Texture and Wavelet Based Classification of Thyroid Nodule in Ultrasound Image Using Kernel Based Support Vector Machines

[1]Wrushali M. Mendre, [2]Mousami V. Munot

Abstract : **The automated detection of thyroid nodule malignancy is a thrust area of research. Literature reports popular use of Fine Needle Aspiration Biopsy [FNAB] to confirm the nodule malignancy, however there is a dreadful need to improve the ability of predicting thyroid malignancy in USG images prior to FNAB. The proposed research aims to design and develop a computerized system to aid the decision making about a malignant thyroid. This study exploits Discrete Wavelet Transform [DWT] to devise a feature vector with seven textural features and additionally fifteen features are derived using Gray Level Co-occurrence Matrix [GLCM]. A Support Vector Machine (SVM) is used as an intelligent classifier to identify the malignant thyroid nodules using the fusion of the derived features. Standard database from Wilmington Endocrinology comprising of 79 ultrasound images of thyroid nodule (51 benign and 28 malignant cases) are considered to examine the efficacy of the proposed algorithm. The performance of the SVM classifier is evaluated considering polynomial, Radial Basis Function [RBF], linear and Multilayer Perceptron [MLP] kernel functions and analyzing the respective Receiver Operating Characteristics [ROC]. Highest classification accuracy of 99.98 with Area under ROC Curve (AUC) of 0.995, employing the polynomial kernel of 3rd degree confirms the superiority of the proposed approach.**

IndexTerms - Ultrasound, Textural Features, GLCM, DWT

I. INTRODUCTION

Thyroid gland is one of the largest endocrine gland and is located below the skin and muscles at the front of the neck [1]. It is one of the most important organisms of our body. As the thyroid hormones are responsible for large part of our body's metabolism, thyroid performance directly affects most of our organs. Therefore, fast and accurate diagnosis of thyroid diseases are of great importance [2].Thyroid nodules are abnormal lumps growing within the thyroid gland which may represent various different conditions including cancer. The incidence of thyroid cancer has increased about five folds in last fifty years. The American Cancer Society estimates for thyroid cancer in the United States for 2015 are about 62,450 new cases of thyroid cancer (47,230 in women, and 15,220 in men) and about 1,950 deaths from thyroid cancer (1,080 women and 870 men) [3,4].

Thyroid ultrasound is the major diagnostic modality for evaluating thyroid nodule. Ultrasound imaging has several advantages over other medical imaging modalities: noninvasive, relatively low cost, safe and allows real time imaging. A thyroid nodule in UltraSonoGraphy [USG] appears as a nodular lesion within the thyroid gland that is distinguishable from adjacent parenchyma. Thyroid nodule can be either benign or malignant nodule. Several ultrasound features such as marked hypoechogenicity, irregular margins, microcalcifications and a taller than wide shape have been introduced as potential predictors for the presence of thyroid malignancies[5].

Fine Needle Aspiration Biopsy [FNAB] is the sensitive method to detect thyroid malignancy [6]. It is a diagnostic procedure used to investigate superficial lumps or masses. In this technique, a hollow needle is inserted into the mass for sampling of cells that, after being stained, for examination under a microscope. It is a minor, invasive surgical procedure. The results of FNAB indicate one of the following findings: benign, non-diagnostic or malignant.

Though USG is a non-invasive and inexpensive imaging technique, it has some limitations. The reliability of diagnosis depends on quality of image and the expertise of the radiologist who interpret the images. USG images are easily affected by echo perturbations and speckle noise which interfere with the correct diagnosis. [7]. These limitations have given rise to extensive research in developing Computer Aided Diagnosis [CAD] Systems so as to obtain accurate diagnosis results using various image processing and pattern recognition methods.

Tsantis et al. [8] extracted morphological and wavelet features to detect thyroid nodule malignancy through USG images and reported the accuracy of 96%. Acharya et al. [9] proposed thyroid nodule classification algorithm using wavelet and textures. Three features are extracted using GLCM and 7 using DWT. Resulting feature vectors are fed to three different classifiers: K-Nearest Neighbour (KNN) and Probabilistic Neural Network [PNN]. Keramidas et al. [10] presented thyroid nodule detection algorithm in which noise resilient features are extracted from thyroid gland tissue and used for analysis of thyroid nodule. Nikita Singh et al. [11] proposed segmentation and classification method for classification of thyroid nodule through USG images.

Zhao et al. [12] proposed segmentation algorithm which combined homomorphic filtering, anisotropic diffusion and fractional differential into the normalization process. This segmentation method could extract nodules and trachea of the thyroid ultrasound image. The edge of this segmentation provided the position of piercing and assisting FNAB to complete discrimination that whether the thyroid nodule is benign or malignant.

