International Journal of Engineering & Technology, 7 (2.33) (2018) 1057-1063

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Extraction of Texture Features and Classification of Renal Masses from Kidney Images

Hima Bindu G[1], Prasad Reddy Pvgd[2], Ramakrishna Murty M[3]

[1]Research Scholar, Dept. of Computer Science and Systems Engineering, Andhra University, India
[2]Dept. of Computer Science and Systems Engineering, Andhra University, India
[3]Dept. of CSE, ANITS(A), Visakhapatnam, India
** Corresponding author E-mail: [1]hima.gopalam@gmail.com, [2]prasadreddy.vizag@gmail.com, [3]ramakrishna.malla@gmail.com*

Abstract

In this work, texture based feature analysis is done on renal masses obtained from kidney CT scan images to determine the type of kidney abnormality. From segmented renal masses, two sets of features are extracted. One set is extracted using first-order statistics(FOS) and second-order statistics(SOS). The second set is extracted using Discrete Wavelet Transformation (DWT). PCA is used as dimensionality reduction technique on DWT feature vector. The classification techniques, SVM Support Vector Machine and K Nearest Neighbors with different parameters (sub classifiers) are used individually on both the features sets. Classification models are trained on training datasets and the models are used for classification of renal masses from CT images. Performance of both the features sets and the classifiers on the CT kidney images are evaluated. It is observed that texture features extracted using FOS and SOS on the image produced better results than wavelet features and Cubic SVM sub classifier produced higher classification accuracy amongst all sub classifiers of SVM and KNN.

Keywords: *Classification; DWT transform; Feature Extraction; First-order statistics(FOS); Gray level co-occurrence matrix; K Nearest Neighbors; PCA, Second-order Statistics(SOS); Support Vector Machine, Texture Analysis*

1. Introduction

Kidney masses are associated with abnormality such as Cysts, Angiomyolipoma (AML) and RCC tumors and these do not give any symptoms most of the times. These masses are accidentally detected during a normal CT scan. In the recent years, kidney RCC tumors have been rated in the top pathologies causing increased death rate. However, when detected in early stages and with proper treatment could help the patient in recovering to normal. Computer aided diagnosis techniques can help in automatic detection and classification of kidney masses. Not much work has been done in the area of automatic renal mass detection and classification from CT scan images. In this paper, feature extraction, selection and classification of renal masses obtained from the CT scan images is carried out using suitable techniques. It is observed that from the literature, that majority of renal masses detected are either Cystic or RCC tumor. AML constitute less than 10% of the renal masses [22].

Features are distinguishing characteristics used for describing specific visual properties of an image. Texture is a core feature that is used in image processing to characterize the structure and surface of a given region and to quantitatively measure the pixel intensities. Texture measurement falls into two major groups: Statistical and Structural. Statistical methods describe texture by analyzing the spatial distribution of gray values, which is one of the important quality of texture. These methods compute local features for all points in the image and using these local features distribution, a set of statistics are derived. Statistical methods are further categorized into first-order statistics(FOS), second- order statistics(SOS) and higher-order statistics. Structural methods define texture by locating the structural primitives and analyzing structural relationships between these primitives. Wavelet transforms help in analyzing the structural primitives. The selected features are used in model construction for classification algorithms [1]. Supervised classification algorithms follows two phases of processing: training and testing phase. In training stage, characteristic properties from selected features are isolated and a training class is created based on the unique description of all classification categories. In testing phase, the identified feature-space partitions are utilised to classify features of the new image data [23].

In this work, features are extracted using both the statistical and structural methods. In statistical feature extraction methods, features are extracted using first-order statistics(FOS) and second-order statistics(SOS). In structural approach, features are extracted using two-dimensional Discrete Wavelet Transform (DWT). Classification is performed using K-NN and SVM algorithms individually on the two features sets. Experiments are conducted and a comparative analysis of feature extraction methods (statistical and structural) using K-NN and SVM classification algorithms is performed.

This paper is organised into eight sections: section I is introduction, Section II details the literature study related to the current work, Section III to VI describes the methodology, statistical and structural feature extraction, feature selection, K-NN and SVM classifiers. Experimental study performed is described in section VII and conclusion remarks is detailed in section VIII.

Copyright © 2018Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2. Related Literature Study

