International Journal of Engineering & Technology, 6 (3) (2017) 50-53

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v6i3.7705
Research paper

Detection and classification of thyroid nodule using Shearlet coefficients and support vector machine

Nanda S. [1]*, M. Sukumar [2]

[1] *Research Scholar, JSS Research Foundation, Sri Jayachamarajendra College of Engineering, Mysuru-570006, Karnataka, India*
[2] *Research Guide, JSS Research Foundation, Sri Jayachamarajendra College of Engineering, Mysuru-570006, Karnataka, India*
**Corresponding author E-mail: nanda_prabhu@sjce.ac.in*

Abstract

Thyroid nodules have diversified internal components and dissimilar echo patterns in ultrasound images. Textural features are used to characterize these echo patterns. This paper presents a classification scheme that uses shearlet transform based textural features for the classification of thyroid nodules in ultrasound images. The study comprised of 60 thyroid ultrasound images (30 with benign nodules and 30 with malignant nodules). Total of 22 features are extracted. Support vector machine (SVM) and K nearest neighbor (KNN) are used to differentiate benign and malignant nodules. The diagnostic sensitivity, specificity, F1_score and accuracy of both the classifiers are calculated. A comparative study has been carried out with respect to their performances. The sensitivity of SVM with radial basis function (RBF) kernel is 100% as compared to that of KNN with 96.33%. The proposed features can increase the accuracy of the classifier and decrease the rate of misdiagnosis in thyroid nodule classification.

Keywords: *Co Occurrence Matrix; Texture Analysis; Thyroid Nodule; Shearlet Transform.*

1. Introduction

Thyroid nodule is common in general population that can be benign or malignant. It is an abnormal growth of thyroid cells as a lump within the thyroid gland. Woman are more affected by the thyroid cancer than men. About 50% of the adults have thyroid nodules, out of which only 5% turns out to be malignant [1]. Ultrasound imaging, a non-invasive imaging modality is more popular in evaluating thyroid nodules. [2]. In most of the cases, benign nodules often have round or ellipsoid shapes, smooth borders and homogeneous internal echoes, whereas malignant nodules often have branch patterns, spiculations, angular borders and heterogeneous internal echoes [3]. These characteristics of ultrasound images are used by the radiologists to differentiate the nodules which is qualitative in nature. The accuracy of the diagnosis is improved if the irregularity of the sonographic findings is quantified correctly through the extracted features. Many researchers have worked towards the goal of automatic detection and classification of thyroid nodules using textural features and morphological features. Stavros Tsantis et. al [4] presented a computer based classification scheme that utilized various morphological and wavelet based features, Michalis Savelonas et. al [5] proposed a method based on boundary features towards malignancy risk evaluation of thyroid nodules in US images. Further the use of gabor filters to characterize the degree of orientation present in ultrasound image textures is discussed by Grigorescu, S. N et. al [6]. Also the representation of complex patterns in an ultrasound image using a parameter fractal dimension obtained from fractal geometry is presented in Yuan Y. Tang et. al [7]. The use of SVMs for the selection of significant textural features and to classify the nodular lesion of a thyroid is discussed by Chan-Yu Chang et. al [8]. An effective method of segmentation of thyroid nodules for assisting fine needle aspiration cytology (FNAC) is presented in Jie Zhao et. al [9]. U. Rajendra Acharya et. al [10] developed an automated identification system for characterizing the intra nodular vascularization of thyroid lesions. U Rajendra Acharya et. al [11] also summarized thyroid cancer tissue characterization and automated classification. K Guo et. al [12] have showed the potentiality of shearlet transform to represent the textural information. The use of shearlet transform to extract the textural features for better classification of breast tumors with an accuracy of around 90% is reported by Shichong Zhou et. al [13]. Hence in this study features based on the shearlet transform are extracted from the ultrasound images and SVM and KNN are used to distinguish the nodules.

