SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM

JBEMi JOURNAL OF BIOMEDICAL ENGINEERING AND MEDICAL IMAGING

ISSN: 2055-1266
VOLUME 3 ISSUE 1

Detection and Classification of Focal Liver Lesions using Support Vector Machine Classifiers

[1]Deepti Mittal, [2]Anju Rani and [3]Ritambhara

Department of Electrical and Instrumentation Engineering, Thapar University, Patiala, Punjab, India.

[1]deeptimit@gmail.com; [2]anju.yaduvanshi@yahoo.com; [3]ritambhara3939@gmail.com

ABSTRACT

In the present work, two computer aided diagnostic systems are designed to detect and classify focal liver lesions such as Cyst, Hemangioma, Hepatocellular carcinoma and Metastases. The work evaluates clinically acquired ultrasound image database. Database contains 111 liver images comprising 95 images of focal lesions and 16 images of normal liver. Images are enhanced and manually segmented into 800 non-overlapping segmented-regions-of-interest. Afterwards, 208 textual features are extracted from each segmented-regions-of-interest. First diagnostic system is designed with one-against-one multiclass support vector machine classification approach showing 93.1% (512/550) overall accuracy on test dataset. Second system is designed with tree structured approach using four binary support vector machine classifiers showing 86.9% (478/550) overall accuracy on test dataset. Out of these two, one-against-one approach based diagnostic system outperforms the neural network based diagnostic system designed for the same purpose by providing 96.6% classification accuracy for typical cases and 85.3% for atypical cases.

Keywords: Ultrasound; Focal liver lesions; Feature extraction; Classification; Support vector machine classifier.

1 Introduction

Focal liver diseases constitute an important public health issue having very high incidence in Asian countries. Among a variety of focal liver diseases, liver cancer is one of the severe liver diseases. Liver cancer is the sixth most common focal malignant tumor in the world and the third most common cause of cancer-related deaths worldwide [1]. Imaging is an effective tool for early detection of liver abnormalities because in many cases it can detect even before they are palpable. Worldwide, ultrasound is always the first preference among various imaging modalities for screening focal liver lesions. Focal liver lesions are mostly found incidentally during ultrasound examinations, especially when it is performed as a part of the follow-up of patients with tumor, abdominal ache or cancer staging. Further, focal liver lesions with their atypical appearances may lead to misdiagnosis and confusion with other lesions even for the expert radiologists.

Present study is performed to design computer aided diagnostic (CAD) system to detect and classify focal liver lesions such as Cyst, Hemangioma (HEM), Hepatocellular carcinoma (HCC) and Metastasis

DOI: 10.14738/jbemi.31.1821
Publication Date: 08th February 2016
URL: http://dx.doi.org/10.14738/jbemi. 31.1821

(MET) using support vector machine (SVM) classifiers. Review of literature reveals that neural network (NN) had been used as a classifier in several studies to design CAD system for detection and classification of focal liver lesions [2-5]. Further, a NN based diagnostic system to classify among five liver tissue categories, viz., four above mentioned focal lesion tissue categories and normal (NOR) liver tissue category, has presented by Mittal *et. al.* [12] showing 86.5% test accuracy. It has also revealed by literature review that SVM classifier had not been used to develop the CAD systems for these focal liver tissue categories in spite of the fact that SVM classifier had been preferred in many other clinical applications, such as breast cancer detection, thyroid nodule detection, liver fibrosis identification, cervical lymph nodes, liver cancer detection using CT images, retinal exudates detection showing good test accuracies [6-11]. Another reason to opt SVM classifier is that the generalized performance of NN based diagnostic system was not found as excellent as its performance on the training and validation sets [12]. The reason for such performance of NN based diagnostic system may be the curse of dimensionality. The curse of dimensionality arises when number of features and that of available training samples are unbalanced. For a classification problem with d features, it is estimated that a minimum of $10d$ samples per training class is necessary for reliable statistical estimates [13]. This requirement is usually not feasible due to constraints of data collection, such as dependency of data collection on frequency of occurrence of patients having specific diseases and so on. Therefore, powerful learner such as NN having small number of training samples with 208 features may cause overfitting especially when performing intensive search for the best system model. On the other side, SVM has the potential to produce high classification accuracy with limited number of training samples having hundreds of features. It is a good candidate for image classification showing its lower sensitivity to the curse of dimensionality. Therefore in the present study SVM classifiers are used in diagnostic system designing.

Furthermore, SVM formulation is originally designed for two class problems therefore it performs well in binary classification. Yet, most of the real life diagnostic problems are not binary. There are some feasible strategies that can extend the formulation of SVM to solve a multi-class classification problem. To solve the multiclass problem, many multiclass SVM methods have already been developed, differing from each other in the definitions of the binary SVMs and in the combining strategy of the binary SVMs [14-25]. These methods can be divided into two types. One, called single machine approach, modifies the binary class objective function and allows simultaneous computation of multiclass classification by directly considering all classes to solve a single optimization problem [16, 20], whereas the other is to construct and combine several binary classifiers. The latter strategy based on the combination of many binary classifiers has been proved to be more efficient than the single machine approach. There are a good number of approaches, which consider a multi-class classification strategy from the combination of many binary SVM classifiers, such as one-against-all (OAA) [14], one-against-one (OAO) [15] and all-and-one (AAO) [22], direct acyclic graph SVM (DAGSVM) [18], the hierarchical tree structure based methods [21, 23, 25], and error correcting output codes (ECOC) methods [17, 19, 24]. Rifkin and Klautau [26] did many carefully controlled experiments and proposed that a simple scheme such as OAA and OAO is preferable to a more complex ECOC methods or single machine schemes, once the best binary classifier available is used. Hsu and Lin [27] suggested that OAO and DAGSVM may be more suitable for practical use after comparing the single machine approaches with OAA, OAO, and DAGSVM. OAO and DAGSVM perform very similar, which is due to the fact that both multiclass SVM methods use the same type of

