ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

SVM-Based Characterization of Focal Kidney Lesions from B-Mode Ultrasound Images.

Shailja Rana[1]*, Shruti Jain[1], and Jitendra Virmani[3].

[1]Jaypee University of Information Technology, Solan, HP, India.
[2]Thapar University Patiala, Punjab, India.

ABSTRACT

Characterization of benign and malignant focal kidney lesions such as angiomyolipomas (AMLs) and renal cell carcinomas (RCCs) from ultrasound images is an intimidating task for radiologists due to their enormously overlaying sonographic presences. In the present work the main focus is on the aspect of variations in texture patterns shown by focal kidney lesions. In order to visualise these textural disparities, texture features are calculated using different methods namely statistical features and spectral texture features. These texture features have been calculated from region of interest (ROIs) extracted from each image within the lesion. The study is performed on 23 ultrasound kidney images with 14 AML lesions and 9 RCC lesions. For classification task support vector machine classifier has been used to classify data based on analysis. It has been seen that by combining Gray level run length statistics and Fourier power spectrum features produces the maximum accuracy of 84% with ICA values of 100 % and 60 % for RCC and AML lesions respectively.

Keywords: Computer aided diagnosis system, Focal kidney lesions, Texture features, Angiomyolipoma, Renal cell carcinoma, SVM classifier.

Corresponding author

ISSN: 0975-8585

INTRODUCTION

Kidney is a vital organ of urinary system having bean shape structure and a parenchyma region. In human body, the expansion of tissue masses is the result of uncontrollable growth of cells called tumors. Kidney tumors are generally solid masses and are considered as an anomalous growth within the kidney. These are of two types: Benign tumors (Non-Cancerous) and Malignant tumors (Cancerous).

Several imaging modalities like X-ray, ultrasound (US), computerized tomography (CT), magnetic resonance imaging (MRI) etc. can be used for identifying kidney irregularities. Ultrasonography is commonly preferred examination for diagnosing the diseases of soft tissue organs such as kidney disorders because of its real-time, non- radioactive and non-invasive nature [1]. US imaging modality help to differentiate the solid masses from simple cysts in the kidneys. The problems associated with US imaging are the artifacts due to the patient movement and apparatus margins. Hence, these limitations limit the clarity of subjective diagnosis. In the recent task, researchers have shown significant interest to provide the radiologists with a computer aided diagnosis (CAD) system for texture analysis so as to provide additional diagnostic information about kidney tissues and to support them in image reading, understanding and decision making [2].

Among kidney focal lesions our study is based on classification of primary benign lesions (AMLs) and primary malignant lesions (RCCs). AML is a most commonly occurring benign kidney lesion and RCC is a most common type of malignant kidney lesion. Texture patterns are generated by the echo being reflected by the kidney tissue [2]. The echogenic appearance of the tumor is assumed to be associated to its fat content. Therefore, echogenicity is the most distinctive ultrasonographic feature of AML. The sonographic appearance of small AMLs (<2cms) is hyperechoic with homogeneous echotexture.For large AMLs (>2cms) sonographic appearance is hyperechoic with heterogeneous echotexture. The sonographic appearance of small RCCs (<2cms) can be hypoechoic, hyperechoic or they may exhibit mixed echogenicity. For large RCCs (>2cms) sonographic appearance is hypoechoic with increasing heterogeneity as they grow in size.
The US images of *(a)* small AML lesion *(b)* large AML lesion *(c)* small RCC lesion *(d)* large RCC lesion are shown in Figure 1.