Tsantis et al. [8] reported 96% accuracy but number of support vectors ranges from 8% to 10% of the number of training points. It is a indicative of comparatively higher complexity. Keramidas et al. [10] could only differentiate between normal thyroid tissue and thyroid nodule tissue. Zhao et al. [12] proposed segmentation method to assist FNAB. It is still a long standing problem which needs further research in the diagnosis of thyroid using USG images before the system can be practically deployed in clinical lab. In this paper, a texture based feature extraction and classification method for the differential diagnosis of malignant and benign thyroid nodule is proposed. There still exist a scope for further experimentation to determine the optimal feature extraction and classification methods .Enhancement of the functionality of proposed algorithms is needed to improve the classification accuracy

by considering the characteristics of malignant thyroid nodule. Evolvement of less complex classifier is also needed for applications in an integrated real time system for the assessment of thyroid gland. As far as correct diagnosis through USG image is concerned, accuracy of classification is a very important parameter which needs to be addressed.

Next section presents material and methods for USG image pre-processing, Region of Interest [ROI] separation, feature extraction approach from thyroid nodule and classification approach with corresponding procedures. Later, the experimental results obtained by the proposed method are discussed and finally the conclusions are presented.

II. PROPOSED METHODOLOGY:

CAD systems can be designed by using feature extraction and classification subsystems. The proposed research is divided into five steps, namely: pre-processing, ROI separation, feature extraction and classification. The first step involves the removal of speckle noise and makes the USG image suitable for further processing. The second step involves the manual separation of ROI that includes only nodular region. The third step includes various feature extraction methods where textural and wavelet information is encoded using feature fusion. In the forth step, classification algorithm is implemented to categorize the nodule as either malignant or benign. Details of the proposed system are described in detail in following sub- sections.

Publically available database containing thyroid nodule USG images from open source [13] are used in this research. It has 51 benign and 28 malignant cases. The images are in JPEG format with 540x410 size, 24bit depth, and true color image.

2.1 Pre-processing:

USG images are often affected by the occurrence of speckle noise which makes the fine details of the image unclear and degrades the quality of the image obtained. Hence filtering the speckle noise becomes important in ultrasonography images. It is an essential step in the preprocessing of medical ultrasonography image. There are various techniques used for the removal of speckle noise. For selecting a best filter, comparison of three techniques namely Frost filter, spatiotemporal filter and SRAD filter is done based on their performance in reducing the speckle noise from the USG images. Among these, Speckle Reducing Anisotropic Diffusion Filter [SRAD] filter gives better performance over the other two methods in terms of mean square error, Peak Signal to Noise Ratio [PSNR] and is therefore chosen in the proposed research [14].

2.2 Feature Extraction:

In this study, texture based features are extracted from USG image of nodular thyroid. The extracted features are fed to SVM classifier. The database of 79 USG images includes 51 benign and 28 malignant images. ROI is separated from the USG image. Total fifteen textural features are extracted from ROI using Gray Level Co-occurrence Matrix (GLCM) and 7 features are extracted using Discrete Wavelet Transform (DWT). The fusion of the features is done to obtain feature vector of 22 features.

The GLCM method is a way of extracting second order statistical features. It is a matrix where the number of rows and columns is equal to the number of gray levels in the image. The matrix element p (i, j | Δx, Δy) is the relative frequency with which two pixels separated by a pixel distance (Δx, Δy), occur within a given neighborhood, with intensity i and j respectively [15]. A number of texture features can be extracted from GLCM [16].

Given an MxN neighborhood of an image containing G gray levels from 0 to G-1, let f(m,n) be the intensity at sample m, line n of the neighborhood. Then its GLCM is given by equation (i), (ii),(iii),(iv) [17].

$$p(i,j\,|\,\Delta x,\Delta y) = WQ(i,j|\Delta x,\Delta y) \quad (i)$$

$$W = \frac{1}{(M-\Delta x)(N-\Delta y)} \text{ (ii)}$$

$$Q(i,j\,|\,\Delta x,\Delta y) = \sum_{n=1}^{N-\Delta y}\sum_{m=1}^{M-\Delta x} A \quad (iii)$$

$$A = \begin{cases} 1 \; if \; f(m,n) = i \; and \; f(m+\Delta x, n+\Delta y) = j \\ 0 \qquad\qquad\qquad\qquad elsewhere \end{cases} \quad (iv)$$

Fifteen textural features are extracted from GLCM. These features are: energy, entropy, dissimilarity, homogeneity, autocorrelation, contrast, correlation, cluster prominence, cluster shade, maximum probability, mean, variance, standard deviation, kurtosis and skewness.