binary SVM classifiers. DAGSVM and OAO have the fastest training time. The OAO can use in training many more binary SVMs than the OAA, even then OAO may require much lesser training time than the OAA. Moreover, OAO can achieve comparable or even lesser test computational complexity compared with OAA due to the large number of shared support vectors between binary SVMs of OAO [28]. These detailed experiments and comparative studies have suggested that OAO is useful and more practical among the mentioned approaches for real application in term of accuracy and computational cost. As a result, OAO is used as a SVM multiclass classification method in this work.

The OAO multiclass classification approach has to solve too many binary decision functions when the multi-class problem has a large number of classes. Also too many binary tests are required to be conducted before a final classification is made. The hierarchical tree-based SVM methods, which combine support vector machine and binary tree architecture, have the characteristics, such as lower number of binary classifiers required to be trained and faster decision speed. These advantages of tree based approach become more and more prominent as the number of the class increases. In multiclass classification problem of N classes, the OAO approach requires $N(N-1)/2$ binary classifiers to be trained, whereas tree structured SVM approach requires only $(N-1)$ binary classifiers to be trained. Therefore, finally one-against-one and hierarchical tree structured based SVM approaches are selected to implement the multiclass classification problem.

Therefore, two CAD systems are designed using two different multiclass SVM classification approach, one is one-against-one (OAO) approach and other is tree structured approach. In the end, both systems are compared with their test performances. Finally, CAD system with best performed multiclass SVM classification approach is compared with NN classification system that was designed previously for the same purpose [12].

2 Materials and Methods

2.1 Subjects

Ultrasound image database is acquired from the patients who underwent a medical examination at the Department of Radio diagnosis, Postgraduate Institute of Medical Education & Research (PGIMER), Chandigarh, India during the period March 2008 to May 2009. After ultrasound examinations on 88 patients, a total of 111 US images comprising 95 focal liver lesion images (17 Cyst, 15 HCC, 18 HEM and 45 MET) and 16 normal liver images were acquired.

2.2 Image Enhancement

Ultrasound image database is enhanced by the regularized MSRAD-template 9 method proposed by Mittal *et. al*. [29]. This image preprocessing is used in order to improve the overall visual quality and contrast of ultrasound images.

2.3 Segmentation of Region-of-Interest

Regions-of-interest marked by an expert radiologist in each ultrasound image are segmented into maximum possible number of non-overlapping 25×25 sized regions and these regions are termed as segmented-regions-of-interest (SROIs). A SROI is used as texture sample representing a liver tissue category. The use of SROI for feature extraction reduces the complexity and calculation time. Non-overlapping SROIs are taken in order to design an accurate and robust system.

The 800 SROIs are segmented within the regions of interest in the acquired liver image database of 111 ultrasound images. A collection of SROIs is termed as dataset. The dataset having 800 SROIs is divided into training and testing sets. Training set contains 250 SROIs whereas remaining 550 SROIs are used for testing the classification systems.

2.4 Feature Extraction and Selection

The five texture based feature extraction methods from first, second and higher order statistics, along with spatial filtering and multi-resolutional approaches are used to achieve good classification accuracy. The methods chosen from these approaches are (i) first order statistics [30, 31], (ii) spatial gray level dependence matrix [32-34], (iii) gray level run length matrix [35, 36], (iv) texture energy measures [37] and (v) Gabor wavelet [38, 39] respectively.

All possible diagnostically important features are considered from the methods and an exhaustive set of 292 (9 of FOS, 80 of SGLDM, 11 of GLRLM, 72 of TEM and 120 of Gabor) features are extracted. The feature set is reduced further with the aim to retain only effective and discriminating features. The feature selection strategy, termed as sensitivity analysis, is designed to identify the features with significant information content from the extracted feature set and to discard the features those are not important from classification point of view. In sensitivity analysis, mean-range plot of each feature is studied to measure the sensitivity in providing discrimination among various liver tissues. Thus, the extracted features were reduced to a set of 208 sensitive features [12].

2.5 Classification

In the present work, two CAD systems are designed using two different multiclass SVM classification approach, one is one-against-one (OAO) approach and other is tree structured approach.

In OAO approach, SVM classifiers for all possible pairs of classes are created from the training set of N classes, each classifier being trained on only two out of N classes, giving a total of $N(N-1)/2$ classifiers. Applying each classifier to the test data gives one vote to the winning class. The data is assigned the label of the class with most votes.

