Original Research Article   **Bioinformatics**

CLASSIFICATION OF FOCAL KIDNEY LESIONS USING WAVELET-BASED TEXTURE DESCRIPTORS

SHAILJA RANA*[1], SHRUTI JAIN[2] AND JITENDRA VIRMANI

[1, 2] Jaypee University of Information and Technology, Solan, H.P, India
[3] Thapar University Patiala, Punjab, India

ABSTRACT

A decision making computer aided diagnostic (CAD) system has been proposed in this paper for the classification of focal kidney lesions using various wavelet energy descriptors .The focal kidney lesions are categorised into two kidney classes namely primary benign i.e. angiomyolipoma (AML) and primary malignant i.e. renal cell carcinoma (RCC). The study is performed on 47 kidney ultrasound images with 22 AML lesions and 25 RCC lesions. The multi-resolution wavelet based texture descriptors are calculated from area of interests (AOIs) of variable sizes by using different types of wavelet filters such as Haar, Daubechies, biorthogonal, symlets and coiflets filters. It has been seen that by using SVM classifier, Daubechies (db4) and symlets (sym3) wavelet energy descriptors gives the best overall classification accuracy of 82.6 %, with the individual class accuracy (ICA) values of 63.6 % for Angiomyolipoma and 100 % for Renal cell carcinoma.

KEYWORDS: Computer aided diagnostic system, Focal kidney lesions, Wavelet texture descriptors, Angiomyolipoma, Renal cell carcinoma, Support vector machine classifier.

SHAILJA RANA
Jaypee University of Information and Technology, Solan, H.P, India

Corresponding Author

INTRODUCTION

Kidney tumors are found to be one of the topmost health problems nowadays and are mostly found in people of age 60. Kidney tumors are typically solid masses having irregular shape and these tumors are formed due to the abnormal growth of cells within the kidney. Kidney tumors are basically of two kinds: Non-Cancerous (Benign tumors) and Cancerous (Malignant tumors). There are various methods for the identification of kidney deformities such as ultrasound (US), magnetic resonance imaging (MRI), X- Ray, computerized tomography (CT) etc. For detecting these solid masses, ultrasound is one of the best imaging modality in medical imaging and soft tissue organ diagnosis because it has low cost, non-radioactive, real-time and non-invasive properties.[1] Ultrasound can easily differentiate between normal and abnormal kidneys and it does not require chemical substance like iodine containing contrast medium (dye) to see the internal functioning of kidney. Although there are few artifacts in US imaging like equipment margin, inter observer variations, image quality degradation etc. which limits its performance.[2] Due of these difficulties for the characterization of kidney tissues, feature extraction becomes tough. To overcome this problem, a second opinion device i.e. a computer aided diagnostic (CAD)