Energy shows how gray levels are distributed. It returns the sum of squared elements in GLCM. Energy is described by equation (v). Heterogeneous nodules have low entropy, while homogeneous nodules have high entropy. Equation (vi) represents the entropy. Dissimilarity is calculated using equation (vii). Local Homogeneity measures the closeness of the distribution of the elements in the GLCM to the diagonal GLCM. Homogeneity is calculated using equation (viii). Autocorrelation and contrast are given by equation (ix) and (x) respectively. Correlation returns a measure of how correlated a pixel is to its neighbor over the whole image. Equation (xi) represents correlation. Cluster Prominence is a measure of asymmetry whereas cluster shade is the measure of skewness of GLCM matrix. When cluster prominence is low, there is a peak in the GLCM around the mean. Cluster prominence and cluster shade are given by equation (xii) and equation (xiii) respectively. Maximum Probability records the maximum P value in GLCM matrix. Mean (μ) is a measure of mean value of P. Variance is the grey level variability of the pixel pairs and is a measurement of heterogeneity. Variance increases when grey scale values differ from their means. Variance is given by equation (xiv). Standard deviation (σ) is a measure used to quantify the amount of variation or dispersion of a set of data values. Standard Deviation close to 0 indicates that the data points tend to be very close to the mean of the set, while high standard deviation indicates that the data points are spread out. Square Root of variance gives the standard deviation. Skewness is the

measure of asymmetry around the mean. The skewness is defined with equation (xv). Kurtosis, K measures the peakness or flatness of a distribution relative to a normal distribution. It is given by equation (xvi).

$$\sum\sum |P(i,j)|^2 \quad (v)$$

$$-\sum\sum [P(i,j)\log[P(i,j)] \quad (vi)$$

$$\Sigma\Sigma([(i-j)P(i,j)]) \quad (vii)$$

$$\sum\sum [(j-uy)^2 * P(i,j)] \quad (viii)$$

$$\sum\sum [(i,j)P(i,j)] \quad (ix)$$

$$\sum\sum |i-j|^2[P(i,j)] \quad (x)$$

$$\Sigma\Sigma[(i-u_x)(j-u_y)P(i,j)\,]/(\sigma_x\sigma_y) \quad (xi)$$

$$\sum\sum \left(i+j-u_x-u_y\right)^4 P(i,j) \quad (xii)$$

$$\Sigma\Sigma(i+j-u_x-u_y)^3 P(i,j) \quad (xiii)$$

$$\Sigma\Sigma(\mathrm{i}-\mu)^2\mathrm{p}(\mathrm{i},\mathrm{j}) \quad (xiv)$$

$$s = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{p(i,j)-\mu}{\sigma}\right)^3 \quad (xv)$$

$$k = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{p(i,j)-\mu}{\sigma}\right)^4 - 3 \quad (xvi)$$

DWT is a useful tool for many image processing applications. It is a linear transformation that operates on a data vector whose length is an integer power of two, transforming it into a numerically different vector of the same length. It is a tool that separates data into different frequency components. It uses finite impulse response filter banks [9]. After applying DWT, four coefficients named LL, HL, LH and HH are derived. LL component contains total energy of the thyroid nodule texture. HL, LH and HH component contains the vertical, horizontal and diagonal orientation of the thyroid nodule texture respectively. Fig. 1 (a) shows the first level of DWT decomposition. Decomposition is further performed on the LL coefficient to get coarser scale of wavelet coefficients. Decomposition is stopped at second level because further decomposition didn't give significant results. Fig.1 (b) shows second level of DWT decomposition. The individual sub- bands are represented as matrixes of size m*n. Theses matrixes are combined to form a feature. The Normalized energies A2 , V2, H2, D2, H1, V1, and D1 are obtained from matrixes LL2, HL2, LH2, HH2, LH, HL and HH respectively .Equations 17 to 23 shows the formation of normalized energies from individual sub-band matrices.

$$A_2 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} LL2_{ij}\right]^2 \quad \text{(xvii)}$$

$$V_2 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} HL2_{ij}\right]^2 \quad \text{(xviii)}$$

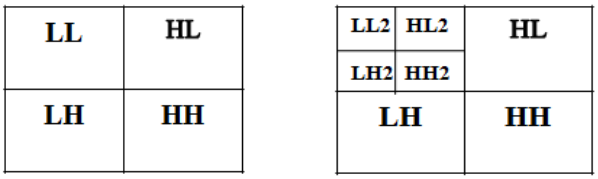
$$H_2 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} LH2_{ij}\right]^2 \quad \text{(xix)}$$

$$D_2 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} HH2_{ij}\right]^2 \quad \text{(xx)}$$

$$H_1 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} LH_{ij}\right]^2 \quad \text{(xxi)}$$

$$V_1 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} HL_{ij}\right]^2 \quad \text{(xxii)}$$

$$D_1 = \frac{1}{m*n}\sum_{i=1}^{m}\left[\sum_{j=1}^{n} HH_{ij}\right]^2 \quad \text{(xxiii)}$$



(a) (b)

Figure 1: 2D sub-band DWT (a) First Level of Decomposition (b) Second Level Decomposition

2.3 Classification:

Features play an important role in the design of a classifier. The fusion of textural and DWT feature vectors extracted from sub-images sampled from the parenchyma of the thyroid gland, are subsequently classified into a predefined set of classes. All 22 features at hand are employed for classification. The output of the classifier is either normal or cancerous thyroid. For the classification phase of the proposed scheme, SVM [18] has been evaluated.