In tree structured SVM approach (N-1) binary classifiers are trained, and each classifier separates a pair of classes at each node of the tree. The performance of tree structured SVM classifier depends on the structure of the binary tree. The closer the occurrence of classification error to the root node, the greater is the overall classification error. Schematic of the designed tree structure for focal lesion detection and classification is shown in Fig. 1.

Figure 1 demonstrates the tree structure with four nodes to detect abnormal (focal diseased) tissues and then classify them into four liver lesion categories. Four-binary SVM classifiers, one at each node, are used for this purpose. The binary-SVM at first node (N1) discriminates SROIs into normal and focal diseased liver tissues. According to pathology, focal diseased liver tissues can be classified into four major categories i.e. Cyst, HCC, HEM and MET. Ultrasound is especially useful in distinguishing Cysts (major category of fluid lesion) from solid lesions. Therefore, binary-SVM at the second node (N2) tries to discriminate all detected SROIs of focal tissues into the fluid lesion and solid lesion. Solid lesion tissues can be further categorized as solid benign and solid malignant.

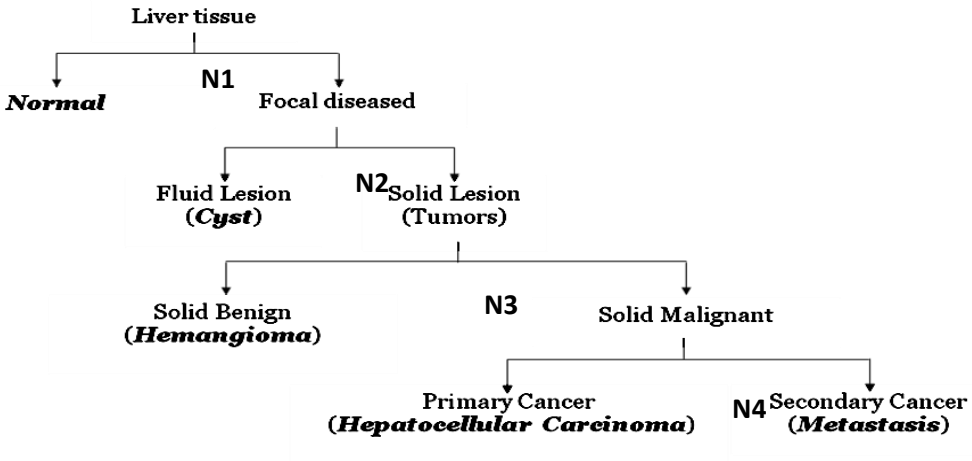

Figure. 1. Schematic of the designed tree structure for multi-class classification using binary SVM classifiers.

HEM is major category of solid benign and it can be detected easily on ultrasound with its typical appearance in comparison to the solid malignant lesions. Therefore, binary SVM at the third node (N3) is used to classify these lesions. Finally binary-SVM at the forth node (N4) is used to classify solid malignant lesions into primary malignant and secondary malignant. HCC is major category of primary malignant and MET is major category of secondary malignant. These liver tissue categories are not comfortably detected on ultrasound in comparison to the rest other mention liver tissues, therefore classified at the last node of the tree. Accordingly, liver tissues can be separated into five categories using the four-binary SVMs in tree structure.

In the designed CAD systems, SVM classifiers are implemented using LIBSVM tool, version 2.6. Radial basis kernel is applied in each SVM classifier. The parameters, i.e., penalty for errors (C) and RBF kernel width (σ), of each SVM classifier are adjusted properly for the optimized results. Selection of proper parameter settings of each binary classifier of the classification system is done with train dataset using cross validation procedure. Proper parameter setting is needed to optimize classification performance and to avoid over fitting and under fitting problems. In the experiments, the performance of a classifier is presented either in terms of confusion matrix with overall accuracy (in percentage) or the area under the ROC curve.

Finally, the performance of the two CAD systems is compared on the basis of their overall accuracy for liver tissue classification using ultrasound data.

3 Results

The results of designed CAD systems are presented in this section. First, the results of designed CAD system using OAO multiclass classification approach are studied. The performances of binary SVM classifiers, which are used in the implementation of this system, are evaluated initially. There are ten different binary classification tasks which can be performed by taking any two liver image categories at a time. Each binary classifier is first trained to get the best classification accuracy with the set of 208 sensitive features, and then the performance of this designed classifier is evaluated on test dataset. Evaluation is performed to assess whether any two liver image categories are adequately differentiated by the designed system or not.