2. Materials and methods

2.1. Data collection

Thyroid ultrasound images used in this study are taken from Digital database of thyroid ultrasound images (DDTI) [14]. Each image in the database contains delineated nodule (done by the expert radiologist). The nature of the nodule is given in terms of margin characteristics along with the TIRADS (Thyroid imaging reporting and data system) levels. Thyroid nodules are classified into 7 levels (1, 2, 3, 4a, 4b, 4c, 5). Database used in this study consists of 60 thyroid images out of which 30 images are of TIRAD level 2 and 3 and 30 are of levels 4c and 5. Each image is annotated in terms of pathological features such as size, shape, margin, composition, calcifications and echogenicity for a given view (sagital or transverse) and their pathologies confirmed by biopsy using BETHESDA (system for reporting thyroid cytopathology) system.

2.2. Feature extraction

Feature extraction plays a major role in the classification of nodules. Features are the descriptors used to characterize the nodule. In this study the textural features extracted using the Shearlet transform. Shearlet transform is a multiscale directional transform that helps in the analysis and representation of an image. It is a method that is used to detect directional features [15], [16] in images. This multidirectional representation, is more powerful in understanding the geometry of images. The continuous shearlet transform of an image f is the mapping

$$f \to SH_{\psi} f(a,s,t) = < f, \psi a, s, t > \quad (1)$$

Where ψ is a generating function, a > 0 is the scale parameter, s ϵ R is the shear parameter, t ϵ R^2 is the translation parameter, and the shearlet basis functions $\psi_{a,s,t}$ is defined as,

$$\psi_{a,s,t}(x) = \left|M_{a,s}\right|^{-\frac{1}{2}}(M_{a,s}^{-1}(x-t)) \quad (2)$$

where

$$M_{a,s} = \begin{bmatrix} a & s\sqrt{a} \\ 0 & \sqrt{a} \end{bmatrix}$$

The shearlets $\psi_{a,s,t}$ are the group of well-localized waveforms at various scales a, orientations s and locations t.

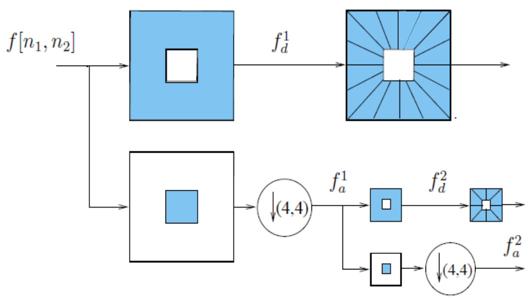

Fig. 1: Shearlet Decomposition of an Image.

Fig.1 illustrates the discrete shearlet decomposition of an image. It consists of two main procedures, the Laplacian pyramid decomposition procedure and the directional filtering. The first level decomposition gives 16 directional subbands and the second level decomposition gives 8 directional subbands. The shearlet basis functions can be more compactly supported in the frequency domain. Thus, finer image detail information can be well captured by this type of basis functions. Multiscale decomposition and directional localization are the two primary steps of Shearlet transform. Because of its superior directional sensitivity at various scales this transform is used for the extraction of features. Hence shearlet based texture feature descriptors can characterize thyroid nodules well.

Discrete shearlet coefficients of the images in the dataset are calculated and co-occurrence matrix of the Shearlet coefficients [17] is computed from each image. This will give the information about the texture of the images, since Shearlet coefficients are good representatives of the heterogeneity of images.

In this study the scaling factor of 2 is selected for two level decomposition of the region of interest in an image using shearlet transform. 11 features namely energy, correlation, entropy, autocorrelation, contrast, cluster prominence, cluster shade, dissimilarity, homogeneity, squared variance, sum average are computed from each level of shearlet coefficients. A total of 22 shearlet features are obtained from each region of interest. Implementation details o these features are as follows.

2.2.1. Energy

Also known as Angular Second Moment is a measure of homogeneity of an image.

$$Energy = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p(i,j)^2 \quad (3)$$

2.2.2. Correlation

Is a measure of the linear dependency of gray levels with the neighborhood pixels.

$$corr = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \frac{(ij)(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y} \quad (4)$$

2.2.3. Entropy

Measures the randomness of the image texture.

$$Entropy = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p(i,j)\log(p(i,j)) \quad (5)$$

2.2.4. Contrast

Is a measure of the local variations in an image.

$$Contrast = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}|i-j|^2 p(i,j) \quad (6)$$

2.2.5. Cluster prominence

Is a measure of asymmetry.

$$CP = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left(i+j-\mu_x-\mu_y\right)^4 p(i,j) \quad (7)$$