SVM is a widely accepted and effective classifier for pattern recognition, machine learning and data mining. It is based on guaranteed risk bounds of statistical learning theory that is the structural risk minimization principle described by Vapnik [19]. SVMs are constructed by locating a set of hyper planes that separate two or more classes of data. It discovers the boundaries between the input classes; the elements of the input data that defines these boundaries are called support vectors [20]. SVMs are maximal margin classifiers. They are free of the optimization headaches of neural networks because they present a convex programming problem [21]. SVM provides a generic technique to fit the surface of the hyper plane to the data points through the use of an appropriate kernel function i.e. data points that cannot be separated by a linear decision boundary in the input space, can be transformed into a high dimensional feature space to make them linearly separable. The choice of kernel function and its parameters to map dataset well in high dimension may depend on specific dataset. To achieve good generalization performance of SVM there is a need to select appropriate kernel function and its parameter, and trade-off parameter C. A properly designed kernel function and its parameters can minimize generalization error, speed up convergence rate, and boost classification accuracy [22]. SVM function uses an optimization method to identify support vectors Si, weights αi and bias b that are used to classify vectors x according to equation (24).

ROC curve is plot of the true positive rate (sensitivity) versus the false positive rate (specificity) for different thresholds over the entire range of each classifier discriminant function output values. Area under the ROC curve (AUC) is the probability of the classifier to correctly classify malignancy from benignancy cases [23].

$$C = \sum_i \alpha_i\, k(s_i, x) + b \quad \text{(xxiv)}$$

Where, K represents a kernel function in equation (24). The proposed research aims in designing of SVM classifier using linear, polynomial, RBF and MLP kernels.

III. EXPERIMENTAL RESULTS

Pattern recognition model, SVM is developed and utilized texture and wavelet features for detecting the malignancy of thyroid nodule. First the USG images are filtered using SRAD filter to make them speckle free. SRAD performance is evaluated based on Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR). MSE is calculated as the mean square error of the corresponding pixel in the input image and the image obtained after despeckling. PSNR value should be higher for the despeckled USG image. SRAD performance is compared with Frost filter and Spatiotemporal filter. Table 1 shows the MSE and PSNR comparison of Frost, SRAD and spatiotemporal filter.

Table. 1. Performance Comparison of Filters

Filter	Mean Square Error (MSE)	Peak Signal to Noise Ratio (PSNR)
Frost	1.0992e-009	137.7542
Spatiotemporal filter	4.3967e-007	111.7336
Speckle Reducing Anisotropic Diffusion (SRAD) filter	2.7479e-010	143.7748

Total 22 features are extracted from nodular region of thyroid. Region of Interest (ROI) is separated by cropping the nodular region. Fig. 2 (a) shows the separation of ROI from USG image of thyroid.

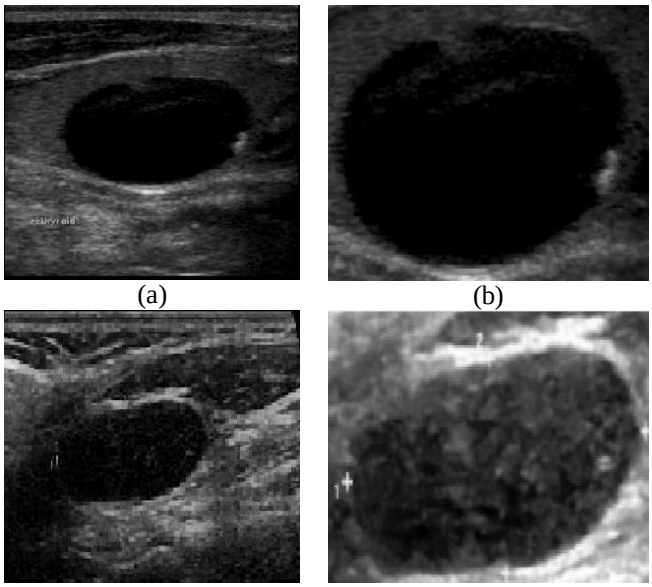


Figure2: Thyroid nodule images (a) Benign cystic thyroid nodule (b) Separation of nodular region (ROI) (c) Malignant solid thyroid nodule (d) ROI separation of malignant thyroid nodule

After pre-processing and ROI separation, the features extraction process has been applied. 15 features are extracted using GLCM. These features are used to design SVM classifier for classifying thyroid nodule into benign and malignant. Percentage accuracy of classification is 90% if all 15 features are used for classification. Minimum Redundancy Maximum Relevance (mRMR) algorithm is used to select the best features for classification to further improve the accuracy. Table 2 displays the performance of mRMR algorithm for best feature combination. It is observed that accuracy is highest if all features are considered in the process of classification. However, this highest accuracy of 90% is not encouraging from the perspective of clinical deployment.. To improve the accuracy further, more features are extracted using Discrete Wavelet Transform (DWT). 7 features are extracted using DWT. These are A2, V2, H2, D2, H1, V1, and D1. Fifteen features obtained from GLCM and 7 features obtained from DWT are fused together and applied as an input to SVM classifier. Significant rise of approximately 9% to10% is observed if 22 feature vectors are used for classification. Further SVM performance is analyzed to select appropriate kernel function.

