Variance(DV)
f_{15}	Information Measure(IMC1)
f_{16}	coefficient of variation(COV)
f_{17}	Peak transition probability(MAX)
f_{18}	diagonal variance(DIAV)
f_{19}	diagonal moment(DIAM)
f_{20}	second diagonal moment(DSM)
f_{21}	triangular symmetry(TRS)
f_{22}	Information Measure(IMC2)

4.2. Feature Selection

The accuracy, speed and interpretation are the most important objectives of image processing and analysis and also machine learning community. All the features in the constructed feature set are not valuable for the image classification systems. Since some of the features may act as curious and will decrease the success rate of the classification system. Therefore, the feature selection algorithm plays vital role in the success of classification system. Feature selection is process of selecting key features for classification and the same time attempting to ignore the (possibly misleading) contribution of irrelevant features for maximizes the accuracy and also simplifies the procedure of classification task, reducing overall complexity [11]. It is also called as feature subset selection, variable selection, or attributes reduction.

In this paper, the QR-ACO feature selection algorithm used for selecting optimum features set from the constructed features set in order to discriminate prostate cancer from normal with prominent classification accuracy. The detailed description of QR-ACO for feature selection is discussed as follows.

4.2.1. Quick Reduct-Ant Colony Optimization Algorithm

Ant Colony optimization (ACO) algorithm is mainly pretty for feature selection as there seems to be no heuristic that can guide search to the optimal minimal subset every time [12]. Additionally, it can be the case that ants discover the best feature combinations as they proceed throughout the search space. However, Rough sets theory is effective method for deals with uncertainty and vagueness of an information system which can reduce decision-making and classification rules so as to establish knowledge model through data analysis and knowledge reduction under the condition of maintaining the ability of classification unchangeable. It requires no additional knowledge except for the supplied data and performs feature selection using only the granularity structure of the data.

QR- ACO: Algorithm
Input: Constructed Features Set
Output: Optimal Features Set

Step1: Take the input as a decision table $S = (U, C, D)$

Step2: Let $Core = \varphi$ and Calculate the $POS_c(D)$

Step3: For $\forall_a C$, calculate $POS_{(c-\{a\})}(D)$. If $POS_{(c-\{a\})}(D) \neq POS_c(D)$ then $CORE = CORE \cup \{a\}$;
Else $C = C - \{a\}$.

Step4: Execute iteratively step 2 until all attributes among C are calculated.

Step5: If $POS_{core}(D) = POS_c(D)$ algorithm stops and return $CORE$ as the result of feature selection; otherwise go to step 6

Step6: The pheromone of each arc (i, j) is assigned to an constant, i.e. $\tau_{ij}(0) = c$

Step7: Some ants (assumed the number of ants is m) are distributed to each core attribute node to conduct feature selection

Step8: Each ant selects next feature node

Step9: Calculate $POS_{core}(D), a \in C - CORE$, if $POS_{core}(D) = POS_c(D)$ algorithm stops and return $FS = CORE \bigcup a_i$ as the result of feature selection; else go to step 10

Step10: Update value of pheromone τ_{ij} for each path link and go to step 8

Fig. 3 ACO with Rough set based QR Feature Selection (QR-ACO)

ACO algorithm for feature selection gets trapped into stagnation situation. The number of ants specified in the algorithm generates continuously the same solution after a certain number of iterations, since pheromone amount intensifies at some points and the difference between pheromone concentrations on paths become very huge. For this reasons, the ACO algorithm stops to generate alternate solutions. The chance of stagnation is increased proportionally when the problem size is increased. The problems of stagnation, premature convergence and the slow convergence speed are exit in ACO algorithm. To overcome these problems in ACO based feature subset selection algorithm, Rough Set Theory based Quick Reduct is hybridized with ACO, and it is named as QR-ACO.

The main advantage of QR-ACO is able to characterize the granulation structure of a rough set using a granulation order. Based on the positive approximation, feature selection process is accelerated. It uses the dependency measure as the stopping criterion. This means that an ant will stop building its features subset when the dependency of the subset reaches the maximum for the dataset. QR-ACO is the best algorithm that found the best features set for prostate cancer classification [31]. It has the ability of avoiding the stagnation situation and improves its solutions with the time. Thus, it provides trustworthy performance, and is also not application specific. The Fig 3 shows QR-ACO feature selection algorithm.