Feature extraction and selection is a core of any Classification algorithms. The accuracy of classification depends both on the selection of features and the efficiency of the classification algorithm itself. Thus it is very important to understand the feature extraction, feature selection and classification algorithms in computer aided detection systems. Mohanaiah P. et al [8] describe a paper for extraction of statistical features using GLCM approach. The RGB images are first converted to gray image and the GLCM is computed. The SOS features namely entropy, correlation, inverse difference moment and angular second moment are computed for different image sizes (64x64, 128x128 and 256x256). It is observed that textural features increased proportionately with the image size and author recommends using 128x128 size for feature extraction so that the loss of information is minimum. Selvarajah, S et al [1], have studied both statistical and structural based texture features in images. They have evaluated the features individually and in combination and based on the experiment results conclude that i) performance is dependent on spatial distribution of the images and GLCM is better in images which are homogenous. ii)In comparative study, texture features based on structure performed better than structural features. iii) Statistical and structure based features when used in combination performed better than individual features. Shijin Kumar P.S et al [2] have used texture feature extraction from GLCM method and shape features from connected regions for classification of tumor in brain MRIs. The shape features are directly extracted from the segmented image. The features are further used for training the classification model. Pauline John [3] in his paper on classification of brain tumour to benign and malignant from MRI has used DWT with Daubechies (db4) wavelet for GLCM formation. Texture features (contrast, correlation, energy, entropy and homogeneity) are then obtained and fed to probabilistic-neural-network (PNN) classifier. The classification accuracy has been noticed to be approximately 100%. Swapnali S et al [4] in their paper on classification of brain tumor has used DWT for extracting features using GLCM. Feature selection was performed utilizing PCA, which is common dimensionality reduction method. The classification is done using PNN method, which is a multi-layer feed-forward network. Y. Zhang et al [5] in his paper on classification of tumor has used 3-level DWT using Haar wavelet for feature extraction and PCA for reducing the dimensions. Kernel SVM (support vector machine) with different Kernels (LIN, IPOL, HPOL, GRB) are experimented and it is found that GRB kernel yielded better accuracy compared to other kernels in classification of abnormal and normal MRI brain images. Yang F et al [6] in study of X-ray images with Esophageal cancer have evaluated features obtained from frequency, textural and complexity domains. A total of 37 features were obtained and the authors used SFS (Sequential forward selection) algorithm and PCA for feature selection. A two stage classification was performed using KNN and SVM methods using all the features with no selection algorithm, SFS and PCA methods. In the comparative analysis using blend of feature selection methods and classification methods, it is noticed that SVM classification using SFS feature selection method performed better than other combinations of classification and feature selection methods. Shailaja R et al [7] have proposed 2D-GWT (Gabor wavelet transform) for extraction of texture features of kidney lesions from the US images. 21 wavelets are obtained from seven orientations and three different scales. From the 21 wavelets, 21 feature images are obtained for the considered ROI. Two statistical features are then computed for each of the feature image, resulting in 42 feature images. The features are fed to the SVM classifier to study the overall classification accuracy for two classes, AML and RCC. Experimental study yielded an overall classification accuracy of 90%. Mariam et al [9] in their study for kidney abnormalities from US images have made a comparative study of features obtained from 1) GLCM of the ROI image (statistical features) and 2) DWT (wavelet based features). ANN classification was used on the training dataset and results with wavelet based features produced higher accuracy compared to the statistical features obtained directly from the ROI of the image.

Not much work is done in the area on detection and classification of kidney masses from CT images. Hence in the current work, suitable feature extraction and selection techniques, classification techniques have been experimented on the CT kidney dataset.

3. Methodology

Renal masses are obtained from kidney CT image using segmentation techniques. The Kidney CT image are first pre-processed for kidney segmentation and the renal mass is segmented using region growing method. The segmented renal mass is used for extracting features and classification. The proposed methodology consists of four main phases. Feature extraction from segmented renal masses, feature selection, classification and performance of classification.



Fig 1: Classification of Renal Mass Workflow

In the current work, features are extracted based on the below two approaches

I. Statistical approach: first-order statistical(FOS) features and second-order statistical(SOS) features are derived from the image

II. Structural approach: Two-dimensional discrete wavelet transform (DWT) is applied on the image and wavelet features are extracted.

PCA is applied to decrease the dimensionality of the feature set obtained from DWT. Finally supervised classification algorithms are trained and applied on the selected feature set. A comparative analysis is performed on the both the features sets obtained from statistical and structural approaches. A second comparative analysis is also performed on K-NN and SVM sub classifiers to select the best classifier for the kidney CT image dataset.

4. Feature Extraction

4.1Statistical Method

Statistical methods compute local features by analysing the spatial distribution of gray levels at all points of the image. Based on the distribution of local features, a collection of statistics is computed. Statistical methods are categorized further into n^{th} order (n pixels) statistics based on the number of pixels that describe the local features [20].

4.2 First Order Statistics (FOS):

These texture measures are computed directly from the pixel values of the original image like mean, variance etc. Pixel neighborhood relationships are not considered in this approach. Histogram of an image represents the count of pixels for each gray level intensity value in the image. The intensity level histogram gives a simple and concise statistical information summary from the image. Gray-level histogram calculation includes single pixels and thus FOS can be obtained from image histogram. The histogram shape also provides additional characteristics of an image. For example, narrowly distributed histogram suggests low contrast image, bi-modal histogram indicates existence of an object with narrow intensity values against a different intensity background etc. Various FOS features can be derived from the histogram. Common FOS features include mean, variance, kurtosis, skewness,

energy, entropy. The FOS are defined by the following equations [21].
First order histogram for a region is expressed as

$$P(j) = \frac{count\ of\ pixels\ with\ gray\ level\ value\ j}{total\ count\ of\ pixels} \quad (1)$$

where j is a variable which represents image gray levels and L is the distinct count of gray levels.