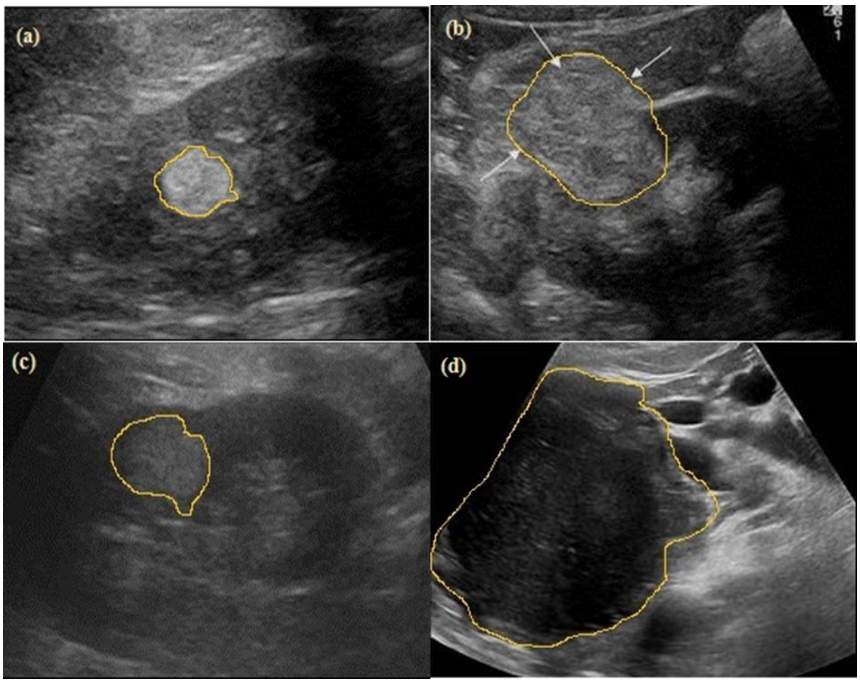

Figure 1: B-mode ultrasound kidney images depicting different cases. *(a)* small AML lesion *(b)* large AML lesion *(c)* small RCC lesion *(d)* large RCC lesion

Note: The sonographic appearance of small AML (<2cms) is hyperechoic with homogeneous echotexture. The sonographic appearance of large AML (>2cms) is hyperechoic with heterogeneous echotexture. The sonographic appearance of small RCC (<2cms) can be hyperechoic, hypoechoic or they may exhibit mixed echogenicity. The sonographic appearance of large RCC (>2cms) is hypoechoic with increasing heterogeneity as they grow in size.

The participating radiologist opined that sonographic differentiation between small AMLs and small RCCs is a challenging task for the radiologists faced during routine clinical practice. Therefore, to correctly identify these lesions, a computerized tissue characterization system i.e. a CAD system can indeed be helpful for differential diagnosis between these lesions. For the characterization of focal kidney lesions, rare studies are described in literature. The description of related studies is shown below in table1:

Table 1: Description of related studies carried out for kidney image classification.

Authors	Dataset Description			
	Kidney image classes	ROI size	No. of images	Features Extracted
Raja et al [1]	Nor, MRD, Cyst	SKT	150	Gabor wavelet
Raja et al [2]	Nor, MRD, Cyst	SKT	150	Statistical, MI, Power spectral and Gabor
Raja et al [3]	Nor, MRD, Cyst	SKT	150	Power spectral
Jose et al [4]	Nor, MRD, Cyst	SKT	35	Histogram and GLCM
Akkasaligar et al [5]	Nor and Cyst	SKT	52	GLCM and GLRLM
Subramanya et al [6]	Nor, MRD, Cyst	32 × 32	35	FOS, MI, GLCM, GLRLM and Laws' mask
Raja et al [7]	Nor, MRD, Cyst	SKT	150	Statistical, MI and Power spectral
Raja et al [8]	Nor, MRD, Cyst	SKT	150	Statistical and MI

Note: MI: Moment invariant features, FOS: First order statistics features, GLCM: Gray length co-occurrence matrix features, GLRLM: Gray level run length matrix features, Nor: Normal, MRD: Medical renal diseases, SKT: Segmented kidney tissue

From the above studies, it is observed that most of the work in recent years on ultrasound kidney images has considered three kidney image classes namely Normal, MRD and Cyst [1-4] [6-8]. It has been observed that only one study reported CAD system design on fixed size ROIs extracted from the kidney parenchyma [6]. In the present work, using SVM classifier a CAD system design has been proposed for the classification of primary benign and primary malignant focal kidney lesions based on their underlying texture characteristics.