The number of input is decided by automatic selection of QR-ACO feature selection algorithms. Further, the reduced dataset is considered as input for the proposed classification methods. The proposed classifiers are discussed subsequent section clearly.

5. Classification

The classification of the image is a significant area of research in various fields including pattern recognition, artificial intelligence medicine and vision analysis. The term image classification refers to the labeling of images into one of a number of predefined categories. Classification is defined as the process of discovering a model that describes and distinguishes the data classes. The trained model can be used to predict the classes of future data instances for which the class label is unknown, and this is often referred to as the predictive task. This paper proposed C-SVM classification technique and compared with SVM. The detailed description of SVM and proposed method is given under here.

5.1 Support Vector Machine Classifier

Support vector machine is based on statistical learning technique which is well-founded in modern statistical learning theory [13]. The Support Vector Machines were introduced by Vladimir Vapnik and his colleagues [16]. SVM is a useful technique for data classification [14]. A classification task usually involves with training and testing data which consist of some data instances. Each instance in the training set contains one target values and several attributes. The goal of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the attributes. Classification in SVM is an example of Supervised Learning. Known labels help indicate whether the system is performing in a right way or not. This information points to a desired response, validating the accuracy of the system, or be used to help the system learn to act correctly. A step in SVM classification involves identification as which are intimately connected to the known classes [15].

Support vector machines use the training data to crate the optimal separating hyperplane between the classes. The optimal hyperplane maximizes the margin of the closest data points. A good separation is achieved by the hyperplane that

has the largest distance to the nearest training features of any class (so-called functional margin). Maximum-margin hyperplane and margins for a SVM trained with samples from two classes. Samples on the margin are called the support vectors [17]. SVM divides the given data into decision surface. Decision surface divides the data into two classes like a hyper plane. Training points are the supporting vector which defines the hyper plane. The basic theme of SVM is to maximize the margins between two classes of the hyper plane (Steve, 1998).

Basically, SVMs can only solve binary classification problems. They have then been extended to handle multi-class problems. The idea is to decompose the problem into many binary-class problems and then combine them to obtain the prediction. To allow for multi-class classification, SVM uses the one-against-one technique by fitting all binary sub classifiers and finding the correct class by a voting mechanism [18]. If K is the number of classes, then K(K - 1)/2 binary classifiers are constructed and trained to separate each pair of classes against each other, and uses a majority voting scheme (max-win strategy) to determine the output prediction. For training data from the i th and the j th classes, we solve the following two-class classification problem [19]. Let us denote the function associated with the SVM model of $\{c_i, c_j\}$ as:

$$g(x)_{ij} = sign(f(x)_{ij})$$

An unseen example, x, is then classified as:

$$f(x) = \arg\max_i \sum_{i=1}^{K} \sum_{j=1 \wedge i \neq j}^{K} V_{ij}(x)$$

where:

$$V_{i,j}(x) = \begin{cases} 1 if \ \ g_{ij}(x) = 1 \\ 0 if \ g_{ij}(x) = -1 \end{cases}$$

Each feature set is examined using the Support Vector Machine classifier. In classification, we use a voting strategy, in which each binary classification is considered to be a voting where votes can be cast for all data points x in the end point is designated to be in a class with the maximum number of votes. In case that two classes have identical votes, though it may not be a good strategy, now we simply choose the class appearing first in the array of storing class names.

The objective of any machine capable of learning is to achieve good generalization performance, given a finite amount of training data, by striking a balance between the goodness of fit attained on a given training dataset and the ability of the machine to achieve error-free recognition on other datasets. With this concept as the basis, support vector machines have proved to achieve good generalization performance with no prior knowledge of the data. The optimal separating hyperplane can be determined without any computations in the higher dimensional feature space by using kernel functions in the input space. Commonly used kernels include:

Linear Kernel: $K(x, y) = x.y$

Radial Basis Function (Gaussian) Kernel: $K(x, y) = e^{(-||x - y||^2 / 2\sigma^2)}$

Polynomial Kernel: $K(x, y) = (x.y + 1)^d$

5.2 Complex Valued Support Vector Machine

Normally real part of the number alone used for image processing and analysis. The imaginary part is also plays significant role in medical image analysis. For each input vector built integral part and imaginary parts and then Construct complex data from real and imaginary components c = a + ib. The output is the same size as the inputs, which must be scalars or equally sized vectors, matrices, or multi-dimensional arrays. Once feature vectors is transformed into complex value then give as input to multi-class support vector machine which is clearly described in previous section.