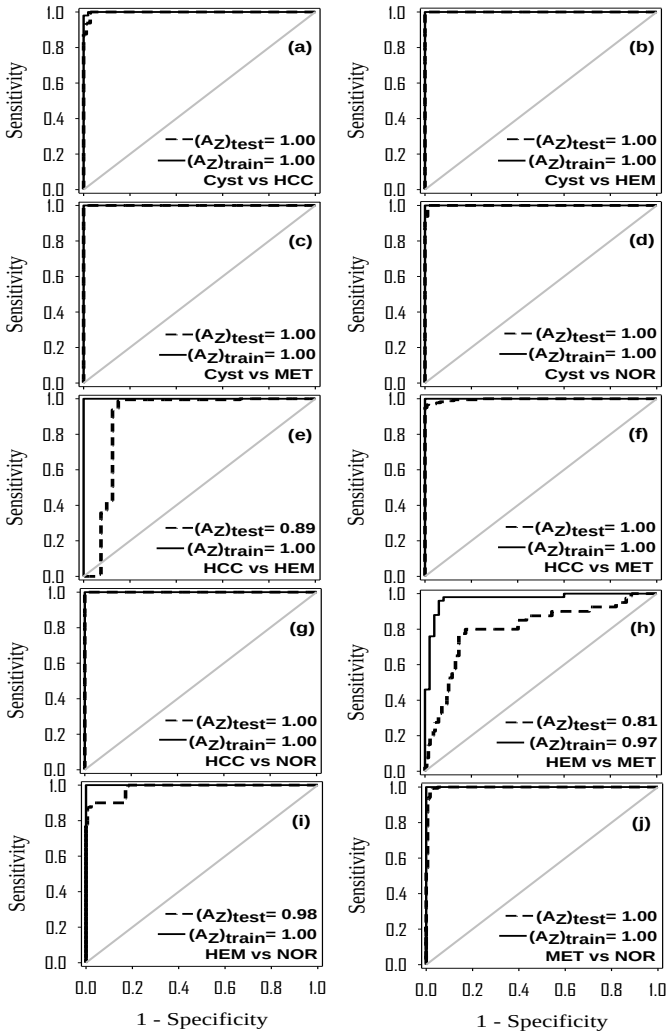

Figure 2. Training and test ROC curves for ten binary classification tasks performed with OAO approach based SVM classifiers.

Figures 2(a)-(j) show ROC curves with train and test datasets for different SVM based binary classification tasks. The results as shown in Fig. 2(a)-(d) depict that Cyst is clearly differentiated from rest other liver image categories having ideal performances ($(A_Z)_{test}$=1.00 and $(A_Z)_{test}$=1.00)) with train and test datasets. In Figure 2(e), the differentiation capability of designed SVM classifier for HCC vs. HEM is ideal on train dataset and A_Z value on test dataset is 0.89 which is reasonably high for classification. Further, HCC is ideally differentiated from MET and NOR liver image categories as shown in Fig. 2(f) and (g). Differentiation provided by the SVM classifier in between HEM and MET is not ideal even on training dataset as can be seen in Fig. 2(h). Still both, training and testing performances are impressively high ($(A_Z)_{train}$=0.97 and $(A_Z)_{test}$=0.81)). Finally, the rest two classification tasks, HEM vs. NOR and MET vs. NOR, depict high performances as shown in Fig. 2(i) and (j) respectively. Hence, it can be said that OAO approach based SVM provides sufficiently high discrimination on almost all two class classification tasks.

Table 1. Confusion matrix on the test results of OAO approach based multi-class SVM classifier.

Known class	Test results of OAO approach based multi-class SVM classifier					Class	Overall
	CYST	HCC	HEM	MET	NOR	Accuracy (%)	Accuracy (%)
CYST	15	1	0	0	0	93.8(15/16)	93.1(512/550)
HCC	2	170	3	1	1	96.0(170/177)	
HEM	0	0	33	7	0	82.5(33/40)	
MET	0	1	18	116	0	85.9(116/135)	
NOR	1	1	1	1	178	97.8(178/182)	

Table 1 shows the test results in the form of confusion matrix when above mentioned binary SVM classifiers are used for multiclass classification using OAO approach. Classification of Cyst is fairly high with one misclassified case out of sixteen. Out of 177 SROIs of HCC, 170 are detected correctly that yields 96% of class accuracy. Similarly high class accuracy for NOR liver tissues is obtained with the correct detection of 178 out of 182 SROIs. Further, the detection rate for HEM is comparatively low as only 32 SROIs out of 40 are detected correctly. Similarly, in case of MET, there is only 116 out of 135 tested SROIs are detected correctly. Finally, the overall accuracy of this CAD system is 93.1% as out of 550, 512 SROIs are classified correctly.

Tree structured multiclass SVM classification system first detects SROIs of diseased class and subsequently distinguishes among different lesion types. Train and test performances of classifier at each node in tree-structured multiclass SVM classification system are presented in form of ROC curves. These curves are shown in Fig. 3.

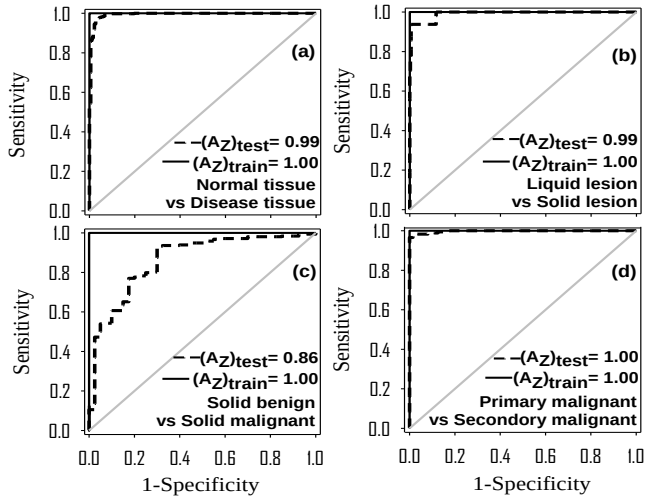

Figure 3. Training and test ROC curves of four binary SVM classifiers for the tasks performed at nodes of the tree structure.