MATERIALS AND METHODS

Data Set Collection and Description

In this work, a publically available benchmark database of ultrasound kidney images has been used. The images in the database are classified into two types namely, AML and RCC. Out of 23 B-mode kidney US images, 14 images are AML images and 9 images are RCC images. Size of each image is 300 × 225 pixels having 96 dpi vertical and horizontal resolutions with a gray scale comprising of 256 tones. The database of 23 images contains the nature and position of irregularity existing in the kidney. The digital images are attained from Aloka ultrasound scanning system and Hitachi high end ultrasound systems [9].

Selection of Region of Interest (ROI)

For calculating various texture features, the selection of ROI size should be chosen carefully in such a way so that it should deliver adequate number of pixels [10]. In previous studies, it has been specified by many researchers that estimate reliable statistics size of ROI must be at least 800 pixels [11]. In the present study, from each lesion maximum non-overlapping ROIs of size 32 × 32 are extracted manually. The participating radiologist recommended that, ROIs should be taken from within the lesions of kidney avoiding the renal sinus region which is a cavity consisting of fats, blood vessels, nerves, etc. Images from the database with marked ROIs are illustrated in Figure 2.

ISSN: 0975-8585

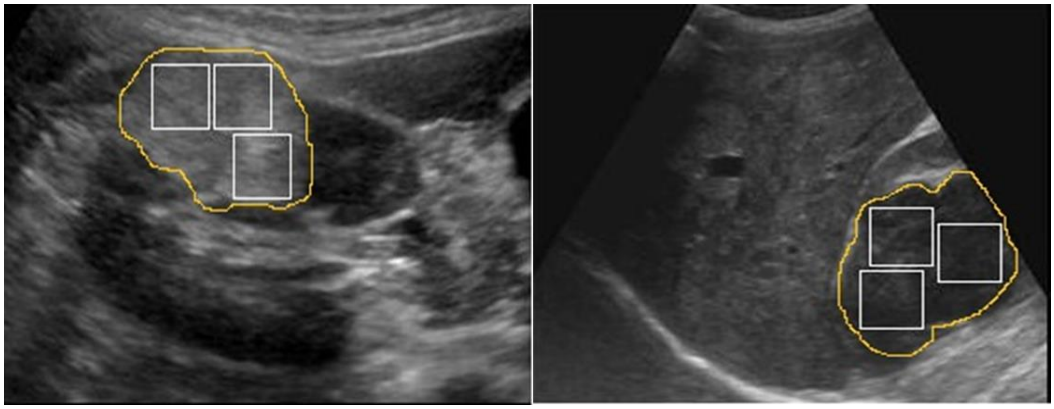

Figure 2: Images with marked ROIs. *(a)* AML image *(b)* RCC image

The splitting of ROIs in training dataset and testing dataset is depicted in Figure 3.