6. Parameters for Evaluation

The algorithms discussed in previous section have been implemented using MATLAB. The entire input features are normalized into the range of [0, 1], whereas the output class is assigned to one for the highest probability and zero for the lowest. Performance evaluation of various classification models are analyzed and discussed based on the following criteria: (i) Sensitivity, (ii) Specificity (iii) Accuracy and (iv) Receiver Operating Characteristic (ROC) curve. The statistical parameters for classification with formula are given in Table 2.

Table 2: Evaluation Measures

S. No	parameters
1	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
2	$Specificity = \frac{TN}{TN + FP}$
3	$Sensitivity = \frac{TP}{TP + FN}$

Sensitivity and Specificity are the two most important characteristics of a medical test. Sensitivity is the probability that the test procedure declares an affected individual affected (probability of a true positive). Specificity is the probability that the test procedure declares an unaffected individual unaffected (probability of a true negative). Accuracy measures the quality of the classification. It takes into account true and false positives and negatives. Accuracy is generally regarded with balanced measure whereas sensitivity deals with only positive cases and specificity deals with only negative cases. TP is number of true positives, FP is number of false positives, TN is number of true negatives and FN is number of false negatives. A confusion matrix provides information about actual and predicted cases produced by classification system. The performance of the system is examined by

demonstrating correct and incorrect patterns. They are defined as confusion matrix in Table 3.

Table 3.Confusion Matrix

Actual	Predicted	
	Positive	*Negative*
Positive	TP	FP
Negative	FN	TN

The higher value of both sensitivity and specificity shows better performance of the system. The constructed features set are given to QR-ACO features selection algorithm. The table 4 shows the list of feature selected.

Table 4. Selected Features List

Algorithms	**Selected Features**
QR-ACO	f_2 ,f_3, f_4 ,f_{10}, f_{13} ,f_{14}

Table 5 Tenfold cross validation for testing classifier

Methods	**SVM**		
Fold	**Sensitivity**	**Specificity**	**Accuracy**
1	0.8559	0.9882	0.8764
2	0.8366	1.0000	0.8618
3	0.9011	1.0000	0.9164
4	0.8645	0.9765	0.8818
5	0.8409	0.9882	0.8636
6	0.9161	1.0000	0.9291
7	0.9441	0.9765	0.9491
8	0.8430	1.0000	0.8673
9	0.8860	1.0000	0.9036
10	0.8994	0.9882	0.8955

Table 6 Tenfold cross validation for testing classifier

Methods	**CSVM**		
Fold	**Sensitivity**	**Specificity**	**Accuracy**
1	0.8817	1.0000	0.9000
2	0.8409	1.0000	0.8655
3	0.9097	1.0000	0.9236
4	0.8860	0.9882	0.9018
5	0.8951	0.9882	0.9094
6	0.9183	1.0000	0.9309
7	0.9656	0.9882	0.9691
8	0.8581	1.0000	0.8800
9	0.9054	1.0000	0.9182
10	0.9699	1.0000	0.9727

Table 7 Computational Results

Methods	**Sensitivity**	**Specificity**	**Accuracy**
SVM	0.87876	0.99176	0.89446
C-SVM	0.90307	0.99646	0.91712

The 10-fold cross-validation is the standard way of measuring the performance of the classifier on a particular dataset. The data is divided by chance into 10 parts. During each run, one of partitions is chosen for testing, while the remaining nine-tenths are used for training. Again, the procedure is repeated 10 times so that each partition is used for training exactly once. Classifier performance is also evaluated by calculating the number of correctly classified instances to total number of instances (Accuracy).