$$\text{Mean } (\mu) = \sum_{j=0}^{L-1} j\,P(j) \quad (2)$$

$$\text{Variance } (\sigma^2) = \sum_{j=0}^{L-1}(i-\mu)^2\,P(j) \quad (3)$$

$$\text{Standard deviation } (\sigma) = \sqrt{\sum_{j=0}^{L-1}(j-\mu)^2\,P(j)} \quad (4)$$

$$\text{Skewness} = \sigma^{-3}\sum_{j=0}^{L-1}(j-\mu)^3\,P(j) \quad (5)$$

$$\text{Kurtosis} = \sigma^{-4}\sum_{j=0}^{L-1}(j-\mu)^4\,P(j) - 3 \quad (6)$$

Other parameters used are

$$\text{Smoothness: R} = 1 - \frac{1}{1+\sigma^2} \quad (7)$$

$$\text{Root Mean Square, RMS} = \frac{1}{N}\sqrt{\sum_{n=1}^{N}|x_n|^2} \quad (8)$$

where x is a vector of N dimensions.

$$\text{L1-norm(Least Absolute Deviations)} = \sum_{i=1}^{N}|y_i - f(x_i)| \quad (9)$$

$$\text{L2-norm (Least Squares)} = \sum_{i=1}^{N}\left(y_i - f(x_i)\right)^2 \quad (10)$$

Where $f(x_i)$, y_i are the estimated and target values respectively.

4.3 Second Order Statistics (SOS):

GLCM (Gray-Level Co-occurrence Matrix) based calculations come under the second-order statistics (SOS) category [1]. GLCM was initially presented by Haralick et al [13]. This is recognized method to obtain statistical texture features by considering the spatial relationship among the pixels of the image. The GLCM characterizes image texture by computing the frequency of occurrence of pixel pairs in a defined spatial relationship and computes measures SOS of image. The size of GLCM is defined by the gray levels present in the image. GLCM is a 2-D matrix where (x,y) element is defined as the frequency of x that occurs with y and is a function of distance and angle (d,Θ). The angle is considered along the four directions - horizontal, vertical and the two diagonals (0, 45, 90, 135). The GLCM matrix G(d,Θ) is represented as

$$G(d,\Theta) = [P(x,y) / d,\Theta)];\ 0<x, y <= L \quad (11)$$

L = no. of distinct gray levels in the image. Four GLCMs are computed for each direction (Θ = 0, 45, 90, 135) and then the average of the GLCMs is considered. From the GLCM matrix, features are extracted. Each feature represents the visually recognizable property of the image. These features include contrast, correlation, entropy, energy, homogeneity, sum of squares (variance), sum variance, sum entropy, sum average, difference variance, difference entropy, maximum correlation coefficient and correlation I & II information measures.

Energy, homogeneity, entropy, contrast, and correlation statistics are computed from GLCM and defined by the following equations [13].

$$\text{Energy} = \sum_i\sum_j\{P(i,j)\}^2 \quad (12)$$

$$\text{Entropy} = -\sum_i\sum_j P(i,j)\log(P(i,j)) \quad (13)$$

$$\text{Contrast} = \sum_{n=0}^{L-1} n^2\left\{\sum_{i=1}^{L}\sum_{j=1}^{L} P(i,j)\right\} \quad (14)$$

$$\text{Homogeneity} = \sum_i\sum_j \frac{1}{1+(i-j)^2}P(i,j) \quad (15)$$

$$\text{Correlation} = \frac{\sum_i\sum_j(i,j)P(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y} \quad (16)$$

Where μ_x,μ_y and σ_x, σ_y are the mean and variance of P respectively.

5. Discrete Wavelet Transformation (Dwt):

Wavelets are useful functions that describe both the temporal and localized frequency details of a pixels in an image and help in extracting the most significant pixels. Hence wavelets have become a very effective mathematical tool in feature extraction. Fourier transform has a disadvantage as it deals with just the frequency component and not the temporal details. Two-dimensional discrete wavelet transform (DWT) is used in this paper for feature extraction from CT images. This is a multi-resolution and multi-scale analysis method that analyses the given signal by decomposing into approximate and detailed coefficients by applying the low-pass and high-pass filters on rows and columns successively [12]. This method helps in reducing the size of wavelet coefficient and produces a feature vectors of much smaller size. This helps in further reducing the complexity of classification algorithms.

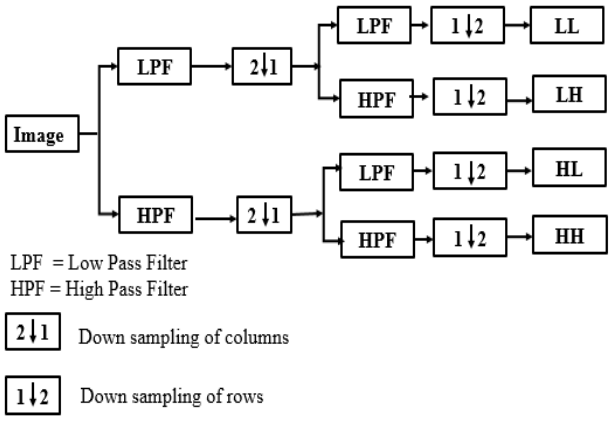


Fig 2:2D Discrete Wavelet Transform

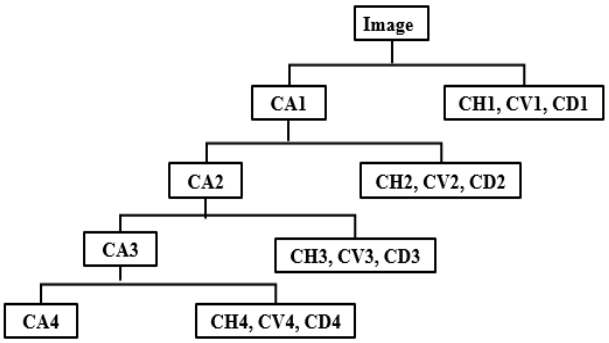


Fig 3: DWT decomposition – Four Level

DWT extracts wavelet coefficient from CT images at multiple levels and scales. Wavelet decomposition produces four frequency subbands: low-low (LL), low-high (LH), high-low (HL), high-high (HH) at each level. LL (low frequency) is the approximation component (CA) of the image and others are high frequency subbands. LH, HL and HH are the detailed components along the three directions and denoted as CH (horizontal), CV (vertical), CD (diagonal) respectively [5]. The mother wavelet utilised in DWT influences the resultant wavelet characteristics. Daubechies (db4) is used in our proposed approach. The approximation component