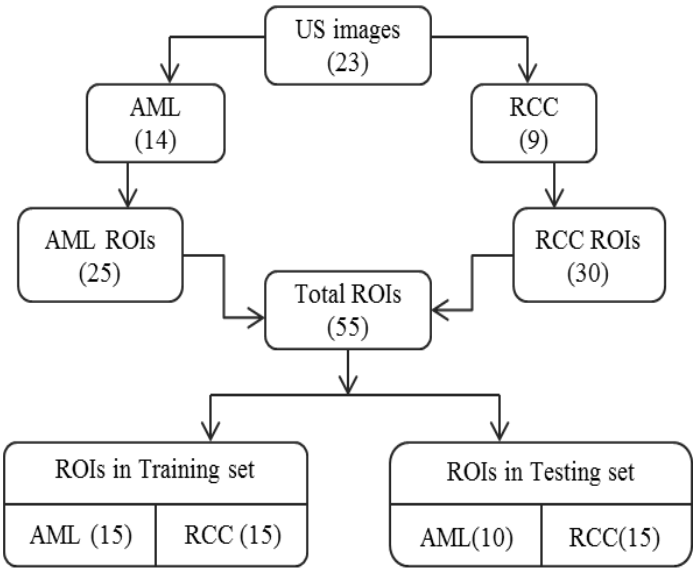

Figure 3: Description of dataset for two class classification

Proposed CAD System Design

In the present study, the Computer aided diagnosis system has been proposed for the classification of primary kidney focal lesions by using kidney US images. The main purpose of emerging such a CAD system is to support radiologists, by providing additional diagnostic information to discriminate between AMLs and RCCs anomalies which might be sometimes missed by radiologists on visual observation. When AMLs are less than 2cms, radiologists may confuse and become doubtful with small RCCs. Thus, distinction between small RCCs and AMLs is crucial. Therefore in this work, a CAD system design has been anticipated for differentiating the primary benign kidney lesions from primary malignant kidney lesions. In Figure 4, the block diagram representation of the proposed CAD system design is shown.

ISSN: 0975-8585

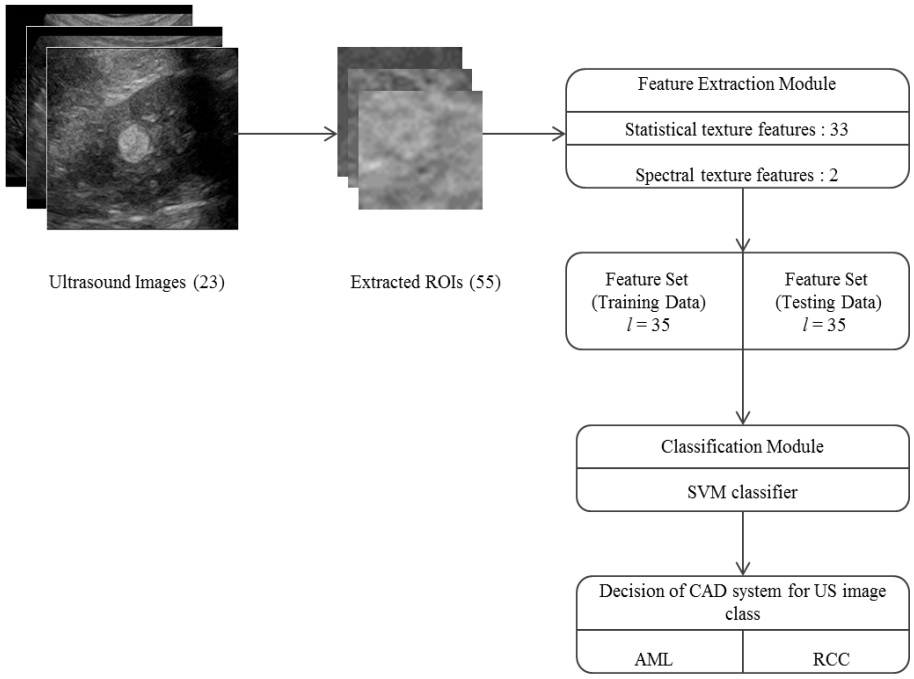

Figure 4: Block diagram representation of proposed Computer aided diagnosis system design for two class classification. Note: *l*: Length of feature set.

The CAD system is comprised of ROI extraction module, feature extraction module and classification module. For executing the CAD system design, out of 23 images of database 55 ROIs are manually extracted. In *ROI extraction module:* many ROIs of size 32 × 32 are taken out from inside the kidney lesion. In *feature extraction module:* various kind of texture features are calculated from ROIs using numerous statistical texture features [12] [13-19] and spectral texture features [15] [20]. Now, the dataset is distributed into training and testing dataset. In *feature classification module*, SVM classifier has been used for the classification task so as to give the suitable results for the two kidney classes.