In Figure 3(a), the differentiation capability of classification system in between normal and diseased tissues is ideal on train dataset and near to ideal ($(A_Z)_{test}$=0.99) on test data set. After the excellent performance at detection stage, all detected lesion SROIs are classified into liquid lesion and solid lesion. It can be observed from Fig. 3(b) that the A_Z values for the classification between liquid and solid lesions are 1.00 and 0.99 with the training and testing datasets, respectively. Further the classification of the detected solid lesion into solid benign (non-cancerous) and solid malignant (cancerous), is 100% on

training dataset having A_Z value of 1.00. This classification stage proves to be more challenging on test set with A_Z value of 0.86 as shown in fig. 3(c). The last stage is classification of solid malignant SROIs into primary and secondary cancer. The results of this classification step as appeared in fig. 3(d) depict that ideal performances (($A_Z)_{train}$=1.00 and $(A_Z)_{test}$=1.00)) are achieved by the system on train and test datasets and suggest perfect separation in between primary and secondary malignant. The performance of tree structured multiclass classification system is also evaluated in terms of class accuracy and overall accuracy and these results are listed in Table 2.

Table 2. Test performance of tree structured multiclass SVM classifier.

Tree structured multiclass SVM classifier		Accuracy	Overall accuracy
N1(first stage)	Normal/ Diseased	96.9% (533/550)	86.9%
N2(second stage)	Liquid lesion/Solid lesion	98.6% (363/368)	(478/550)
N3(third stage)	Solid benign/Solid malignant	87.5% (308/352)	
N4(last stage)	Primary malignant/Secondary malignant	98.1% (306/312)	

Table 3. Confusion matrices for all four stages of the designed tree structured SVM classification approach based CAD system.

	Focal diseased	**Normal**
Focal diseased (368)	361	7
Normal (182)	10	172
	Solid lesion	Liquid lesion
Solid lesion (352)	348	4
Liquid lesion (16)	1	15
	Solid Malignant	Solid Benign
Solid Malignant (312)	280	32
Solid Benign (40)	12	28
	Secondary Malignant	Primary Malignant
Secondary Malignant (135)	135	0
Primary Malignant (177)	6	171

Table 3 shows the confusion matrices reporting these results. The first stage of classification system is mass screening for lesion detection. At this stage, 96.9% (533/550) test SROIs are correctly classified as either normal or diseased. The specificity is 94.5% (172/182), while the sensitivity is 98.1% (361/368). In the next stage the accuracy achieved for the classification of detected lesions in between liquid lesions and solid lesions is 98.6% (363/368). System specificity and sensitivity to classify solid lesions are 93.8% (15/16) and 98.9% (348/352) which are high values in performances.

The results of confusion matrix for solid benign and solid malignant show that the test accuracy of this important stage is 87.5% (308/352). Solid malignant SROIs are classified with sensitivity 89.7% (280/312) and specificity 70% (28/40). Last stage is another important classification stage as here designed system is evaluated for distinction in between cancer classes, as primary malignant and secondary malignant. At this stage, system accuracy is 98.1% showing sensitivity in detection of primary lesion as 96.6% and sensitivity in detection of secondary malignant lesion as 100%. Finally, out of 550 SROIs, 478 are classified correctly and the overall accuracy comes out to be 86.9%.

The test results of OAO and tree structured multiclass SVM classification approaches are compared in terms of overall accuracy. Both of the CAD systems are designed and evaluated with same database and the set of 208 sensitive features, therefore the nature of the problem is same. The overall correct decisions with OAO approach based multiclass SVM classifier are 512 out of 550 SROIs yielding 93.1% overall accuracy. On the other hand, overall correct decisions with tree structured approach based CAD system are 478 out of 550 yielding 86.9% overall accuracy. Therefore OAO approach based CAD system is better as compared to tree structured based CAD system in terms of overall accuracy.

Table 4. Confusion matrix on the test results of CAD system based on OAO multi-class SVM classification approach for typical cases.

Known class	Test results of CAD system based on OAO multi-class SVM classification approach (Typical cases)					Class accuracy (%)	Overall accuracy (%)
	CYST	HCC	HEM	MET	NOR		
CYST	13	0	0	0	0	100.0(13/13)	96.6(367/380)
HCC	0	115	2	1	1	96.6(115/119)	
HEM	0	0	7	0	0	100.0(7/7)	
MET	0	1	4	54	0	91.5(54/59)	
NOR	1	1	1	1	178	97.8(178/182)	

Table 5 Confusion matrix on the test results of CAD system based on OAO multi-class SVM classification approach for atypical cases.