system has been proposed to provide output of computer as a second option which makes the final decisions for radiologists in image reading.[3] In this study we have seen that echogenicity plays a very important role to reflect the tissue characteristics of kidneys. The term echogenicity of kidney means the ability of numerous functional parts of kidney to produce echo signals, during ultrasonography. On the basis of echogenicity focal kidney lesions are classified into two classes: (a) Angiomyolipomas (i.e. (AMLs) primary benign lesion) (b) Renal cell Carcinomas (i.e. (RCCs) primary malignant lesion). These are the most commonly occurring lesions of kidney. For small size AMLs (i.e. <2cms) the echotexture of lesion is homogeneous with high echogenicity (having hyperechoic sonographic appearance). For large size AMLs (i.e. > 2cms) the echotexture of lesion is heterogeneous with high echogenicity. For small size RCCs (i.e. <2cms) the echotexture of lesion is homogeneous and their echogenicity may be hyperechoic, isoechoic, hypoechoic or may exhibit mixed echogenicity. For large size RCCs (i.e. > 2cms) the echotexture of lesion is heterogeneous with low echogenicity (having hypoechoic sonographic appearance).The sonographic appearances of AML and RCC lesions are depicted in Figure 1.
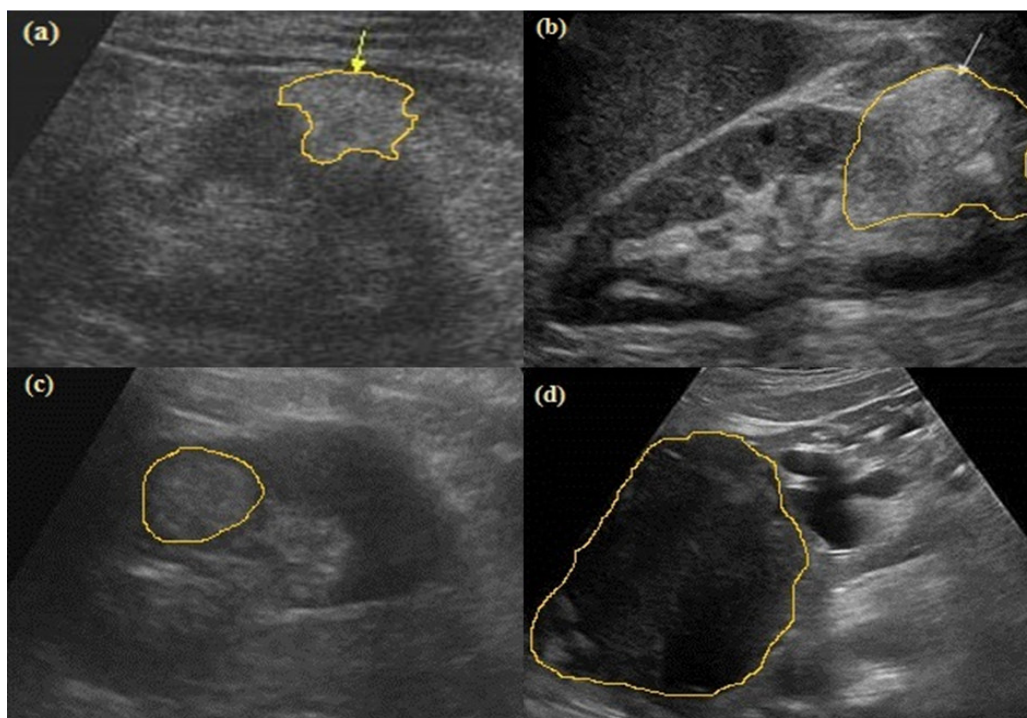

Figure 1
Ultrasound kidney images showing different cases.
(a) small AML lesion (b) large AML lesion
(c) small RCC lesion (d) large RCC lesion

Note: Small size AML (<2cms) having hyperechoic sonographic appearance with homogeneous echotexture. Large size AML (>2cms) having hyperechoic sonographic appearance with heterogeneous echotexture. Small size RCC (<2cms) sonographic appearance can be hyperechoic, isoechoic, hypoechoic or they may exhibit mixed echogenicity. Large size RCC (>2cms) having hypoechoic sonographic appearance with increasing heterogeneity as they grow in size.

The participating radiologist suggested that, on the basis of echogenicity, it has become difficult to absolutely differentiate between small size AMLs and small size

RCCs lesions and it is a tough task for radiologists faced during daily medical practice. Therefore, to appropriately analyze and identify these deformities a CAD system is

extremely expected to help radiologists in medicinal environment. For the classification of focal kidney lesions, very few studies are labeled in literature. The brief depiction of related studies is shown in table1:

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

Authors	Dataset Description			
	Kidney image classes	AOI size	No. of images	Features Extracted
Raja et al [1]	Nor, MRD, Cyst	SKT	150	Power spectral
Raja et al [2]	Nor, MRD, Cyst	SKT	150	Gabor wavelet
Raja et al [3]	Nor, MRD, Cyst	SKT	150	Statistical, MI, Power spectral and Gabor
Akkasaligar et al [4]	Nor and Cyst	SKT	52	GLCM and GLRLM
Jose et al [5]	Nor, MRD, Cyst	SKT	35	Histogram and GLCM
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: AOI: Area of interest, Nor: Normal, MRD: Medical renal diseases, SKT: Segmented kidney tissue, MI: Moment invariant features, FOS: First order statistics features, GLCM: Gray length co-occurrence matrix features, GLRLM: Gray level run length matrix features

The related studies for classification of renal diseases using ultrasound images are few and most of these studies have considered normal, medical renal disease and cyst image classes. In the present work a CAD system has been proposed for the classification of focal kidney lesions using wavelet energy descriptors.

MATERIALS AND METHODS

2.1 Data Set Collection and Description
The present work has been done on publically available benchmark database.[9] The database contains 47 B-mode kidney US images (22 images of AML lesion and 25 images of malignant lesion). Each image is of the size of 300 × 225 pixels having 96 dpi horizontal and vertical resolutions with 256 gray scale tones.