Table. 2. Performance of mRMR algorithm for best feature combination in proposed system

Best Features	Accuracy	Feature Combination
2	80	Cluster Prom., Dissimilarity
3	78	Cluster Prom., Dissimilarity, Homogeneity
5	86	Cluster Prom., Dissimilarity, Homogeneity, Mean, Variance
6	86	Cluster Prom., Dissimilarity, Homogeneity, Mean, Variance, Entropy
7	88	Cluster Prom., Dissimilarity, Homogeneity, Mean, Variance, Entropy, Std. Dev.
8	86	Cluster Prom., Dissimilarity, Homogeneity, Mean, Variance, Entropy, Std. Dev., Contrast
10	86	Cluster Prom., Dissimilarity, Homogeneity, Mean, Variance, Entropy, Std. Dev., Contrast, Correlation, Cluster Shade
15	90	All GLCM features

In training a SVM we need to select appropriate kernel function and its parameters and value of trade-off parameter C. For a given feature vectors, SVM are trained using linear, RBF, MLP and polynomial kernels. The corresponding classification accuracies are analyzed and compared. Fig. 3 reports the accuracies for all kernels under test for SVM classifier. For polynomial kernel function, the polynomial degree variations are also done to get the highest classification accuracy. Variation in classification accuracy for polynomial degree variation can also be observed in Fig.3.

The choice of the trade-off parameter C is also vital to obtain good classification results. Accuracy depends on the trade-off between a high-complexity model which may over-fit the data and a large-margin which will incorrectly classify some of the training data in the interest of better generalization. The number of support vectors can range from very few to every single data

point if data is completely over-fitted. This trade-off is controlled via C and through the choice of kernel and kernel parametersIn this study, trade off (C) variations is done from 0.1 to 0.3 with 0.1 differences. It is observed from Fig. 4 that for our feature space, accuracy is optimal for trade-off factor of 0.2.

Fig. 5 and Fig.6 reports the results of ROC analysis of SVM classifier. AUC represents the probability that a random pair of malignant and benign thyroid nodules will be correctly classified. Sensitivity (SN)depends only on measurements of malignant nodules and specificity(SP) only on benign nodules. Highest classification accuracy of 99.98% is achieved with AUC of 0.995, employing the polynomial kernel of 3rd degree. The parameter C that controls the trade off is set at 0.2. This combination gives the sensitivity and specificity of 0.99 and 0.99 respectively. Fig. 7 shows the variation of support vectors for various kernel functions. The support vectors ranges from 1.29% to 2.45% of the number of training points. Computational complexity of the model is linear in the number of support vectors. Fewer support vectors means faster classification of test points.Since the decision hyper-plane of the SVM classifier is determined by the support vectors, their minimum number is indicative of the SVM low complexity and differentiation capability.

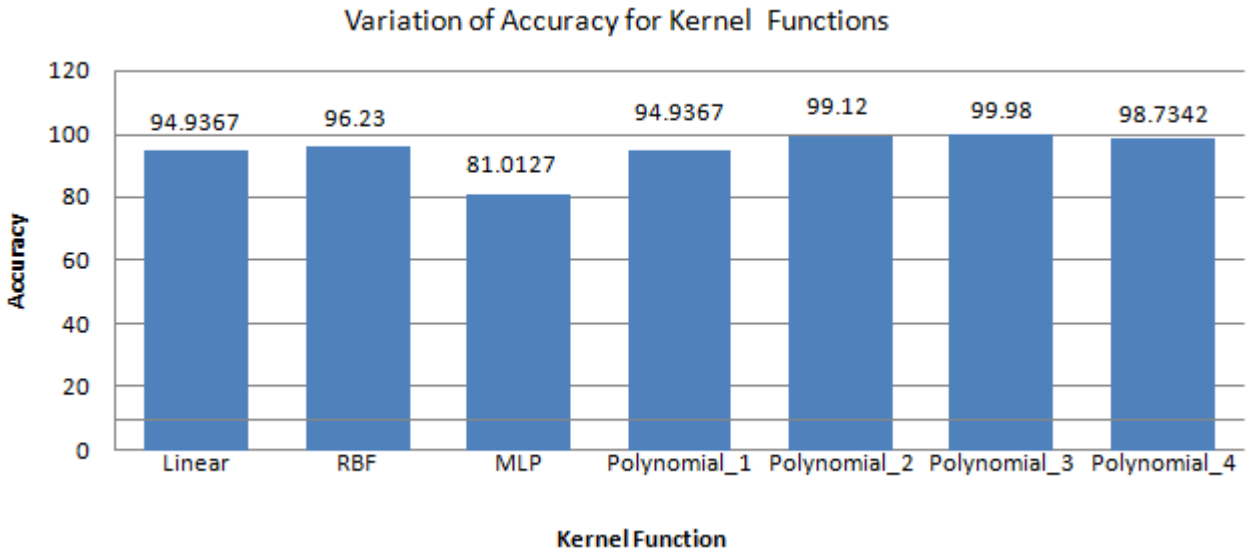


Figure 3: Performance comparison of linear, RBF, MLP and polynomial (with variation in polynomial degree) kernels for SVM classifier