Known class	Test results of CAD system based on OAO multi-class SVM classification approach (Atypical cases)					Class accuracy (%)	Overall accuracy (%)
	CYST	HCC	HEM	MET	NOR		
CYST	2	1	0	0	0	66.7(2/3)	85.3(145/170)
HCC	2	55	1	0	0	94.8(55/58)	
HEM	0	0	26	7	0	78.8(26/33)	
MET	0	0	14	62	0	81.6(62/76)	
NOR	0	0	0	0	0	0	

The performance of OAO approach based CAD system is also evaluated by examining class accuracies of typical and atypical SROIs in test set with separate confusion matrices as shown in Tables 4 and 5 respectively. As given in Table 4, the class accuracies for typical SROIs of Cyst, HCC, HEM, MET and NOR are 100% (13/13), 96.6% (155/119), 100.0% (7/7), 91.5% (54/59) and 97.8% (178/182) respectively with the overall accuracy of 96.6% (367/380). Table 5 shows the test results from atypical set of SROIs. The overall accuracy of 85.3% (145/170) is observed for atypical cases. The class accuracies for atypical SROIs of Cyst, HCC, HEM and MET are 66.7% (2/3), 94.8% (55/58), 78.8% (26/33) and 81.6% (62/76) respectively.

Finally test results of OAO approach based CAD system are compared with the NN based CAD system [12]. This comparison is made to decide which CAD system out of these two shows the better performance in classification among liver tissues with the selected set of 208 sensitive features. It is needed to mention that a validation set is used in process of development of NN based CAD system [12]. By definition, validation set is the test set that is used to evaluate the performance of neural network model during training process for the best model selection and to control the over-fitting problem of designed system. Therefore, the results of NN based CAD system on validation set and test set are put together for the comparison of these results with the results of SVM based CAD system. A confusion

matrix is compiled with the test results of 550 test SROIs to show the exact performance comparison in between NN and SVM based CAD systems.

Table 6 Confusion matrix representing the comparison of the test results on the set of 550 SROIs with NN and SVM based CAD systems

Known class		Test results on the set of 550 SROIs					Class accuracy (%)	Overall accuracy (%)	
		CYST	HCC	HEM	MET	NOR		NNM	SVM
CYST	NN-CAD	15	1	0	0	0	93.8(15/16)	87.6	93.1
	SVM-CAD	15	1	0	0	0	93.8(15/16)	(482/550)	(512/550)
HCC	NN-CAD	2	165	6	3	1	93.2(165/177)		
	SVM-CAD	2	170	3	1	1	96.0(170/177)		
HEM	NN-CAD	0	2	33	5	0	82.5(33/40)		
	SVM-CAD	0	0	33	7	0	82.5(33/40)		
MET	NN-CAD	2	4	25	100	4	74.1(100/135)		
	SVM-CAD	0	1	18	116	0	85.9(116/135)		
NOR	NN-CAD	0	0	5	8	169	92.9(169/182)		
	SVM-CAD	1	1	1	1	178	97.8(178/182)		

Table 6 shows the confusion matrix with the combined test results of these two CAD systems. The result analysis of both the systems reveals that one SROI of Cyst is misclassified as HCC, yielding equal class accuracy for both systems. The class accuracy of HCC with SVM system is 2.8% higher than that with NN based CAD system. The results of HEM classification are same for both the systems. Out of 135 SROIs of MET, 100 and 116 SROIs are correctly classified by NN and SVM NN based CAD systems respectively, thus yielding 11.9% higher class accuracy with SVM system. Similarly the class accuracy of NOR liver tissue with SVM system is 7% higher than that with NN system. Thus the OAO approach based SVM system outperforms the NN system with their overall accuracy of 93.1% and 87.6% respectively. Finally as SVM system is better in performance with 30 more correct decisions in comparison with NN system, it can be said that the test results obtained by SVM classifier are more robust as compared to the NN system for the same feature set and dataset used.

4 Discussions

Ten binary classifiers are designed and tested to implement OAO approach based CAD system. There is the limitation in providing clear differentiation in between HEM and MET in the designed binary classifier with given feature set which can be seen in Fig. 2(h) showing no ideal performance on training dataset. Still in the case, both training and testing performances are imposingly high to accept the designing of this classifier. Table 1 shows the performance of OAO approach based CAD system using these binary classifiers for Cyst, HCC, and NOR liver tissues are sufficiently high with class accuracy of 93.8%, 96.0%, and 97.8% respectively. This SVM classification approach shows reasonable accuracy of 80.0% and 85.9% for HEM and MET respectively.

The performance of tree structured based CAD system depends on the designing of binary classifiers at tree nodes. Binary classifiers at tree nodes are trained to separate SROIs into the desired classes. Tree structured based CAD system for mass screening is highly promising with high sensitivity in lesion detection on test dataset missing only 1.9% (7/361) of SROIs from diseased category. Further, system sensitivity for solid lesion classification form diseased cases is quite high missing only 1.1% (4/352) cases. However, system performance is limited at the third stage of the tree structured CAD system. The

results of confusion matrix, as shown in Table 3, reflect that solid benign (non-cancerous) and solid malignant (cancerous) are particularly hard to separate for the system. The test accuracy of this important stage is 87.5% (308/352) which shows performance degradation in comparison to the previous stages, but still this accuracy is satisfactory as images of these classes appear most of the time with overlapping characteristics on ultrasound images and create confusion even to the expert radiologist during visual evaluation. The ability to correctly identify cancer SROIs within the testing set is promising, considering the difficulty of the task. Finally, at the last stage system performance is excellent with accuracy and sensitivity values on test dataset.