Feature Extraction Module

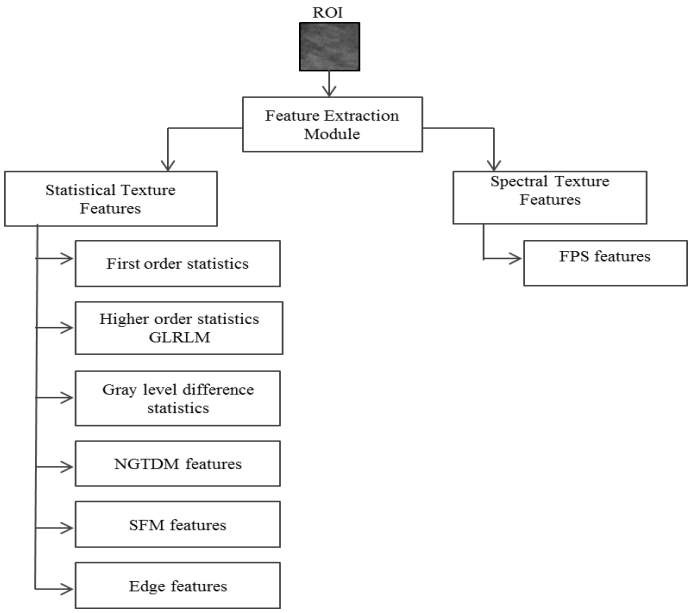

Figure 5: Texture features computed for each ROI image.

Note: GLRLM: Gray level run length matrix, NGTDM: Neighborhood gray tone difference matrix, SFM: Statistical feature matrix, FPS: Fourier power spectrum

In present work, feature extraction module is used to calculate the mathematical descriptors for understanding the fundamental ROI textural properties. These descriptors are either texture based features or shape based features (morphological features). The radiologist tells that, for the differential analysis between focal kidney lesions, only texture features gives the appropriate information. The different types of feature extraction methods are illustrated in Figure 5.

First order statistics (FOS) features

The first order statistics (FOS) are derived from the gray level intensity histograms of the image. For each ROI, six FOS texture features namely average gray level, smoothness, standard deviation, third moment, uniformity and entropy are calculated for each ROI [12].

Higher order statistics

Gray level run length matrix (GLRLM) features are used to calculate the higher order statistics. By using different combinations of intensities texture features are calculated. Gray level run is made up of the set of successive pixels of gray levels and the run length (means the number of times a run occurs). Total 11 GLRLM features are calculated i.e. low gray level run emphasis, high gray level run emphasis, short run emphasis, long run emphasis, short run low gray level emphasis, short run high gray level emphasis, long run low gray level emphasis, long run high gray level emphasis, gray level non uniformity, run length non-uniformity and run percentage **[13] [14]**.

Gray Level Difference Statistics (GLDS) features

On the basis of co-occurrence of a pixel pair having a certain absolute difference in gray-levels, five GLDS features are extracted such as homogeneity, energy, contrast, entropy and mean **[15] [16]**.

Neighbourhood Gray Tone Difference Matrix (NGTDM) Features

With the neighbouring pixels and with a definite gray scale, NGTDM imitates a grayscale difference between pixels Five NGTDM features are calculated such as complexity, coarseness, contrast, strength and busyness **[17] [18]**.

Statistical Feature Matrix (SFM) features:

SFM calculates the statistical properties of pixels such as (1) the matrix size is dependent on the maximum distance used instead of the number of gray-levels and (2) the matrix can be extended very easily at different distances. SFM computes four features namely periodicity, coarseness, roughness and contrast.

Edge features

More information about texture is always present at the edges. In order to measure the spatial variation of gray levels across an image, gradient of an image is used. If there is no abrupt change in gray level then gradient will be low otherwise it will be high. Two texture features, i.e. absolute gradient variance and absolute gradient mean are computed in edge features [19].