2.2 Selection of Area of Interests (AOIs)
The size of Area of Interest (AOI) has been selected wisely in such a manner that it must deliver a good statistical population for making the correct decisions in CAD system.[10, 11] In the present work, single AOI of variable sizes are extracted manually from inside the lesions of kidney. Sample images with extracted AOIs from focal kidney lesions are shown in Figure 2. Depending on the size of lesions, the largest rectangular AOI has been cropped from the region inside the lesion. The participating radiologist opined that the shape or margin features are not important for differential diagnosis between focal kidney lesions. Therefore an attempt has been made to take the largest AOI from each lesion leaving the boundary of the lesion so as to compute reliable estimates of texture features.
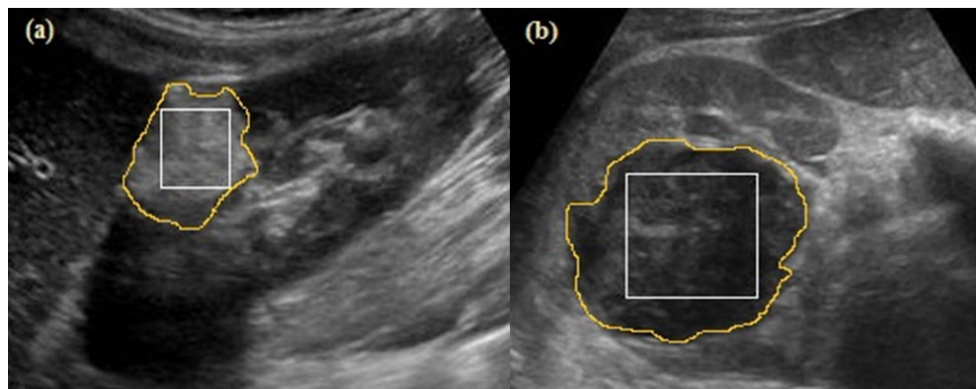

Figure 2
Images of extracted AOIs. (a) AML image (b) RCC image

2.3 Proposed CAD System Design
For enhancing the confidence level of radiologists, the CAD systems are used in the medical imaging as a second opinion tool. Figure 3 shows the block diagram of the proposed CAD system design for the prediction of kidney lesions.
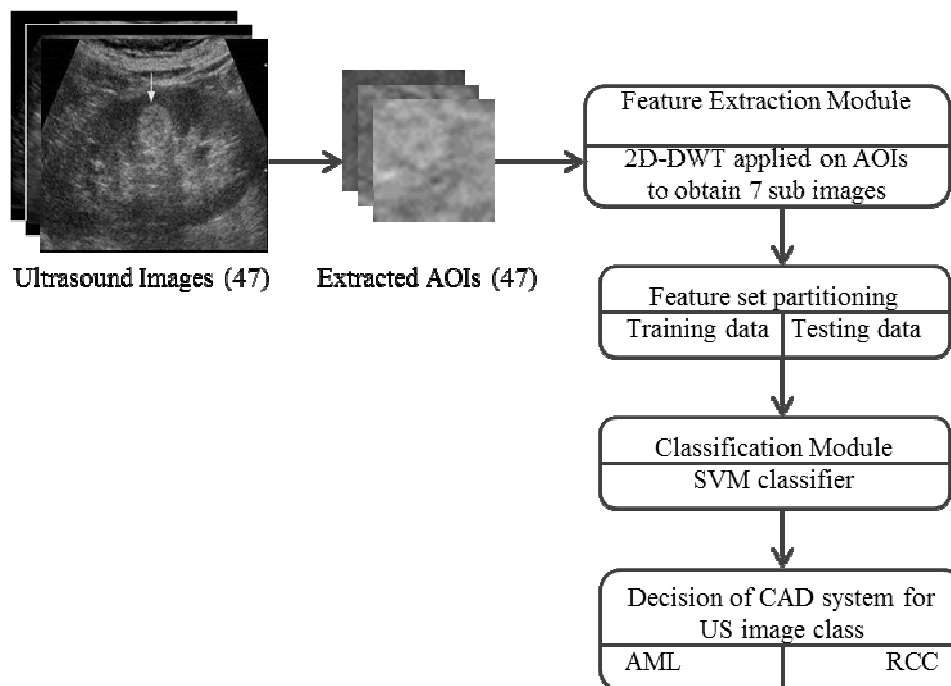
Figure 3
Block diagram representation of proposed CAD system design