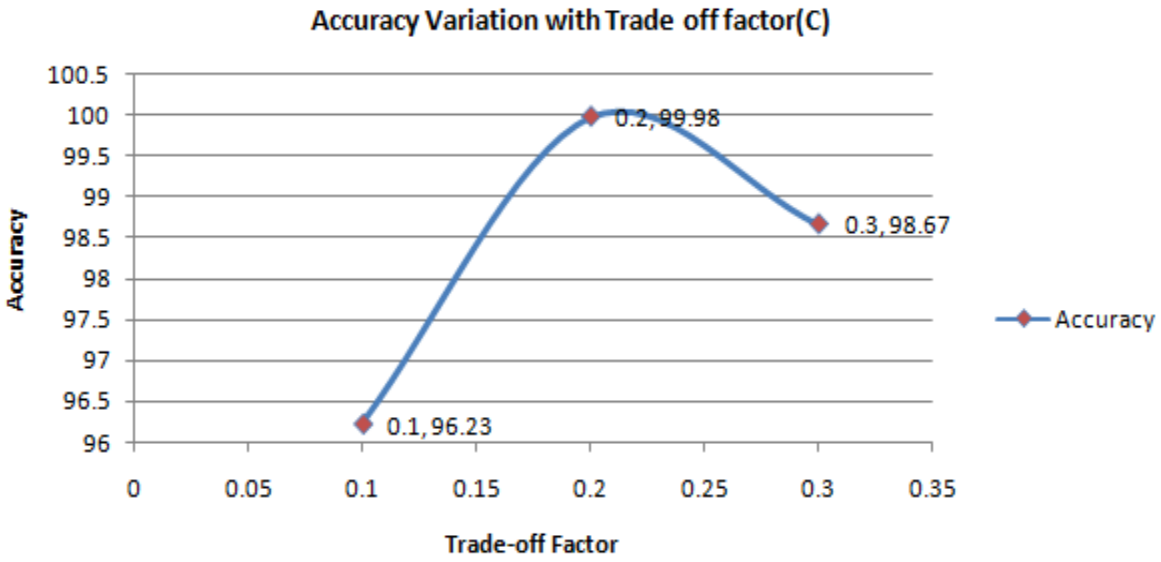


Figure4: Variation in accuracy with trade off factor

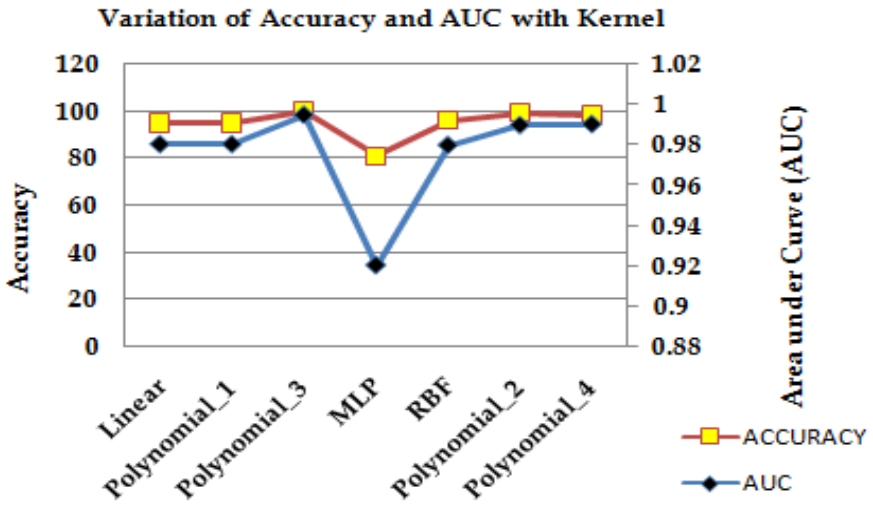


Figure 5: Variation of accuracy and AUC with kernel function

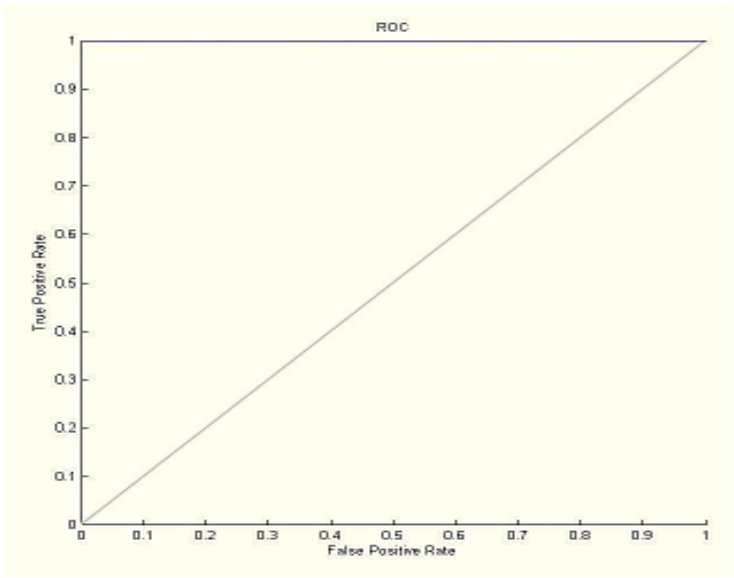


Figure 6: Receiver Operating Characteristic (ROC) Curve of SVM classifier with polynomial of 3rd degree