2.2.6. Cluster shade

Is a measure of the skewness and is used to gauge the perceptual concepts of uniformity.

$$CS = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left(i+j-\mu_x-\mu_y\right)^3 p(i,j) \quad (8)$$

2.2.7. Dissimilarity

Is a measure that defines the variation of gray level pairs in an image.

$$Dissim = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}|i-j|p(i,j) \quad (9)$$

2.2.8. Homogeneity

This statistic is also called as inverse difference moment and measures image homogeneity.

$$Hom = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{p(i,j)}{1+(i-j)^2} \quad (10)$$

2.2.9. Squared variance

It refers to the gray level variability of the pixel pairs and is a measurement of heterogeneity.

$$Var = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i-\mu_x)^2 p(i,j) + \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left(i-\mu_y\right)^2 p(i,j) \quad (11)$$

2.2.10. Sum average

It is a secondary feature.

$$Sumavearge = \sum_{i=2}^{2(N-1)} ip_{x+y}(i) \quad (12)$$

where p(i,j) is the normalized co occurrence matrix obtained from the shearlet coefficients, μ_x and μ_y are the mean of p_x and p_y respectively, σ_x and σ_y are the standard deviation of p_x and p_y respectively. The parameter

$$p_{x+y}(k) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p(i,j) \qquad \text{for k = 0,1, ...2(N-1)} \tag{13}$$

2.3. Classification

Classification of thyroid nodules as benign or malignant is carried out using SVM and KNN classifiers.

2.3.1 Support vector machine

SVM is primarily a two class supervised learning model. This model constructs an hyperplane which separates two classes with a margin [18], [19]. The instances that are very close to the hyperplane are called support vectors. The hyperplane that separates two classes is represented as

$$F(x) = w_0 + w_1\, a_1 + w_2 a_2 \tag{14}$$

Where w_0, w_1 and w_2 represent weights and a1 and a2 are the attributes.
The hyperplane in terms of support vectors for maximum margin is represented as

$$f(x) = b + \sum_{i=1}^{l} \propto_i\, y_i < x_i\, .x > \tag{15}$$

Where i is the support vector, y_i = +1 or -1 represents the class value, x_i is the i^{th} support vector, x is the test vector, $<x_i.x>$ gives the dot product, b and α's are parameters of the hyperplane.
SVM is also used to separate overlapping data by transforming input feature space to a new space through a nonlinear transformation [20] . This nonlinear transformation results in a large dimensional space with large number of attributes which is not desirable. Hence a kernel function can be used to the instances in the input space which brings the same effect as linear transformation. Different kernel functions can be used to construct various learning models of SVM. In our study experimentation was done with polynomial and RBF kernels and RBF kernel performed good.
The polynomial kernel of degree d is represented as

$$K(x, x_i) = (x.x_i + 1)^d \tag{16}$$

RBF kernel is given by

$$K(x, x_i) = exp\frac{-|x-x_i|^2}{2\sigma^2} \tag{17}$$

Where σ refers to the width of the Gaussian function.

2.3.2 K nearest neighbor (KNN)

KNN classifier is an instance based classifier in which the classification of an unknown sample is done by relating the unknown to a known sample based on some distance or similarity criteria [21]. Here the class is assigned to a sample which is the most common among its K-nearest neighbors. The distance considered is the Euclidean distance which is represented as

$$d_E\,(x, x_i) = \sum_{i=1}^{N} \sqrt{x^2 - {x_i}^2} \tag{18}$$

The 22 features calculated from the co occurrence matrix of the shearlet coefficients are given as the input to these two classifiers. Performances of SVM and KNN classifiers are evaluated with the help of confusion matrix shown in Fig.2.

True Negative (TN)	False Positive (FP)
False Negative (FN)	True Positive (TP)

Fig. 2: Confusion Matrix.