The objective of emerging a decision making CAD system is, to support radiologists by providing extra diagnostic information in image reading and to differentiate between various anomalies. In the present study, it has been described that when size of AMLs are less than 2cms, radiologists get confused with small size RCCs and thus the chances of uncertainty increases for the discrimination between small RCCs and small AMLs. Therefore, CAD system has been proposed for differential diagnosis between focal kidney lesions by using kidney US images. The CAD system mainly consists of three blocks namely AOI extraction module, feature extraction module and classification module. For the CAD system execution, total of 47 images of database has been taken, out of which total 47 single AOIs are extracted manually.

AOI extraction module
In the present work, single AOI of variable sizes such as 32 × 32, 40 × 40, 64 × 64, 96 × 96, 100 × 100, 128 × 128 etc. are manually extracted from within the lesion. The data set contains total 47 AOIs with 22 AML AOIs (from 22 AML kidney images) and 25 RCC AOIs (from 25 RCC images)

Feature extraction module
In this study, various kinds of texture features are calculated from AOIs using numerous wavelet energy descriptors. The dataset is distributed into training and testing dataset. By using multi-resolution scheme such as discrete wavelet transform (DWT), feature extraction is carried out in transform domain over various scales. In various studies, it has been shown that the effective depiction of texture depends upon the wavelet filter properties. Therefore in order to obtain the best wavelet filter for the present task of kidney lesions classification, ten dissimilar compact support wavelet filters such as Biorthogonal (bior3.1, bior3.3 and bior4.4), Coiflets (coif1 and coif2), Haar (db1), Daubechies (db4 and db6) and Symlets (sym3 and sym5) filters have been used to obtain sub-band images for each AOI image.[11-14] Each compact support wavelet filter consists of seven texture feature vectors. By using 2D-DWT (Discrete wavelet transform), when AOI image is decomposed up to 2nd level, then one approximate sub image and six orientation detail sub images are generated. From the obtained sub images, normalized energy is calculated. Total seven wavelet energy descriptors chosen as texture feature vectors (TFVs) are shown in Table 2. TFV1 consists of 7 features, TFV2 consists of 3 features, TFV3 consists of 6 features, TFV4 consists of 4 features, TFV5 consists of 4 features, TFV6 consists of 4 features and TFV 7 consists of 3 features are shown in Table 2.

Table 2
Description of Seven wavelet energy descriptors as TFVs.

TFV	Wavelet energy descriptors in TFVs	l
TFV1	$\left(\dfrac{\|A_2\|_F^2}{area(A_2)}, \dfrac{\|D_2^{(h)}\|_F^2}{area(D_2^{(h)})}, \dfrac{\|D_2^{(v)}\|_F^2}{area(D_2^{(v)})}, \dfrac{\|D_2^{(d)}\|_F^2}{area(D_2^{(d)})}, \dfrac{\|D_1^{(h)}\|_F^2}{area(D_1^{(h)})}, \dfrac{\|D_1^{(v)}\|_F^2}{area(D_1^{(v)})}, \dfrac{\|D_1^{(d)}\|_F^2}{area(D_1^{(d)})} \right)$	7
TFV2	$\left(\dfrac{\|D_1^{(h)}\|_F^2}{area(D_1^{(h)})}, \dfrac{\|D_1^{(v)}\|_F^2}{area(D_1^{(v)})}, \dfrac{\|D_1^{(d)}\|_F^2}{area(D_1^{(d)})} \right)$	3
TFV3	$\left(\dfrac{\|D_2^{(h)}\|_F^2}{area(D_2^{(h)})}, \dfrac{\|D_2^{(v)}\|_F^2}{area(D_2^{(v)})}, \dfrac{\|D_2^{(d)}\|_F^2}{area(D_2^{(d)})}, \dfrac{\|D_1^{(h)}\|_F^2}{area(D_1^{(h)})}, \dfrac{\|D_1^{(v)}\|_F^2}{area(D_1^{(v)})}, \dfrac{\|D_1^{(d)}\|_F^2}{area(D_1^{(d)})} \right)$	6
TFV4	$\left(\dfrac{\|D_1^{(h)}\|_F^2}{area(D_1^{(h)})}, \dfrac{\|D_1^{(v)}\|_F^2}{area(D_1^{(v)})}, \dfrac{\|D_1^{(d)}\|_F^2}{area(D_1^{(d)})}, \dfrac{\|D_2^{(d)}\|_F^2}{area(D_2^{(d)})} \right)$	4
TFV5	$\left(\dfrac{\|A_2\|_F^2}{area(A_2)}, \dfrac{\|D_1^{(h)}\|_F^2}{area(D_1^{(h)})}, \dfrac{\|D_1^{(v)}\|_F^2}{area(D_1^{(v)})}, \dfrac{\|D_1^{(d)}\|_F^2}{area(D_1^{(d)})} \right)$	4
TFV6	$\left(\dfrac{\|A_2\|_F^2}{area(A_2)}, \dfrac{\|D_2^{(h)}\|_F^2}{area(D_2^{(h)})}, \dfrac{\|D_2^{(v)}\|_F^2}{area(D_2^{(v)})}, \dfrac{\|D_2^{(d)}\|_F^2}{area(D_2^{(d)})} \right)$	4
TFV7	$\left(\dfrac{\|D_2^{(h)}\|_F^2}{area(D_2^{(h)})}, \dfrac{\|D_2^{(v)}\|_F^2}{area(D_2^{(v)})}, \dfrac{\|D_2^{(d)}\|_F^2}{area(D_2^{(d)})} \right)$	3