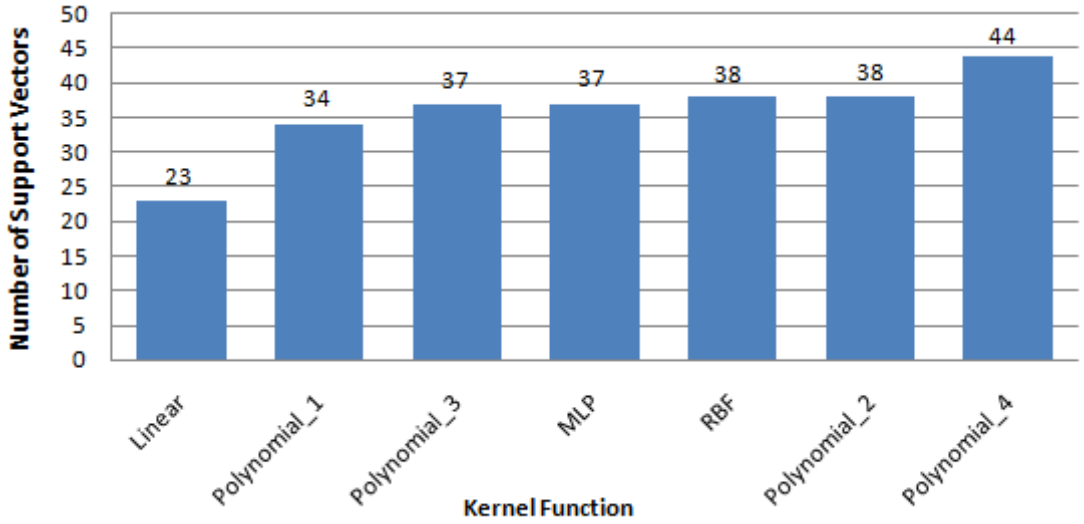


Figure 7: Number of Support Vectors for SVM kernels

IV. DISCUSSION AND CONCLUSION:

This research develops a computerized system to aid the decision making about a malignant thyroid. A feature vector comprising of fifteen features derived using GLCM and seven textural features derived using DWT is generated to achieve high classification accuracy. Literature reports use of DWT and GLCM to achieve classification accuracy of 98.9%.However, this

research presents a novel approach that fuses the DWT and GLCM based features to achieve comparatively very high classification accuracy.

Coarse calcification and micro calcification in predominantly solid nodule, solid vs. cystic, texture of the nodule, blood flow within the nodule [25]are some of the prominent characteristics of a nodule, as identified by a radiologist that play a vital role in its probable classification as malignant. Calcifications are small flecks of calcium within a thyroid nodule, usually seen as bright spots on USG image. These white spots, containing specific orientation, cover the whole nodule region. Derived DWT features extract this structure orientation, which can be in horizontal, vertical or diagonal directions. GLCM also plays very important role for finding the differentiating texture information of malignancy. Energy, entropy, dissimilarity, homogeneity, autocorrelation, contrast, correlation, cluster prominence, cluster shade, maximum probability, mean, variance, standard deviation, kurtosis and skewness are calculated from GLCM matrix, which in itself contain information about the texture of nodular region. Therefore features derived in the proposed approach imbibe the expert knowledge in the system thus ensuring superior classification accuracy. The designed intelligent system thus aids the decision making of the radiologist about malignancy.

Standard database from Wilmington Endocrinology comprising of 79 ultrasound images of thyroid nodule (51 benign and 28 malignant) are considered to examine the efficacy of the proposed algorithm. The performance of the SVM classifier is evaluated considering polynomial, RBF, linear and MLP kernel functions and analyzing the respective ROCs. Highest classification accuracy of 99.98% with AUC of 0.995, employing the polynomial kernel of 3rd degree confirms the superiority of the proposed approach.

This research thus presents a comprehensive study that aims to extract several second order statistical texture features using GLCM and wavelet based features using DWT together with the design of powerful SVM classifier, in order to evaluate the malignancy of thyroid nodule in USG image. SVM is used as an intelligent classifier to identify the malignant thyroid nodules using the fusion of the derived features. It is worth noting that the algorithm, presented in [10]is limited to the detection of nodule tissue from thyroid region, whereas the detection of malignancy is very important issue and is not addressed in their study. Methodologies reported in [8] also used SVM for classification but the number of support vectors ranges from 8% to 10% of the number of training points. It is indicative of the SVM complexity and less differentiation capability. It must be emphasized, that the reported methods are all tested on independent data sets. A direct comparison may not necessarily justify their effectiveness and efficiency but a general idea of the performance might be obtained. The use of mRMR algorithm in the proposed system helps us to come to the conclusion that all extracted features equally contribute in detection of thyroid malignancy. So the summarized results are merely indicative and prove the efficacy of the approach for its clinical deployment. Our further research aims to deploy the designed system in clinics for aiding the decision making of a radiologist when applied to practical cases.