Further, the choice for a multiclass approach depends on the problem in hand and it needs the consideration of the accuracy requirement along with the computational time, the resources available and the nature of the problem in hand. In classification problem with large number of classes, the number of binary SVM classifiers in OAO approach are much higher than those in tree structured based approach and therefore the processing time in OAO approach based multi-class classification would be higher than that in tree structured based approach. However, in classification problem with small number of classes like the present one where the number of classes is five, only ten binary classifiers are needed in OAO approach, so difference of processing time in both the types of approach is not significant and therefore one should go for the approach which gives the better results.

Finally, it can be concluded that one-against-one approach based CAD system provides higher overall accuracy, with correct decision on 34 more SROIs showing an improvement of 6% in correct decisions. Moreover, overall accuracy and class accuracies, all are sufficiently high to demonstrate the effective performance of the designed CAD system with typical cases. The test performances of atypical cases are lesser than that of typical cases; still these results are promising for the application of OAO approach based CAD system with atypical cases too.

Results of SVM system are better than that of NN system even after improvement of NN classifier by introducing two-step classification concept. Thus it can be said that as compared to NN system, the SVM system has high anticipation level in the diagnosis of focal liver disorders. The test results of each SROIs from both the systems are further analyzed to go in more details. It can be seen that there is one misclassification case out of 16 SROIs of Cyst by both the systems. But the SROI that is misclassified by NN system is different from the SROI that is misclassified by SVM system. That means the SROI that is misclassified by SVM system is correctly classified by the NNM system. Several similar instances have been observed on analyzing the test results of each SROIs. Thus it can be said that NN system has a potential to detect those cases which are not correctly detected by the SVM system, therefore importance of NN system cannot be ignored. Even, the problems of overfitting likely occur because the complexity of the diagnostic model is not controlled in empirical risk minimization on which neural network is based. To be useful in practice where only limited training examples are available, the non-asymptotic analysis of the quality of empirical risk minimization is necessary. To address these issues, structural risk minimization has been suggested in SVM theory. Structural risk minimization minimizes the upper bound on the generalization error, as against empirical risk minimization which minimizes the error on the training data

5 Conclusion

The results leads to the conclusion that SVM-CAD system with tree structured approach can produce a high accuracy rate of 86.9% in focal liver lesion detection and classification. The approach is valuable to improve the accuracy of the diagnosis of liver cancer and to reduce the number of biopsies. CAD system with OAO approach based multi-class SVM classification in comparison to that of tree structured approach based multi-class SVM classification provides the better results in terms of overall accuracy of 93.1% with the selected set of 208 sensitive features. The proposed OAO-SVM system outperforms the system designed with neural network using the same database. The overall classification accuracy with this system has been achieved as 96.6% for typical cases and 85.3% for atypical cases.

ACKNOWLEDGMENTS

Author acknowledges the Prof. Vinod Kumar, Indian Institute of Technology, Roorkee, India and Department of Radiodiagnosis and Imaging, Postgraduate Institute of Medical Education and Research, Chandigarh, India for their support in carrying out this research work.

REFERENCES

[1] Herszenyi, L., Tulassay, Z., 2010. Epidemioligy of gastrointestinal and liver tumors. European Review for Medical and Pharmacological Sciences 14, 249-258.

[2] Sujana, H., Swarnamani, S., Suresh, S., 1996. Application of artificial neural networks for the classification of liver lesions by image texture parameters. Ultrasound in Medicine and Biology 22, 1177-1181.

[3] Yoshida, H., Casalino, D.D., Keserci, B., Coskun, A., Ozturk, O., Savranlar, A., 2003. Wavelet-packet-based texture analysis for differentiation between benign and malignant liver tumors in ultrasound images. Physics in Medicine and Biology 48, 3735–3753.

[4] Balasubramanian, D., Srinivasan, P., Gurupatham, R., 2007. Automatic classification of focal lesions in ultrasound liver images using principal component analysis and neural networks. In: Proceedings of 29th International Conference IEEE EMBS, Lyon, France, pp. 2134-2137.

[5] Poonguzhali, S., Ravindran, G., 2008. Automatic classification of focal lesions in ultrasound liver images using combined texture features. Information Technology Journal 7, 205-209.

[6] Krishna, M.M.R., Banerjee, S., Chakraborty, C., Chakraborty, C., Ray, A.K., 2010. Statistical analysis of mammographic features and its classification using support vector machine, Expert Systems with Applications 37, 470-478.

[7] Tsantis, S., Cavouras, D., Kalatzis, I., Piliouras, N., Dimitropoulos, N., Nikiforidis, G., 2005. Development of a support vector machine-based image analysis system for assessing the thyroid nodule malignancy risk on ultrasound. Ultrasound in Medicine and Biology 31, 1451-1459.

[8] Cao, G., Shi, P., Hu, B., 2005. Liver fibrosis identification based on ultrasound images. In: Proceedings of 27th IEEE Annual Conference on Engineering in Medicine and Biology, Sanghai, China, pp. 6317-6320.

[9] Zhang, J., Wang, Y., Dong, Y., Wang, Y., 2007. Ultrasonographic feature selection and pattern classification for cervical lymph nodes using support vector machine. Computer Methods and Programs in Biomedicine 88, 75-84.

[10] Krishan, A., Mittal, D., 2015. Detection and Classification of Liver Cancer using CT Images.International Journal on Recent Technologies in Mechanical and Electrical Engineering 2, 93-98.