In our experiment, the total 5500 instances are divided into 10 disjoint sets with 550 cases in each. For each experiment, 9 of these sets are used as training data, while the 10th is reserved for testing. The experiment is repeated 10 times so that every case appears once as part of a test set. The constructed classification algorithms, such as SVM and CSVM are evaluated with tenfold cross-validation using selected features set and their classification results are recorded in the table 5 and table 6. The individual performance measures of ten folds for the reduced feature sets by the classification algorithms are exposed in Fig 4, Fig 5, and Fig 6. The average classification measures of each classifier are provided in table 7.

The relative performance measures for the proposed methods are demonstrated in Fig. 7. The classification results in Table 5 and Table 6 for the ten folds clearly show that the different classification models discriminate malignant and benign with different accuracy. And also confirm that the classification accuracy achieves using C-SVM is much better than SVM.

From the table 7, we observed that the C-SVM and SVM classifier achieve different classification accuracies. The accuracy of 88 % is arrived by the SVM, while 87% sensitivity and 99% specificity. The C-SVM classifier achieved an accuracy of 92% where 87% sensitivity and 99% specificity. The accuracy of C-SVM Classifier is 2% higher than SVM classifier.

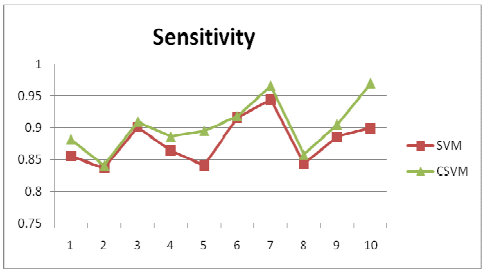

Fig. 4 Performance of Sensitivity

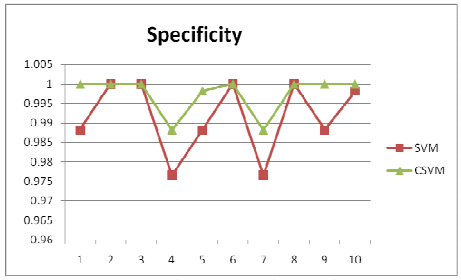

Fig. 5 Performance of Specificity

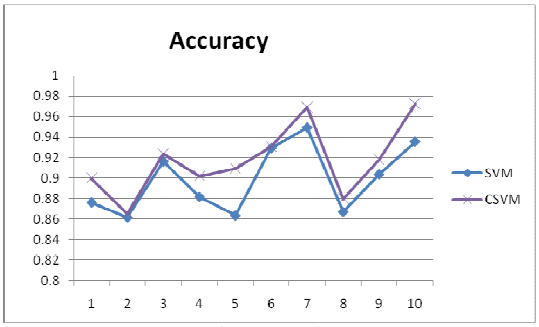

Fig. 6 Performance of Accuracy

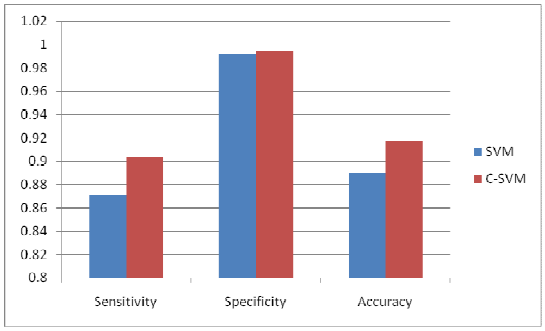

Fig. 7 Relative performance measures

ROC (Receiver operating characteristic) analysis is also used to evaluate the performance of the classifiers with respect to the prostate cancer classification task. It is a based on statistical decision theory, developed in the context of electronic signal detection, and has been applied extensively to diagnostic system in clinical medicine. ROC curve is a plot of sensitivity against 1-specificity at differing thresholds encountered during the classifier testing. The 1- Specificity is the probability of incorrectly classifying a normal as tumor. Similarly, sensitivity is the probability of correctly classifying a normal as normal and tumor as tumor. To make ROC graph, the X-axis is 1- Specificity and the Y-axis is the sensitivity. ROC curve is generated for the results of classifiers and exposed in the Fig. 8.