Note: TFV: Texture feature vector, I: length of TFV, A: Approximate sub image, D: Detail sub image, h: horizontal direction, v: Vertical direction, d: Diagonal direction, F: Frobenius norm. A_i or D_i i is the level of decomposition

Feature classification module

Classification examines the numerical properties of numerous image features and arranges data into classes. It has been established that the classifier designs which use regularization like support vector machines are less prone to over-fitting and obtain good generalization performance to even without feature space dimensionality reduction. In kNN classifier, an unknown instance is classified on the basis of a majority vote of its neighbors and most common class is selected amongst its k nearest neighbors. The performance of kNN classifier is adversely affected by the presence of outliers in the data. Therefore in the present work, SVM classifier has been chosen for the classification task. Classification procedures normally do processing in two phases: *training* phase and *testing phase*. The classifier named as SVM classifier has been used for the classification task so as to provide the appropriate results for the two classes of kidney. This classifier is a supervised machine learning classifier because the classes are already defined for the training sets. For the present study LibSVM library has been used for the execution of SVM classifier.[15] Kernel based SVM is a best classifier for the class separability with very less no. of errors. Kernels such as polynomial and Gaussian radial basis function are normally used kernels. By using these kernel functions non - linear training data is mapped from input to higher dimension space. The right choice of the kernel parameter γ and regularization parameter C is important in order to obtain a good performance and the best values of C and γ has been obtained using grid research procedure. In this study, for each combination of (C, γ), ten-fold cross validation is done on the training data such that $C \in \{2^{-4}, 2^{-3} \ldots 2^{15}\}$ and $\gamma \in \{2^{-12}, 2^{-11} \ldots 2^{4}\}$.[6, 11, 13-30] SVM classifier is considered as a best classifier because of its more speed and it gives more accuracy than k-nearest neighbors (kNN) and probabilistic neural network (PNN) classifiers.

RESULTS

The results are depicted in Table 3 and Table 4. From table3, it is observed that out of all seven TFVs, the maximum accuracy is attained from db4 (Daubechies) and sym3 (Symlets) wavelet filters. It is also seen that highest overall classification accuracy (OCA) of 82.6 % is achieved for db4 and sym 3 filters with TFV1 and TFV3 using SVM classifier. The minimum classification accuracy of 39.1 % is achieved for db4 wavelet filter using SVM classifier.

Table 3
Classification performance for the seven TFVs using SVM classifier.

TFV	I	Max. Acc. (%)	Wavelet filter	Min. Acc. (%)	Wavelet filter
TFV1	7	82.6	db4	60.8	bior3.1, bior3.3
TFV2	3	52.1	bior3.3	39.1	db4
TFV3	6	82.6	sym3	52.1	bior3.1, db1
TFV4	4	78.2	coif2, db4	43.4	bior3.3
TFV5	4	78.2	bior3.1	47.8	db4, coif2
TFV6	4	69.5	bior3.3	56.5	bior4.4, coif1, sym3
TFV7	3	65.2	bior3.1	52.1	coif1, coif2, db4, db6, sym5

Note: I: Length of TFV, Max. Acc.: Maximum accuracy, Min. Acc.: Minimum accuracy

Table 4

Classification performance for the best TFVs using SVM classifier.