can further be decomposed to produce next level of approximation and detailed components. As CA is decomposed to next level, it will be further distributed to four components and each component will have M/2xM/2 coefficients. In the following level decomposition again, the components will have M/4xM/4 coefficients and the process will repeat at every next level [10]. This wavelet decomposition is continued until the required level is obtained. The detailed components at the last level is retained and all previous levels are discarded as they contain too much of noise and very less useful information. Different features can be extracted from coefficients obtained at the required level.

6. Feature Selection

Principal Component Analysis (PCA) is a robust statisticaltool which is used for interpreting the data. It finds patterns in data and then reduces the dimensions without much information loss. PCA in conjunction with wavelet transform has also been used in areas of medical images [5] and face recognition [14] for dimensionality reduction. In this paper, PCA is used to decrease the dimensions of wavelet coefficients obtained from DWT of renal mass CT images. This helps in better performance of classification algorithms. Dimensionality reduction using PCA involves the following steps 1) get the data (wavelet coefficients) 2) normalize the data by deducting mean from the individual data dimensions. 3) compute the co-variance matrix 4) compute the eigen vectors and values for the co-variance matrix 5) Sort eigen vectors by eigen values (highest to lowest) 6) select top eigen vectors and discard the low significant eigen vectors. 7) Feature vector is resultant of the selected top eigen vectors.

7. Classification

Classification methods have evolved from pattern recognition and is a process of classifying data into definite set of classes by the virtue of relationship of data to classes. Classification is broadly classified into two approaches – supervised approach and unsupervised approach. Unsupervised classification does not need prior information and it uses clustering techniques to determine related pixels and group them into classes. Analyst uses a posteriori information obtained to label relevant classes of interest. Some of the common unsupervised classification methods are K means clustering and ISODATA clustering methods. Supervised classification uses learning process from the trained data. The trained data consists the input data and the class labels from which a classification model is built. The model is then tested on the new data to obtain predicted responses. Commonly used supervised classifiers are Decision tress, Discriminant Analysis, Naïve Bayes, nearest neighbors

In this paper, KNN and SVM classification techniques are evaluated to check the best suited classifier for renal mass data obtained from the kidney CT images.

7.1 K Nearestneighbors (KNN)

The KNN algorithm is one of the simplest and widely used classification algorithms. It is first proposed in 1968 by Cover and Hart. It is suitable technique and can assign weights to the neighbors such that the contributions of the nearer neighbors are significantly high to the average compared to the farthest ones. KNN is a non-parametric classifier as it does not make any assumptions about the structure of the data. In K-NN classification, the neighbors are selected from an object set (considered as training set) for which the class is known. Explicit training step is not required here. KNN stores the complete training dataset and makes the predictions based on the training data. KNN makes just in time predictions by computing the similarity of the input data to each of training data samples. Predictions for the input sample is made by searching the complete training data set for the K best similar neighbors using distance measurea. Various distance measures that can be used with KNN are Euclidean, Hamming, Manhattan, Tanimoto, Mahalanobis, Jaccard, Minkowski(cubic) and cosine distance [18]. Minkowski distance d(u,v) between two points u and v is defined as [17]

$$d(u,v) = \left(\sum_{i=1}^{n}(v_i - u_i)^a\right)^{1/a} \quad (17)$$

Minkowski distance is called Manhattan and Euclidean distances for a=1 and=2 respectively.a=3 for Minkowski (cubic) distance.

$$\text{Cosine distance} = 1 - \text{Similarity} \quad (18)$$

$$\text{Similarity } (\cos(\Theta)) = \frac{u \cdot v}{\|u\|\|v\|} \quad (19)$$

Depending on the number of nearest neighbors, distance metric and distance weights, different variants of K-NN sub classifiers are formulated. Fine, Medium and coarse KNN subclassfiers use different number of neighbors with coarse considering large set of neighbors. Inverse or sqaured inverse distance weights can be used with weighted KNN.

The algorithm for weighted KNN is as follows [24]

For a training dataset $T_s = \{(v_i, u_i), i = 1,2,...,n\}$, let U_i represent the observations and $V_i$represent the class membership. For a new observation (u), the class label (v) is to be predicted from the class membership.
Using the distance function $d(u, u_i)$, $k+1$ nearest neighbors of x are computed
$(k+1)^{th}$neighbor is applied for standardizing the k smallest distances using the formula

$$D_{(i)} = D(u, u_i) = d(u, u_i)/d(u, u_{k+1}) \quad (20)$$

The normalised distance $D_{(i)}$is transformed into weights W(i) = K $(D_{(i)})$ using kernel function K(.)
The prediction of class label y is choosen from a class of maximum weights of K nearest neighbours.