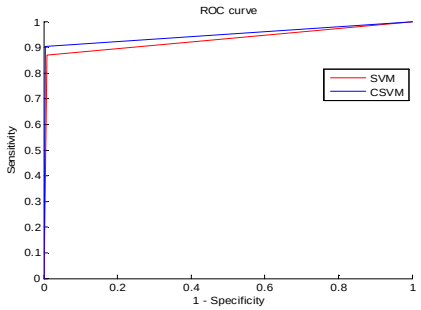

Fig. 8 ROC Curves for SVM, CSVM Classifier

Area under the ROC curve is another important criterion for classifier evaluation. The area under the ROC curve represents the probability of a random positive sample to receive a better score than a random negative sample. The value of AUC ranges from 0.5 to 1.0 that indicates chance to perfect discrimination. The diagnostic test is more accurate when area is larger. Commonly used classification using AUC for a diagnostic test is summarized in table 8. The computed value of ACU for each classifier is recorded in Table 9.

Table 8 Range of AUC for a diagnostic test

AUC Range	Classification
0.9 < AUC < 1.0	Excellent
0.8 < AUC < 0.9	Good
0.7 < AUC < 0.8	Worthless
0.6 < AUC < 0.7	Not good

Table 9 Performance of Classification Algorithms

Algorithms	Az Value
SVM	0.931259
C-SVM	0.948779

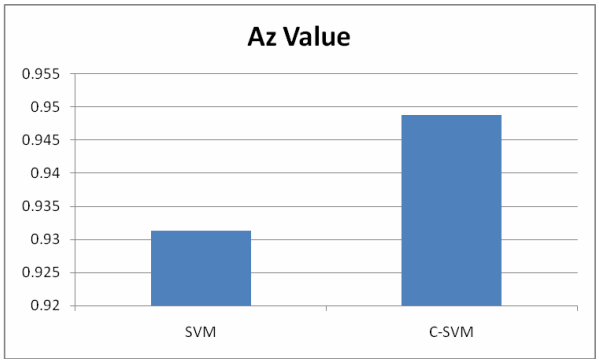

Fig. 9 Area (AZ) under ROC curves

From the table 9, the best area under curve value is 0.948779 for CSVM and SVM is 0.931259 respectively. Fig. 9 represents the areas under ROC curves for the proposed classifiers.

7. Conclusion

Classification accuracy is more important in the field of medical diagnosis using images. The prostate region from the TRUS images is extracted by DBSCAN Clustering with morphological operations. From the identified region of interest twenty two features are extracted. Then the relevant features are selected for classification using QR-ACO feature selection technique. In this paper, the two classifiers have been constructed, namely SVM, C-SVM and investigate for the task performance of prostate cancer classification using TRUS images. The classification accuracy of C-SVM is higher compare to SVM. The experimental result reveals that the C-SVM achieves better classification accuracy which is 2% higher than SVM.

References

[1] Pingkun Yan, Sheng Xu, Baris Turkbey, and Jochen Kruecker. "Adaptively Learning Local Shape Statistics for Prostate Segmentation in Ultrasound", IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 58, NO. 3, pp. 633 -641, MARCH 2011.

[2] Christopher M. Coley, Michael J. Barry, Craig Fleming, John H. Wasson, Marianne C. Fahs, Joseph E. Oesterling, "Should medicare provide reimbursement for prostate-specific antigen testing for early detection of prostate cancer? Part II: Early detection strategies", Urology, Volume 46, Issue 2, pp 125-141, ISSN 0090-4295, 10.1016/S0090-4295(99)80181-2, August 1995.

[3] Mohamed, S., Youssef, A., El-Saadany, E. and Salama, M. "Primary Title: Artificial Life Feature Selection Techniques for Prostrate Cancer Diagnosis Using TRUS Images", Image Analysis and Recognition, Lecture Notes in Computer Science., Springer Berlin / Heidelberg, pp. 903 - 913, Vol. 3656, Copyright: 2005.

[4] Dar-Ren Chen.T, Ruey-Feng Chang, Chii-Jen Chen, Ming-Feng Ho, Shou-Jen Kuo, Shou-Tung Chen, Shin-Jer Hung, and Woo Kyung Moon. "Classification of breast ultrasound images using fractal feature", Journal of Clinical Imaging vol. 29, pp. 235–245, 2005.