FPS features

By using Fourier power spectrum features , two spectral features namely angular sum and radial sum of discrete Fourier transform has been computed **[15] [20]**.

From the above methods, in Table 2 the following texture feature vectors (TFVs) have been computed.

ISSN: 0975-8585

Table 2: Description of extracted TFVs for characterization of AML and RCC lesions

Features	TFV	Texture Features	*l*
Statistical features	TFV1	FOS features	6
	TFV2	GLRLM features	11
	TFV3	GLDS features	5
	TFV4	NGTDM features	5
	TFV5	SFM features	4
	TFV6	Edge features	2
Spectral features	TFV7	FPS features	2

Note: TFV: Texture feature vector, *l*: Length of TFV

Feature Classification Module

To predict the class membership of indefinite data, a supervised machine learning technique i.e. Support vector machine (SVM) classifier has been used in this paper. The data set used in this study comprises of 55 ROIs. The training set contains 30 ROIs and testing set contains 25 ROIs.

Support Vector Machine (SVM) Classifier

The SVM classifier refers to a class of supervised machine learning algorithms. It is based on the idea of decision planes that label the decision boundary. SVM classifier is a kernel based classifier which uses kernel functions for mapping the data from input space to higher dimensional feature space. For the present study LibSVM library has been used for the execution of SVM classifier [21]. The best values of regularization parameter C and kernel parameter γ has been obtained using grid research procedure. In this study, on the training data ten-fold cross validation is approved out for each grouping of (C, γ), such that C ϵ $\{2^{-4}, 2^{-3} \ldots 2^{15}\}$ and γ ϵ $\{2^{-12}, 2^{-11} \ldots 2^{4}\}$. Hence, the selection of kernel parameter γ and regularization parameter C is very important for better performance of the classifier [6] [11] [20] [22-29].

Classifier performance evaluation criteria

The Overall classification accuracy (OCA) and individual class accuracy (ICA) parameters are used to measure the performance of CAD system. Calculations of OCA and ICA are done by using the confusion matrix (CM).

RESULTS AND DISCUSSION

For measuring the performance of the proposed CAD system design, rigorous experimentation has been done in this work to differentiate different kidney lesions using SVM classifier on the basis of features extracted.

Experiment 1: To find the classification performance of AML and RCC Kidney lesions using different Statistical features by SVM classifier.

In this experiment, different statistical features classification performance is measured using SVM classifier. The results of experiment 1 are depicted in Table 3.

Table 3: Result of Classification performance of statistical features using SVM classifier for two-classes of Kidney

Features	CM			OCA (%)	ICA_{AML} (%)	ICA_{RCC} (%)
FOS		AML	RCC	64.0	50.0	73.3
	AML	5	5			
	RCC	4	11			

GLRLM	AML	6	4	76.0	60.0	86.6
	RCC	2	13			
GLDS	AML	3	7	24.0	30.0	20.0
	RCC	12	3			
NGTDM	AML	4	6	52.0	40.0	60.0
	RCC	6	9			
SFM	AML	4	6	36.0	40.0	33.3
	RCC	10	5			
Edge	AML	5	5	72.0	50.0	86.6
	RCC	2	13			

Note: CM: Confusion matrix, AML: Angiomyolipoma class, RCC: Renal Cell Carcinoma class, OCA: Overall classification accuracy, ICA_{AML}: Individual class accuracy for angiomyolipoma, ICA_{RCC}: Individual class accuracy for renal cell carcinoma.

From the table 3 it has been observed that, from GLRLM features, the highest OCA of 76.0 % is achieved for statistical features using SVM classifier. Out of 25 testing cases, 6 cases (6/25) are not correctly classified in GLRLM feature. It is also observed that the maximum ICA for AML is 60.0 % and maximum ICA for RCC is 86.6 %.