A good K value can be choosen by cross-validation method; by executing the nearest neighbor(NN) classifier on the training data set only. Larger value of K reduces noise effect on classification but also reduces the distinction of boundaries between classes. [6]. We can determine that the new instance of test data belongs to a certain class when it has more representatives compared to other classes or by computing the average class numbers of the K nearest neighbors. KNN is widely used in pattern-recognition for to its simplicity in implementation and eff ectiveness of classification.

7.2 Support Vector Machine (SVM)

SVM is first presented by Boser et al in 1992 and is a robust supervised machine learning algorithm. SVM is used in classification and is built using statistical learning theory. The simplest method to separate dataset with two classes is to find an optimal hyperplane with maximum margin. For non-linearly separable datasets, the data is mapped to higher dimensional space using kernel function so that linear surfaces can be used to classify the data easily [6,19]. This is also called kernel trick. The margin distance between the two-classes represents the maximum distance between nearest data points of the selected classes. Larger margins result in lower generalization error of the classification algorithms.

The kernel function maps data to new space and takes the inner dot product of the new vector space. The kernel function is symbolized as [19]

$$K(U_i,U_j) = \varphi(U_i).\ \varphi(U_j) \quad (21)$$

Three Kernel functions [19] are shown below:

1. Linear Kernel: $K(U_i,U_j) = (U_i . U_j)$ *(22)*

2. Polynomial Kernel: $K(U_i,U_j) = (1 + (U_i . U_j))^d$ *(23)*

d =2 and =3 for quadratic and cubic Kernel functions respectively

3. Gaussian Radial Basis Kernel:

$$K(U_i,U_j) = \exp\left(-\frac{\|u_i-u_j\|^2}{2\sigma^2}\right) \quad (24)$$

Depending on the value of kernel scale (σ), the classifier is defined as low, medium and coarse with σ taking values from low to high respectively.

7.3 Classification Metrics

The performance of classification is measured using the following three metrics: Sensitivity, Specificity and Accuracy. Sensitivity is a measure of true positives compared to the total predicted positives. Mathematically it is expressed as [15,16]

$$\text{Sensitivity} = \frac{True\ Positives}{True\ Positves + False\ Negatives} \quad (25)$$

Specificity is a measure of true negatives compared to the total predicted negatives. Mathematically it is expressed as

$$\text{Specificity} = \frac{True\ Negative}{True\ Negatives + False\ Positives} \quad (26)$$

Accuracy is a measure of correctly classified objects from the total objects.

$$\text{Accuracy} = \frac{Correctly\ classified\ cases}{Total\ cases} \quad (27)$$

Sensitivity is a measure of how well the positives are classified, while specificity indicates how well the negatives are classified. Accuracy determines the performance of the overall classification. For a good classification, all the three values – Sensitivity, Specificity and accuracy should be high.

8. Experimental Result

Dataset: CT images containing renal mass are obtained from medical hospitals and from medical databases. Dataset containing four classes is considered: Cystic, AML, RCC Tumor and Normal. Implementation of feature extraction, selection and classification is performed using Matlab.

Features are extracted using

1. First-order statistics(FOS) and second-order statistics(SOS) as described in section 4.1.
2. 2D-DWT. Texture features of the 4th level wavelet decomposition is performed using daubechies wavelet.
3. The texture features obtained, form the input vectors for training and test phases.

For the training phase, a database with extracted features was compiled consisting of renal mass (Cystic, AML, RCC) and without any renal mass (from a normal kidney). All K-NN and SVM classifiers are trained using the training dataset. The test data is then run against the trained models (both K-NN and SVM) individually.

K-NN Classification is executed using different values for number of nearest neighbors, distance metric and distance weight as given in the below table.

Table 1: K-NN Sub Classifiers with parameter values

K-NN Sub Classifiers						
Parameters	**Fine**	**Medium**	**Coarse**	**Cosine**	**Cubic**	**Weighted**
No. of nearest neighbors	1	10	100	10	10	10
Distance Metric	Euclidean	Euclidean	Euclidean	Cosine	Minkowski (Cubic)	Euclidean
Distance weight	No weight	No weight	No weight	No weight	No weight	Squared Inverse

SVM Classification is executed using different kernel functions and scale as given in the below table.

Table 2: SVM Sub Classifiers with parameter values

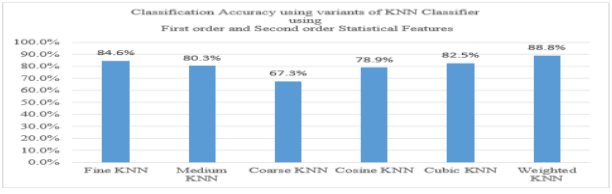
SVM Sub Classifiers						
Parameters	**Linear**	**Quadratic**	**Cubic**	**Fine Gaussian**	**Medium Gaussian**	**Coarse Gaussian**
Kernel Function	Linear	Quadratic	Cubic	Gaussian	Gaussian	Gaussian
Kernel Scale	Automatic	Automatic	Automatic	0.9	3.6	14

The below table shows the classification accuracy obtained with different variants of SVM and K-NN classifier using statistical features and wavelet features.