TP: No. of malignant nodules detected as malignant
TN: No. of benign nodules detected as benign
FP: No. of benign nodules detected as malignant
FN: No. of malignant nodules detected as benign

The following performance measures are calculated from the confusion matrix.
Sensitivity: is the ability of the classifier to correctly identify the malignant nodules (true positive rate).

$$Sensitivity = \frac{TP}{(TP+FN)}\ X\ 100 \tag{19}$$

Specificity: is the ability of the classifier to correctly identify the benign nodules (true negative rate).

$$Specificity = \frac{TN}{(FP+TN)}\ X\ 100 \tag{20}$$

F1 Score: is the weighted average of precision (positive predictive value) and recall (sensitivity). This score considers both false positives and false negatives.

$$F1\_Score = \frac{2TP}{(2TP+FP+FN)}\ X\ 100 \tag{21}$$

Accuracy: is the ability of the classifier to correctly identify the malignant nodules as malignant and benign nodules as benign.

$$Accuracy = \frac{TP+TN}{(TP+FN+FP+TN)}\ X\ 100 \tag{22}$$

3. Results and discussion

The ultrasound images of 60 patients comprising of 30 benign and 30 malignant thyroid nodules are taken from DDTI. As Shearlet transform is highly effective at detecting both the location and orientation of edges it is used to know the textural variation of thyroid ultrasound images. Two level decomposition of the region of interest using shearlet transform is done. 11 features, energy, correlation, entropy, autocorrelation, contrast, cluster prominence, cluster shade, dissimilarity, homogeneity, sum of squared variance, sum of average are computed from each level of shearlet coefficients. From each region of interest 22 shearlet features are obtained.
In the classification phase 10 fold cross validation is used to select the images for training and testing. This is repeated 10 times and the average classification results are obtained and tabulated. In SVM both polynomial kernel and RBF kernel are used. We chose three values [1, 2, 3] for the polynomial kernel degree and the polynomial kernel of degree 3 is giving good result. The average classification results of SVM using polynomial kernel are reported in Table 1.

Table 1: Average Classification Results Using SVM with Polynomial Kernel of Three Degrees

Degree of	Performance measures (%)			
Polynomial	Sensitivity	Specificity	F1_score	Accuracy
1	92.33	89.00	91.07	90.66
2	92.66	96.33	94.02	94.50
3	96.33	96.00	96.08	96.16

From Table 1 it is clear that polynomial kernel of degree 3 is correctly classifying 96.33% of the malignant nodules as malignant with an overall accuracy of 96.16%.
The selection of parameters C (soft margin constant) and sigma (width of the Gaussian kernel) of SVM plays a major role in the classification accuracy. In this study experimentation has been carried out by taking five different values for both the parameters (sigma of the Gaussian kernel and hyper parameter C). The better accuracy is obtained for C=1 and sigma=5.

Table 2: Results of Classifiers

		Sensitivity	Specificity	F1_Score	Accuracy
SVM	Polynomial kernel	96.33	96.00	96.08	96.16
	RBF kernel	100	92.67	96.85	96.33
KNN		96.33	92.67	94.65	94.50

Table 2 reports performance measures of different classifiers and Fig. 3 shows the graphical representation of the statistics given in Table 3. From the results it is inferred that SVM with RBF kernel is 100% sensitive (classifying all the malignant nodules as malignant) in classification. 92.67% of the benign nodules are classified as benign and 7.33% of benign nodules are misclassified as malignant. Further classification accuracy of KNN is 94.50% with the sensitivity of 96.33%.

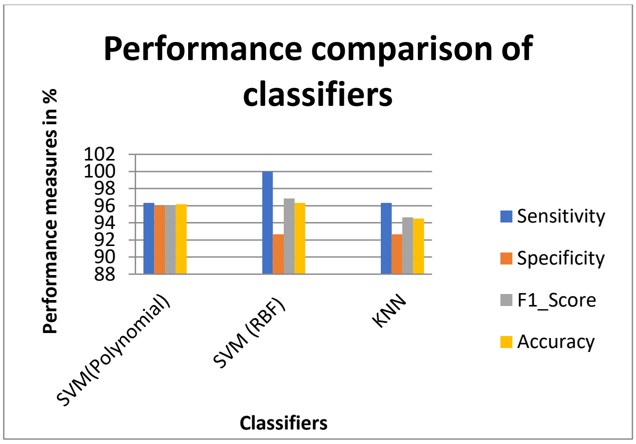

Fig. 3: Comparison of Performance Measures.