[5] Bing-Yu Sun, Xiao-Ming Zhang, Jiuyong Li, and Xue-Min Mao. "Feature Fusion Using Locally Linear Embedding for Classification", IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 21, NO. 1, pp 163 -168, January 2010.

[6] Yanong Zhu, Stuart Williams, Reyer Zwiggelaar, "Computer technology in detection and staging of prostate carcinoma: A review, Medical Image Analysis", Volume 10, Issue 2, Pages 178-199, April 2006.

[7] Thangavel K, Manavalan R and Laurence Aroquiaraj I. "Removal of Speckle Noise from Ultrasound Medical Image based on Special Filters: Comparative Study", ICGST-GVIP Journal, ISSN 1687-398X, Volume (9), Issue (III), June 2009, PP 25 – 32.

[8] Manavalan R, and Thangavel K., "TRUS Image Segmentation Using Morphological Operators and DBSCAN Clustering", 2011 World Congress on Information and Communication Technologies, 978-1-4673-0125-1© 2011 IEEE, PP. 902 – 907, 2011.

[9] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural Features for Image Classification, IEEE Transactions on Systems, Man, and Cybernetics, vol. 3, no. 6, pp. 610–621, 1973.

[10] Manavalan R, and Thangavel K. "Evaluation of Textural Feature Extraction Methods for Prostate Cancer TRUS Medical images", International Journal of Computer Applications (0975 – 8887) Volume 36– No.12, pp.33 – 39, December 2011.

[11] R. Duda, P. Hart, D. Stork, "Pattern Classification", John Wiley and Sons.2001.

[12] Debahuti Mishra, Dr. Amiya Kumar Rath, Dr. Milu Acharya and Tanushree Jena. "Rough ACO: A Hybridized Model for Feature Selection in Gene Expression data" Int. J. of Computer Communication and Technology, Vol. 1, No. 1, pp. 85 – 98, 2009.

[13] Devendran V et. al., "Texture based Scene Categorization using Artificial Neural Networks and Support Vector Machines: A Comparative Study," ICGST-GVIP, Vol. 8, Issue IV, pp. 45-52, December 2008.

[14] Vapnik V.,The Nature of Statistical Learning Theory, chapter 5. Springer-Verlag, New York. 1995.

[15] Vapnik.V, Statistical Learning Theory. John Wiley & Sons, New York, 1998.

[16] Vapnik V N, "Estimation of Dependencies Based on Empirical Data". Nauka, Birmingham, AL. 1979.

[17] Fatima Eddaoudi, Fakhita Regragui, "Microcalcifications Detection in Mammographic Image Using Texture Coding", Applied Mathematical Sciences, Vol.5, no. 8, 381-393, 2011.

[18] Knerr, S., Personnaz, L., and Dreyfus, G., "Single-layer learning revisited: A stepwise procedure for building and training neural network". Neurocomputing: Algorithms, Architectures and Applications, NATO ASI, Berlin: Springer-Verlag., 1990.

[19] Metehan Makinacı, "Support Vector Machine Approach for Classification of Cancerous Prostate Regions", World Academy of Science, Engineering and Technology, pp 166 – 169, 2005.

Mr. MANAVALAN R Obtained M.Sc., Computer Science from St.Joseph's College of Bharathidasan University,Trichy, Tamilnadu, India, in the year 1999, and M.Phil., in Computer Science from Manonmaniam Sundaranar University, Thirunelveli, Tamilnadu, India in the year 2002. He works as Asst.Prof & Head, Department of Computer Science and Applications, KSR College of Arts and Science, Thiruchengode, Nammakal, Tamilnadu, India. He pursues Ph.D in Medical Image Processing. His areas of interest are Medical image processing and analysis, soft computing, pattern recognition and Theory of Computation.

D**r. K. THANGAVEL** received the Master of Science from Department of Mathematics, Bharathidasan University in 1986, and Master of Computer Applications Degree from Madurai Kamaraj University, India in 2001. He obtained his Ph. D. Degree from the Department of Mathematics, Gandhigram Rural University in 1999. Currently he is working as Professor and Head, Department of Computer Science, Periyar University, Salem. His areas of interests include medical image processing, artificial intelligence, neural network, fuzzy logic, data mining, pattern recognition and mobile computing.