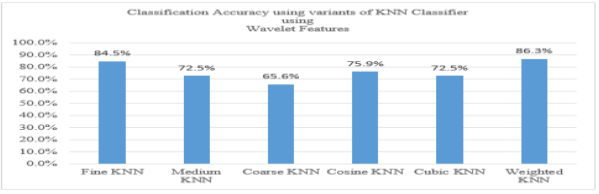
Table 3: Accuracy of KNN Sub Classifiers and SVM Classifiers on features sets

K-NN Sub Classifiers	Statistical Features (FOS and SOS)	Wavelet Features (DWT)
Fine KNN	84.6%	84.5%
Medium KNN	80.3%	72.5%
Coarse KNN	67.3%	65.6%
Cosine KNN	78.9%	75.9%
Cubic KNN	82.5%	72.5%
Weighted KNN	88.8%	86.3%

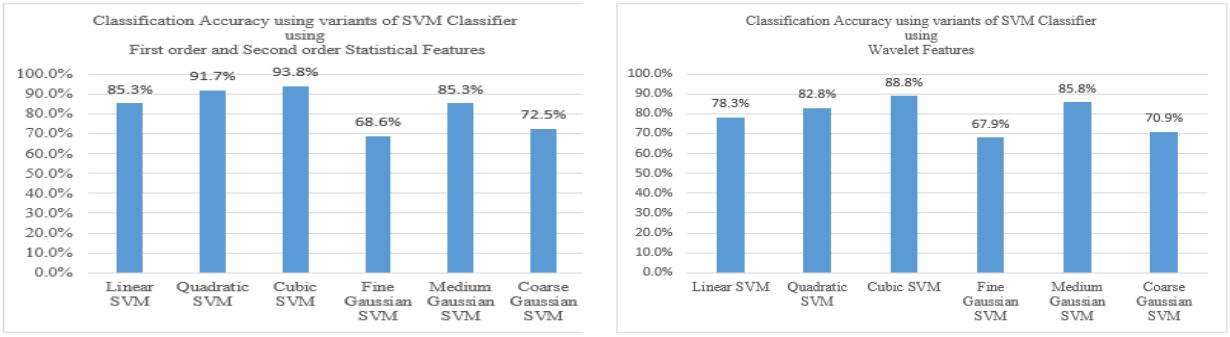
SVM Sub Classifiers	Statistical Features (FOS and SOS)	Wavelet Features (DWT)
Linear SVM	85.3%	78.3%
Quadratic SVM	91.7%	82.8%
Cubic SVM	93.8%	88.8%
Fine Gaussian SVM	68.6%	67.9%
Medium Gaussian SVM	85.3%	85.8%
Coarse Gaussian SVM	72.5%	70.9%



(a)



(b)



(c) **(d)**

Fig 4: Classification Accuracy of a) KNN sub classifiers on Statistical features b) KNN sub classifiers on wavelet features c) SVM sub classifiers on Statistical features d) SVM sub classifiers on wavelet features

In all the sub classifiers of

- K-NN, Weighted KNN has produced classification results with higher accuracy for both the statistical and wavelet feature set.
- SVM classifier, Cubic SVM has produced classification results with higher accuracy for both the statistical and wavelet feature set.

Detailed results of Cubic SVM and Weighted KNN for the different feature sets are shown in table 4 and table 5 below along with graphical representation in Fig 5.

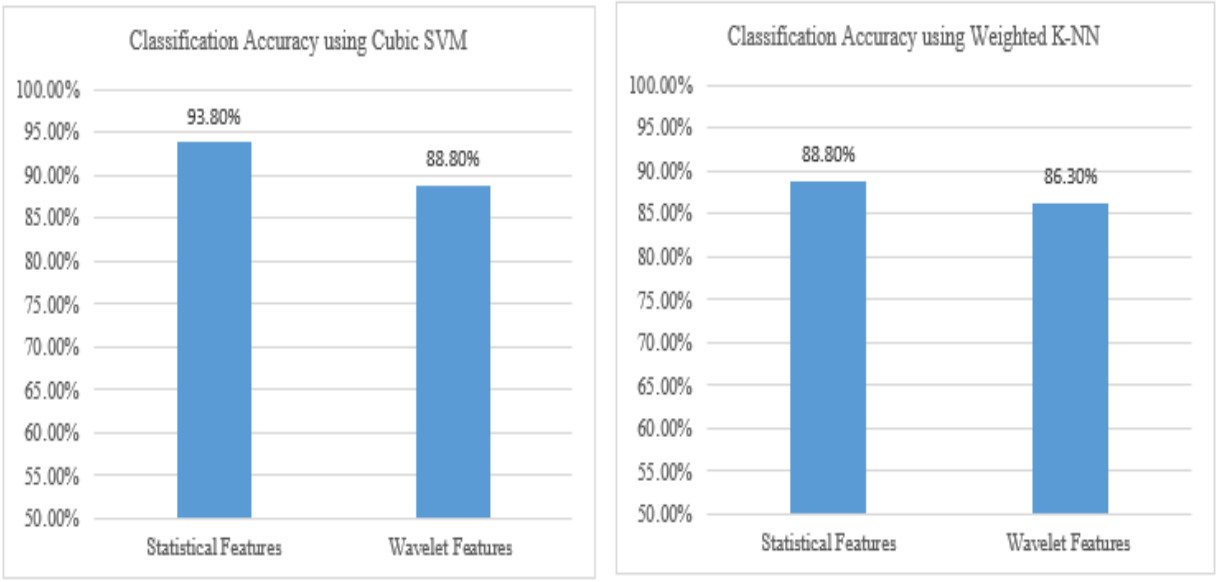
Sensitivity and Specificity for Cubic SVM for the two different feature set is calculated and graphical representation is shown in Fig 6 for comparison study.