TFV	CM			OCA (%)	ICA_{AML} (%)	ICA_{RCC} (%)
		AML	RCC			
TFV1	AML	7	4	82.6	63.6	100.0
	RCC	0	12			
TFV3	AML	7	4	82.6	63.6	100.0
	RCC	0	12			

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

A proficient CAD system always provides good precision with reduced number of features used for classifier design.

DISCUSSION

For evaluating the performance of the proposed CAD system design, severe experimentation has been done in this work to obtain the classification performance and to discriminate primary kidney lesions using various wavelet energy descriptors with SVM classifier.The papers in literature are few and all the papers are related to 3 kidney class namely normal, medical renal disease (MRD) and cyst. There is no work done on benign and malignant lesions and we have done the work on focal kidney lesions in order to capture maximum information of texture features. The result in this paper is different and classification has been done to capture maximum information of texture features and this is a first attempt to classify the focal lesions of the kidney. The best results obtained from the above experiments are summarised in Table 4.Therefore, the obtained results of proposed CAD system design specify their usefulness to help radiologists for the differential diagnosis of AML and RCC kidney lesions during routine medical check-ups.

CONCLUSION

From the above experimentation it can be concluded that, for the characterization of focal kidney lesions, wavelet energy descriptors i.e. db4 and sym3 filters contain significant information. These wavelet filters provides the highest OCA of 82.6 % and gives the highest individual class accuracy (ICA) of 100 % for RCC.

ACKNOWLEDGEMENT

REFRENCES

1. Raja BK, Madheswaran M, Thyagarajah K. Ultrasound kidney image analysis for computerized disorder identification and classification using content descriptive power spectral features. J Med Syst. 2007; 31(5): 307–317.
2. Raja BK, Madheswaran M, Thyagarajah K. Texture pattern analysis of kidney tissues for disorder identification and classification using dominant Gabor wavelet. Mach Vision and Appl. 2010; 21(3): 287–300.
3. Raja BK, Madheswaran M, Thyagarajah K. A hybrid fuzzy-neural system for computer-aided diagnosis of ultrasound kidney images using prominent features. J Med Syst. 2008; 32(1): 65–83.
4. Akkasaligar TP, Biradar S. Classification of Medical Ultrasound Images of Kidney. International Conference on Information and Communication Technologies (ICICT). 2014: 0975 – 8887.
5. Jose JS, Sivakami R, Maheswari NU, Venkatesh R. An Efficient Diagnosis of Kidney Images Using Association Rules International Journal of Computer Technology and Electronics Engineering (IJCTEE). 2012; 2(2): 2249-6343.
6. Subramanya MB, Kumar V, Mukherjee S, Saini M. SVM-Based CAC System for B-Mode Kidney Ultrasound Images. Journal of Digital Imaging. Society for Imaging Informatics in Medicine. 2014.
7. Raja BK, Madheswaran M, Thyagarajah K. Analysis of ultrasound kidney images using content descriptive multiple features for disorder identification and ANN based classification. International conference on Computing, Theory and Applications. ICCTA'07. 2007: 382–388.
8. Raja BK, Madheswaran M, Thyagarajah K. Evaluation of tissue characteristics of kidney for diagnosis and classification using first order statistics and RTS invariants. In Proceedings of IEEE International Conference on Signal Processing, Communications and Networking. ICSCN'07. 2007: 483–487.
9. Ultrasound cases.info [online]; 2015 [cited 2015 June]. Available from: http://www.ultrasoundcases.info/Category.aspx?cat=8 7
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. The mammographic image analysis society digital mammogram database. In: Gale EG et al. (Eds.). Digital Mammography. Berlin, Heidelberg: Springer; 1994. p. 375-378.
11. Virmani J, Kumar V, Kalra N, Khandelwal N. SVM-based characterization of liver ultrasound images using wavelet packet texture descriptors. Journal of Digital Imaging. 2013; 26(3): 530-543.
12. Li X, Tian Z. Wavelet energy signature: comparison and analysis. In: King I, Wang J,

Chan LW, Wang DL, (Eds.). Neural Information Processing 2006, Part II. LNCS, Heidelberg: Springer; 2006. p. 474-480.