[11] Mishra, S. K., Mittal., D., sunkaria, R. K., 2015. Designing of Computer Aided Diagnostic System for the Identification of Exudates in Retinal Fundus Image. Journal of Biomedical Engineering and Medical Imaging 2, 29-40.

[12] Mittal, D., Kumar, V., Saxena, S.C., Khandelwal, N., Kalra, N., 2011. Neural network based focal liver lesion diagnosis using ultrasound images. Computerized Medical Imaging and Graphics 35, 315–323.

[13] Jain, A.K., Duin, R.P.W., Mao, J., 2000. Statistical pattern recognition: a review. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 4–37.

[14] Bottou, L., Cortes, C., Denker, J., Drucker, H., Guyon, I., Jackel, L., LeCun, Y., Muller, U., Sackinger, E., Simard, P., Vapnik, V., 1994. Comparison of classifier methods: A case study in handwriting digit recognition. In: Proceedings of International Conference on Pattern Recognition, pp. 77–87.

[15] Friedman, J., 1996. Another approach to polychotomous classification. Stanford, CA: Department of Statistics and Stanford Linear Accelerator Center, Stanford University, available at http://www-stat.stanford.edu/~jhf/ftp/poly.pdf.

[16] Vapnik, V.N., 1998. Statistical Learning Theory. Wiley, New York.

[17] Allwein, E.L., Schapire, R.E., Singer, Y., 2000. Reducing multiclass to binary: A unifying approach for margin classifiers. Journal of Machine Learning Research 1, 113–141.

[18] Platt, J.C., Cristianini, N., Shawe-Taylor, J., 2000. Large margin DAGs for multiclass classification, in: Sollam, S.A., Leen, T.K., Mu¨ller, K.-R. (Eds.), Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA, pp. 547–553.

[19] Dietterich, T.G., Bakiri, G., 1995. Solving multiclass learning problems via error-correcting output codes. Journal of Artificial Intelligence Research 2, 263–286.

[20] Crammer, K., Singer, Y., 2001. On the algorithmic implementation of multiclass kernel-based vector machines. Journal of Machine Learning Research 2, 265–292.

[21] Chen, Y., Crawford, M.M., Ghosh, J., 2004. Integrating support vector machines in a hierarchical output space decomposition framework. In: Proceedings of International Geoscience and Remote Sensing Symposium, pp. 949–952.

[22] García-Pedrajas, N., Ortiz-Boyer, D., 2006. Improving multiclass pattern recognition by the combination of two strategies. IEEE Transactions on Pattern Analysis and Machine Intelligence 28, 1001-1006.

[23] Fei, B., Liu, J., 2006. Binary tree of SVM: A new fast multiclass training and classification algorithm. IEEE Transactions on Neural Network 17, 696–704.

[24] Pujol, O., Radeva, P., Vitià, J, 2006. Discriminant ECOC: A heuristic method for application dependent design of error correcting output codes. IEEE Transactions on Pattern Analysis and Machine Intelligence 28, 1007-1012.

[25] Chen, J., Wang, C., Wang, R., 2008. Combining support vector machines with a pairwise decision tree. IEEE Geoscience and Remote Sensing Letters 5, 409–413.

[26] Rifkin, R., Klautau, A., 2004. In defense of one-vs-all classification. Journal of Machine Learning Research 5, 101–141.

[27] Hsu, C., Lin, C., 2002. A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Network 13, 415–425.

[28] Chen, J., Wang, C., Wang, R., 2009. Using stacked generalization to combine SVMs in magnitude and shape feature spaces for classification of hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing 47, 2193-2205.

[29] Mittal, D., Kumar, V., Saxena, S.C., Khandelwal, N., Kalra, N., 2010. Enhancement of the ultrasound images by modified anisotropic diffusion method. Medical & Biological Engineering & Computing 48, 1281–1291.

[30] Golemati, S., Tegos, T.J., Sassano, A., Nikita, K.S., Nicolaides, A.N., 2004. Echogenicity of B-mode sonographic images of the carotid artery. Journal of Ultrasound in Medicine 23, 659–669.

[31] Sonnad, S.S., 2002. Describing data: statistical and graphical methods. Radiology 225, 622–628.

[32] Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Texture features for image classification. IEEE Transactions on Systems Man and Cybernetics SMC-3, 610-621.

[33] Soh, L.-K., Tsatsoulis, C., 1999. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Transactions on Geoscience and Remote Sensing 37(2), 780-795.

[34] Clausi, D.A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. Canadian Journal of Remote Sensing 28, 45-62.

[35] Galloway, M.M., 1975. Texture analysis using gray level run lengths. Computer Graphics and Image Processing 4, 172-179.

[36] Tang, X., 1998. Texture information in run-length matrices. IEEE Transactions on Image Processing 7(11), 1602-1609.

[37] Laws, K.I., 1980. Rapid texture identification. SPIE 238, 376-380.

[38] Arivazhagan, S., Ganesan, L., Priyal, S.P., 2006. Texture classification using Gabor wavelets based rotation invariant features. Pattern Recognition Letters 27, 1976-1982.

[39] Manjunath, B.S., Ma, W.Y., 1996. Texture features for browsing and retrieval of image data. IEEE Transactions on Pattern Analysis and Machine Intelligence 18, 837–842.