Table 4: Comparison of Cubic SVM and Weighted K-NN classifier on wavelet features (DWT)

Image Class	Wavelet Features (DWT)			
	Cubic SVM Classification Rate = [71/80] (88.8%)		Weighted K-NN Classification Rate = [69/80] (86.3%)	
	Sensitivity	Specificity	Sensitivity	Specificity
CYST	[32/35] (91%)	[39/45] (87%)	[31/35] (89%)	[38/45] (84%)
NORMAL	[12/14] (86%)	[59/66] (89%)	[12/14] (86%)	[57/66] (86%)
AML	[14/14] (100%)	[57/66] (86%)	[14/14] (100%)	[55/66] (83%)
TUMOR	[13/17] (76%)	[58/63] (92%)	[12/17] (71%)	[55/63] (87%)

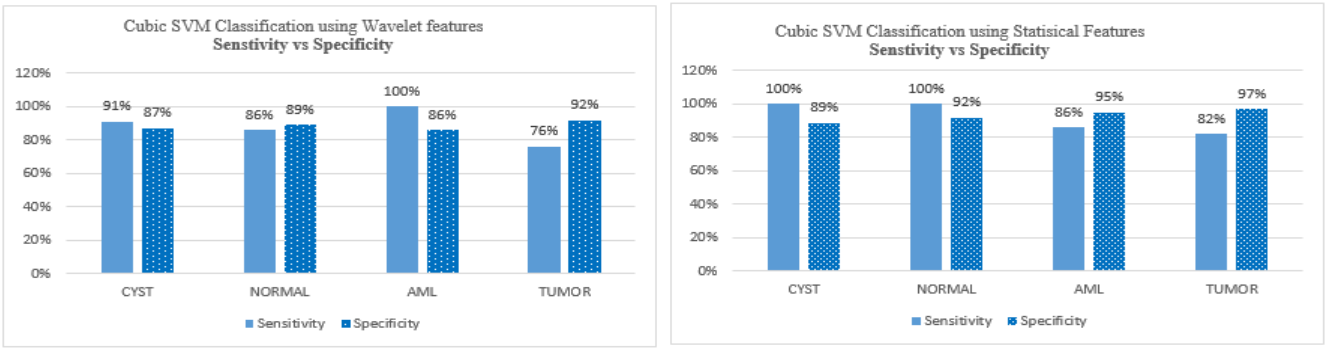
Table 5: Comparison of Cubic SVM and Weighted K-NN classifier on Statistical features

Image Class	Statistical features (FOS and SOS)			
	Cubic SVM Classification Rate = [75/80] (93.8%)		Weighted K-NN Classification Rate = [71/80] (88.8%)	
	Sensitivity	Specificity	Sensitivity	Specificity
CYST	[35/35] (100%)	[40/45] (89%)	[35/35] (100%)	[36/45] (80%)
NORMAL	[14/14] (100%)	[61/66] (92%)	[13/14] (93%)	[58/66] (88%)
AML	[12/14] (86%)	[63/66] (95%)	[10/14] (71%)	[61/66] (92%)
TUMOR	[14/17] (82%)	[61/63] (97%)	[13/17] (76%)	[58/63] (92%)



(a) (b)

Fig 5: Classification Accuracy of a) Cubic SVM b) Weighted KNN



(a) (b)

Fig 6: Comparison of Sensitivity and Specificity for a)wavelet features b)statistical features using Cubic SVM classifier

9. Conclusions and Future Work

In this paper, computer aided algorithm has been implemented for detection and classification of renal masses from CT images. Two sets of features are extracted from renal mass segmented from Kidney CT images. Features are extracted using Statistical (FOS and SOS) and DWT methods. PCA is used for reducing the dimension of matrix obtained from DWT. In training phase, various SVM and K-NN sub classifiers are trained using 50% of the available dataset with k=5 in cross validation method. The SVM and K-NN sub classifiers are then used to test the remaining 50% of test dataset. From the experiment, it is observed that Cubic SVM produced higher classification accuracy results amongst all the SVM sub classifiers and weighted KNN produced higher classification accuracy results amongst all the KNN sub classifiers. Further analysis of the experimental results shows Cubic SVM classifier with statistical features produced classification accuracy of 93.8% while weighted K-NN classifier resulted in accuracy of 88.8%. Both the classifiers are also run on wavelet features and results shows that 1) Statistical features (FOS and SOS) are better compared to wavelet features 2) The accuracy of SVM classifier is higher than K-NN classifier.

To optimize of the classification models, bio-inspired and its variants along with hybrid techniques can be applied in the future work.

9. References

[1] Selvarajah, S., and S. R. Kodituwakku. "Analysis and comparison of texture features for content based image retrieval." International Journal of Latest Trends in Computing 2, no. 1 (2011).