Experiment 2: To find the classification performance of AML and RCC Kidney lesions using different Spectral texture features (Fourier Power Spectrum (FPS) texture features) by SVM Classifier. In this experiment, FPS features classification performance is assessed using SVM classifier. The results of experiment 2 are depicted in Table 4.

Table 4: Result of Classification performance of Fourier Power Spectrum features using SVM classifier for two-classes of kidney

Features	CM			OCA (%)	ICA_{AML} (%)	ICA_{RCC} (%)
		AML	RCC			
FPS	AML	8	2	72.0	80.0	66.6
	RCC	5	10			

Note: CM: Confusion matrix, AML: Angiomyolipoma class, RCC: Renal Cell Carcinoma class, OCA: Overall classification accuracy, ICA_{AML}: Individual class accuracy for angiomyolipoma, ICA_{RCC}: Individual class accuracy for renal cell carcinoma

From the table 4 it has been observed that, for Fourier Power Spectrum texture features, the highest OCA 72.0 % is achieved using SVM classifier. It is also observed that the maximum ICA for AML is 80.0 % and maximum ICA for RCC is 66.6 %. Out of 25 testing cases, 7 cases (7/25) are not correctly classified in FPS feature by using SVM classifier. From this experiment it can be observed that amongst statistical texture features, higher order statistics based GLRLM features (TFV2) yield the highest OCA of 76% and the second highest OCA of 72% is achieved with spectral domain FPS features (TFV7). Another experiment, experiment no 3 was conducted to evaluate the performance of combined feature vector (CFV) consisting of higher order GLRLM features and FPS features (CFV = TFV2 + TFV7).

Experiment 3: To find the classification performance of AML and RCC Kidney lesions using combined texture feature by SVM classifier.

Table 5: Result of Classification performance of combination of statistical and spectral features using SVM classifier for two-classes of kidney

Features	CM			OCA (%)	ICA_{AML} (%)	ICA_{RCC} (%)
		AML	RCC			
GLRLM + FPS	AML	6	4	84.0	60.0	100.0
	RCC	0	15			

Note: CM: Confusion matrix, AML: Angiomyolipoma class, RCC: Renal Cell Carcinoma class, OCA: Overall classification accuracy, ICA_{AML}: Individual class accuracy for angiomyolipoma, ICA_{RCC}: Individual class accuracy for renal cell carcinoma

In this experiment, combined features (statistical and spectral features) classification performance is measured using SVM classifier. The results of experiment 3 are depicted in Table 5.

From the table 5 it has been observed that, by CFV comprising of GLRLM and FPS features, the maximum OCA of 84.0 % is achieved with ICA of 100 % for RCC and with ICA of 60% for AML. Out of 25 testing cases, 21 cases (21/25) are correctly classified and 4 cases (4/25) are not correctly classified by using SVM classifier.

CONCLUSION

This work indicates that a study on the characterization of focal kidney lesions has been carried out. From the rigorous experimentation it can be observed that, for the classification of focal kidney lesions, the CAD system gives the highest OCA of 84.0 % from SVM classifier. It can be concluded that, for the differential analysis between primary benign and primary malignant kidney lesions, the performance of CFV i.e. combined feature vectors are obtained by combining TFV2 (Gray level run length matrix (GLRLM) feature) with TFV7(Fourier Power Spectrum (FPS) feature) provide the highest OCA of 84.0 %. CFV also give the highest individual class accuracy (ICA) of 100 % for RCC. Therefore, combination of these two features (GLRLM+FPS) is important for achieving best results which accounts for the textural variations shown by the primary focal lesions. The attained results of proposed CAD system design specify their effectiveness to support radiologists for the differential diagnosis of AML and RCC kidney lesions during regular medical check-ups.

ACKNOWLEDGEMENT

The authors are faithful to Dr. Shruti Thakur, Department of Radiology, Indira Gandhi Medical College (IGMC), Shimla (H.P) for meaningful discussions regarding the understanding of sonographic appearances of focal kidney lesions.