References

i. Sholosh, Biatta, and Amir A. Borhani. "Thyroid ultrasound part 1: technique and diffuse disease." *Radiologic Clinics of North America* 49.3 (2011): 391-416.

ii. Rouhani M,Mansouri K. Comparison of several ANN architectures on the thyroid diseases grades diagnosis. 2009 International Association of computer science and information technology-spring conference; 2009. pp 526-528.

iii. NCI (National Cancer Institute) on thyroid cancer. Information available at http://seer.cancer.gov/statfacts/html/thyro.html (last accessed January 2016)

iv. ACS (American Cancer Society). Information available at http://www.cancer.org/cancer/thyroidcancer/detailedguide/thyroid-cancer-key-statistics (last accessed January 2016).

v. Hoang, Jenny K., et al. "US Features of Thyroid Malignancy: Pearls and Pitfalls 1." *Radiographics* 27.3 (2007): 847-860.

vi. Amedee, Ronald G., and Nina R. Dhurandhar. "Fine-Needle Aspiration Biopsy." *The Laryngoscope* 111.9 (2001): 1551-1557.

vii. Acharya, U. Rajendra, et al. "A review on ultrasound-based thyroid cancer tissue characterization and automated classification." Technology in cancer research & treatment 13.4 (2014): 289-301.

viii. Tsantis, Stavros, et al. "Morphological and wavelet features towards sonographic thyroid nodules evaluation." *Computerized Medical Imaging and Graphics* 33.2 (2009): 91-99.

ix. Acharya, U. R., Faust, O., Sree, S. V., Molinari, F., Garberoglio, R., Suri, J. S. Cost-effective and non-invasive automated benign and malignant thyroid lesion classification in 3D contrast-enhanced ultrasound using combination of wavelets and textures: A class of thyroscan algorithms. Technol Cancer Res Treat 10, 371-380 (2011).

x. Keramidas, Eystratios G., Dimitris Maroulis, and Dimitris K. Iakovidis. "TND: a thyroid nodule detection system for analysis of ultrasound images and videos." *Journal of medical systems* 36.3 (2012): 1271-1281.

xi. Nikita Singh and Alka Jindal "A Segmentation and Comparison of Classification Method for Thyroid Ultrasound Images." International Journal of Computer Applications (0975-8887) Volume 50-No 11,.July 2011

xii. Zhao J, Zheng W, Zhang L, Tian H. Segmentation of Ultrasound Images of Thyroid Nodule for Assisting Fine Needle Aspiration Cytology. Health Information Science and Systems; 2012. 1:5.

xiii. www.wilmingtonendo.com

xiv. Mendre, Wrushali, R. D. Raut, and Sayali Malwadkar. "Performance Comparison of Speckle Reducing Filters for Ultrasonography USG) Images of Thyroid Glands." *International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCSE)* 4 (2015): 33.

xv. Haralick, Robert M., KarthikeyanShanmugam, and Its' HakDinstein. "Textural features for image classification." *Systems, Man and Cybernetics, IEEE Transactions on* 6 (1973): 610-621.

xvi. Conners, Richard W., Mohan M. Trivedi, and Charles A. Harlow. "Segmentation of a high-resolution urban scene using texture operators."*Computer Vision, Graphics, and Image Processing* 25.3 (1984): 273-310.

xvii. Albregtsen, Fritz. "Statistical texture measures computed from gray level coocurrence matrices." *Image processing laboratory, department of informatics, university of oslo* (2008): 1-14.

xviii. Chapelle, Olivier, Patrick Haffner, and Vladimir N. Vapnik. "Support vector machines for histogram-based image classification." *Neural Networks, IEEE Transactions on* 10.5 (1999): 1055-1064.

xix. Vapnik, Vladimir. *The nature of statistical learning theory*. Springer Science & Business Media, 2013.

xx. Sivanandam, S. N., and S. N. Deepa. *Introduction to neural networks using Matlab 6.0*. Tata McGraw-Hill Education, 2006.

xxi. Burges, Christopher JC. "A tutorial on support vector machines for pattern recognition." *Data mining and knowledge discovery* 2.2 (1998): 121-167.

xxii. Muller K-R, Mikas, Ratsch G, Tsuda K, Scholkopf B. An introduction to kernel- based learning algorithm, IEEE Transaction in Neural Networks, 2001; (12): 181-202

xxiii. Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. Journal of Biomedical Informatics 2005;38:404-15

xxiv. Frates, Mary C., et al. "Management of Thyroid Nodules Detected at US: Society of Radiologists in Ultrasound Consensus Conference Statement 1."*Radiology* 237.3 (2005): 794-800.