13. Virmani J, Kumar V, Kalra N, Khandelwal N. Prediction of liver cirrhosis based on multi-resolution texture descriptors from B-mode ultrasound. Int. J. Converg. Comput. 1. 2013: 19–37.

14. Kumar I, Bhadauriab H S, Virmani J. Wavelet Packet Texture Descriptors based four-class BIRADS Breast Tissue Density Classification. 4th International Conference on Eco-friendly Computing and Communication Systems, Procedia Computer Science 70. 2015: 76 – 84.

15. Chang CC, Lin CJ. LIBSVM, a library of support vector machine [Internet]; 2016 [cited 2016 Jan]. Available from: http://www.csie.ntu.edu.tw/~cjlin/libsvm

16. Jain S. Communication of signals and responses leading to cell survival / cell death using Engineered Regulatory Networks. PhD Thesis, Jaypee University of Information Technology, Solan, Himachal Pradesh, India. 2012.

17. Jain S, Naik PK. System Modeling of cell survival and cell death: A deterministic model using Fuzzy System. International Journal of Pharma and BioSciences (IJPBS). 2012; 3(4): 358-373.

18. Jain S, Chauhan DS. Linear and Non Linear Modeling of Protein Kinase B/ AkT. International Conference on Information and Communication Technology for Sustainable Development (ICT4SD - 2015), Ahmedabad, India. 2015.

19. Burges CJC. A tutorial on support vector machines for pattern recognition. Data Min Knowl Disc. 1998; 2(2): 1-43.

20. Jain S. Mathematical Analysis and Probability Density Function of FKHR pathway for Cell Survival /Death. Control System and Power Electronics – CSPE 2015, Bangalore. 2015 Aug 1.

21. Jain S, Naik PK, Bhooshan SV. Nonlinear Modeling of cell survival/ death using artificial neural network. International Conference on Computational Intelligence and Communication Networks (CICN2011), Gwalior, India. 2011 Oct 7: 565-568.

22. Virmani J, Kumar V, Kalra N, Khandelwal N. A comparative study of computer-aided classification systems for focal hepatic lesions from B-mode ultrasound. Journal of Medical Engineering and Technology. 2013; 37: 292-306.

23. Rana S, Jain S, Virmani J. Classification of Kidney Lesions using Gabor Wavelet Texture Features. In Proceedings of the 10th INDIACom 3rd 2016 International Conference on Computing for Sustainable Global Development. 2016 March 16: 2528-2532.

24. Virmani J, Kumar V, Kalra N, Khandelwal N. Prediction of cirrhosis from liver ultrasound B-mode images based on Laws' masks analysis. In Proceedings of IEEE International Conference on Image Information Processing, ICIIP-2011, Waknaghat, HP, India. 2011: 1–5.

25. Bhusri S, Jain S, Virmani J. Classification of Breast Lesions based on Laws' Feature Extraction Techniques. In Proceedings of the 10th INDIACom 3rd 2016 International Conference on Computing for Sustainable Global Development. 2016 March 16: 2523-2527.

26. Bhusri S, Jain S, Virmani J. Breast Lesions Classification using the Amalagation of morphological and texture features. International Journal of Pharma and BioSciences. 2016 Apr-Jun; 7(2): 617-624.

27. Virmani J, Kumar V, Kalra N, Khandelwal N. SVM based characterization of liver cirrhosis by singular value decomposition of GLCM matrix. International Journal of Artificial Intelligence and Soft Computing. 2013; 3: 276-296.

28. Kriti, Virmani J, Dey N, Kumar V. PCA-PNN and PCA-SVM based CAD systems for breast density classification. In: AE Hassanien et al. (Eds.). Applications of Intelligent Optimization in Biology and Medicine: Current Trends and Open Problems; 2015. p. 159-180.

29. Jain S. Regression analysis on different mitogenic pathways. Network Biology. 2016 June; 6(2): 40-46.

30. Jain S. Mathematical Analysis using Frequency and Cumulative Distribution functions for Mitogenic Pathway. Research Journal of Pharmaceutical, Biological and Chemical Sciences. 2016 May- June; 7(3).