REFERENCES

[1] Raja BK, Madheswaran M, Thyagarajah K. Mach Vision and Appl, 2010; 21: 287–300.
[2] Raja BK, Madheswaran M, Thyagarajah K. J Med Syst, 2008; 32: 65–83.
[3] Raja BK, Madheswaran M, Thyagarajah K. J Med Syst, 2007; 31: 307–317.
[4] Jose JS, Sivakami R, Maheswari NU, Venkatesh R. International Journal of Computer Technology and Electronics Engineering, 2012; 2.
[5] Akkasaligar TP, Biradar S. International Conference on Information and Communication Technologies, 2014; 0975 – 8887.
[6] Subramanya MB, Kumar V, Mukherjee S, Saini M. Journal of Digital Imaging, Society for Imaging Informatics in Medicine, 2014.
[7] Raja BK, Madheswaran M, Thyagarajah K. International conference on Computing, Theory and Applications, ICCTA'07, 2007, pp. 382–388.
[8] Raja BK, Madheswaran M, Thyagarajah K. In Proceedings of IEEE International Conference on Signal Processing, Communications and Networking, ICSCN'07, 2007, pp. 483–487.
[9] Ultrasound cases from the teaching files from the Gelderse Vallei Hospital in Ede, the Netherlands,http://www.ultrasoundcases.info/.
[10] Suckling J, Parker J, Dance DR, Astley S, Hutt I, Boggis CRM, Ricketts I, Stamatakis E, Cerneaz N, Kok SL, Taylor P, Betal D, Savage J. Springer, 1994; 1069 : 375-378.
[11] Virmani J, Kumar V, Kalra N, Khandelwal N. Journal of Digital Imaging, 2013; 26: 530-543.
[12] Virmani J, Kumar V, Kalra N, Khandelwal N. IEEE International Conference on Multimedia, Signal Processing and Communication Technologies, IMPACT, Aligarh, India, 2011; pp. 212–215.
[13] Xu DH, Kurani AS, Furst JD, Raicu DS. Heart, 2004; 2 : 25-30.
[14] Albregtsen F, Image, 1995; 1: 3-8.
[15] Weszka JS, Dyer CR, Rosenfeld A. 1976; 6: 269-285.
[16] Kim JK, Park HW. IEEE Transactions on Medical Imaging, 1999; 18: 231-238.
[17] Castella C, Kinkel K, Eckstein MP, Sottas PE, Verdun FR, Bochud F. Academic Radiology, 2007; 14: 1486-1499.
[18] Amadasun M, King R. IEEE Transactions on Systems, Man and Cybernetics, 1989; 19.1264-1274.
[19] Castellano G, Bonilha L, Li LM, Cendes. 2004; 59 : 1061-1069.

[20] Virmani J, Kumar V, Kalra N, Khandelwal N. Journal of Digital Imaging, 2013; 26.
[21] Chang CC, Lin CJ. LIBSVM, a library of support vector machine. Software available at http://www.csie.ntu.edu.tw/~cjlin /libsvm, Accessed 15 Jan 2016.
[22] Jain S. PhD Thesis, Jaypee University of Information Technology, 2012.
[23] Jain S, Naik PK. System International Journal of Pharma and BioSciences, 2012; .3: 358-373.
[24] Jain S, Chauhan DS. International Conference on Information and Communication Technology for Sustainable Development, Ahmedabad, 2015, pp 81-88.
[25] Jain S, Chauhan DS. International Journal of Pharma and Bio Sciences, 2015; 6: 164-176.
[26] Burges CJC. Data Min Knowl Disc, 1998; 2: 1-43.
[27] Jain S., Control System and Power Electronics, Bangalore, 2015: pp 84-93.
[28] Jain S, Naik PK, Bhooshan SV. International Conference on Computational Intelligence and Communication Networks, Gwalior, 2011; pp 565-568
[29] Virmani J, Kumar V, Kalra N, Khandelwal N. Journal of Medical Engineering and Technology, 2013; 37, 292-306.