[2] Shijin Kumar P.S., and Dharun V. S. "Extraction of Texture Features using GLCM and Shape Features using Connected Regions." International Journal of Engineering and Technology (IJET), Vol 8 No 6 Dec 2016-Jan 2017

[3] John, Pauline. "Brain tumor classification using wavelet and texture based neural network." International Journal of Scientific & Engineering Research 3, no. 10 (2012): 1-7.

[4] Sawakare, Swapnali, and Dimple Chaudhari. "Classification of Brain Tumor Using Discrete Wavelet Transform, Principal Component Analysis and Probabilistic Neural Network." International Journal for Research In Emerging Science And Technology 1 (2014).

[5] Zhang, Yudong, and Lenan Wu. "An MR brain images classifier via principal component analysis and kernel support vector machine." Progress In Electromagnetics Research 130 (2012): 369-388.

[6] Yang, Fang, Murat Hamit, Chuan B. Yan, Juan Yao, Abdugheni Kutluk, Xi M. Kong, and Sui X. Zhang. "Feature Extraction and Classification on Esophageal X-Ray Images of Xinjiang Kazak Nationality." Journal of Healthcare Engineering 2017 (2017).

[7] Rana, Shailja, Shruti Jain, and Jitendra Virmani. "Classification of kidney lesions using Gabor wavelet texture features." In Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on, pp. 1705-1709. IEEE, 2016.

[8] Mohanaiah, P., P. Sathyanarayana, and L. GuruKumar. "Image texture feature extraction using GLCM approach." International Journal of Scientific and Research Publications 3, no. 5 (2013): 1.

[9] Attia, Mariam Wagih, F. E. Z. Abou-Chadi, Hossam El-Din Moustafa, and Nagham Mekky. "Classification of ultrasound kidney images using PCA and neural networks." International Journal of Advanced Computer Science and Applications 6, no. 4 (2015): 52-57.

[10] Ahmad, Mubashir. "Mahmood ul-Hassan, Imran Shafi, Abdelrahman Osman, Classification of Tumors in Human Brain MRI using Wavelet and Support Vector Machine." IOSR Journal of Computer Engineering 8, no. 2 (2012): 25-3.

[11] Yadav, Arvind R., R. S. Anand, M. L. Dewal, and Sangeeta Gupta. "Performance analysis of discrete wavelet transform based first-order statistical texture features for hardwood species classification." Procedia Computer Science 57 (2015): 214-221.

[12] Rajpoot, Kashif M., and Nasir M. Rajpoot. "Wavelets and support vector machines for texture classification." In Multitopic Conference, 2004. Proceedings of INMIC 2004. 8th International, pp. 328-333. IEEE, 2004.

[13] Haralick, Robert M., and Karthikeyan Shanmugam. "Textural features for image classification." IEEE Transactions on systems, man, and cybernetics 6 (1973): 610-621.

[14] Chitaliya, N.G. and Trivedi, A.I., 2010, March. Feature extraction using wavelet-pca and neural network for application of object classification & face recognition. In Computer Engineering and Applications (ICCEA), 2010 Second International Conference on (Vol. 1, pp. 510-514). IEEE.

[15] Devasena, C. Lakshmi, and M. Hemalatha. "Efficient computer aided diagnosis of abnormal parts detection in magnetic resonance images using hybrid abnormality detection algorithm." Central European Journal of Computer Science 3, no. 3 (2013): 117-128.

[16] Ramteke, R. J., and Y. Khachane Monali. "Automatic medical image classification and abnormality detection using K-Nearest Neighbour." International Journal of Advanced Computer Research 2, no. 4 (2012): 190-196.

[17] Hu, Li-Yu, Min-Wei Huang, Shih-Wen Ke, and Chih-Fong Tsai. "The distance function effect on k-nearest neighbor classification for medical datasets." SpringerPlus 5, no. 1 (2016): 1304.

[18] Chomboon, Kittipong, Pasapichi Chujai, Pongsakorn Teerarassammee, Kittisak Kerdprasop, and Nittaya Kerdprasop. "An empirical study of distance metrics for k-nearest neighbor algorithm." In The 3rd International Conference on Industrial Application Engineering 2015 (ICIAE2015). 2015.

[19] Jakkula, Vikramaditya. "Tutorial on support vector machine (svm)." School of EECS, Washington State University 37 (2006).

[20] Srinivasan, G. N., and G. Shobha. "Statistical texture analysis." In Proceedings of world academy of science, engineering and technology, vol. 36, pp. 1264-1269. 2008.

[21] Materka, Andrzej, and Michal Strzelecki. "Texture analysis methods–a review." Technical university of lodz, institute of electronics, COST B11 report, Brussels (1998): 9-11.

[22] http://www.radiologyassistant.nl/en/p571eea20ec282/kidney-solid-masses.html

[23] Xie, Xiaoyuan, Joshua WK Ho, Christian Murphy, Gail Kaiser, Baowen Xu, and Tsong Yueh Chen. "Testing and validating machine learning classifiers by metamorphic testing." Journal of Systems and Software 84, no. 4 (2011): 544-558.

[24] Hechenbichler, Klaus, and Klaus Schliep. "Weighted k-nearest-neighbor techniques and ordinal classification." (2004).

[25] Murty, M. Ramakrishna, J. V. R. Murthy, and Prasad Reddy PVGD. "Text Document Classification based-on Least Square Support Vector Machines with Singular Value Decomposition." International Journal of Computer Applications 27, no. 7 (2011).