4. Conclusion

Precise characterization and classification of thyroid nodules are very essential which would assist the radiologists for accurate diagnosis. A comprehensive study has been done that aimed at the extraction of shearlet coefficients based features along with the help of two classifiers (SVM and KNN) to differentiate the thyroid nodule as benign or malignant in ultrasound images. Results show that SVM with RBF kernel gives a better accuracy of 96.33%, sensitivity of 100%, specificity of 92.67% and 96.85% F1_score. Polynomial kernel of degree three gives good accuracy of 96.16% and KNN gives an accuracy of 94.50%. Hence the combination of this quantitative analysis and the qualitative (visual) analysis done by the radiologists results in the improved diagnostic accuracy thus reducing the number of cases going for fine needle aspiration. Future perspective is to improve the classification accuracy using a better classifier and also through the feature selection.

References

[1] D Ghassi, A Donato, Evaluation of the thyroid nodule", *Postgraduation Medical Journal*, Vol. 85, No. 1002, (2009), pp. 190-195, https://doi.org/10.1136/pgmj.2008.072140.
[2] K T Wong, Anil T Ahuja, "Ultrasound of Thyroid Cancer", *Cancer Imaging, BioMed central*, Vol. 5, No. 1, (2005), pp. 157-166.
[3] Young Hun Lee, Dong Wook Kim, Hyun Sin in et al. "Differentiation between benign an malignant solid thyroid nodules using an US classification system", *Korean Journal of Radiology*, Vol. 12, No.5, (2011), pp. 559-567. https://doi.org/10.3348/kjr.2011.12.5.559.
[4] Stavros Tsantis, Nikos Dimitropoulos, Dionisis Cavouras, George Nikiforidis, "Morphological and wavelet features towards sonographic thyroid nodules evaluation", *Computerized Medical Imaging and Graphics, Elsevier*, Vol. 33, No. 2, (2009), pp 91-99. https://doi.org/10.1016/j.compmedimag.2008.10.010.
[5] Michalis Savelonas, Dimitris Maroulis, Manolis sangriotis, "A computer aided system for malignancy risk assessment nodules in thyroid US images based on boundary features", *Computer Methods and programs in Biomedicine*, Elsevier, Vol. 96, No. 1 (2009), pp. 25-32.
[6] Grigorescu, S. N., Petkov, N., Kruizinga, P. "Comparison of Texture Features Based on Gabor Filters", *IEEE Transactions on Image processing*, Vol. 11, No. 10, (2002), pp. 1160-1167. https://doi.org/10.1109/TIP.2002.804262.
[7] Yuan Y. Tang, Yu Tao, Ernest C.M. Lam, "New method for feature extraction based on fractal behavior", *Pattern Recognition*, Vol. 35, No.5 (2002), pp. 1071–1081, https://doi.org/10.1016/S0031-3203(01)00095-4.
[8] Chuan-Yu Chang, Shao-Jer Chen, Ming-Fong Tsai, "Application of Support vector machine based method for feature selection and classification of thyroid nodules in ultrasound images", *Pattern Recognition, Elsevier*, Vol. 43, No. 10, (2010), pp 3494-3506. https://doi.org/10.1016/j.patcog.2010.04.023.
[9] Jie Zhao, Wei Zheng, Li Zhang and Hua Tian, "Segmentation of ultrasound images of thyroid nodule for assisting fine needle aspiration cytology", *Journal on Health Information Science and Systems*, Vol. 1, No. 1, (2012), pp. 1-12.
[10] U Rajendra Acharya, Oliver Faust, S Vinitha sree, Filippo Molinari, Jasjit S Suri, "Thyroscreen system: High resolution ultrasound thyroid image characterization int benign and malignant classes using novel combination of texture and discrete wavelet transform", *Computer Methods and programs in biomedicine, Elsevier*, Vol. 107, No.2, (2012), pp 233-241.
[11] U Rajendra Acharya, G Swapna, S Vinitha Sree, Filippo Molinari, Savita Gupta, Ricardo H Bardales, Agnieszka Witkowska and Jasjit S Suri, " A review on ultrasound based thyroid cancer tissue characterization and automated classification", *Technology in cancer research and treatment*, Vol. 13, No. 4, (2014), pp 288-301.
[12] K Guo , D Labate, "Optimally sparse multidimensional representation using Shearlets", *SIAM Journal on Mathematical Analysis*, Vol 30, No. 1, (2007), pp 298-318. https://doi.org/10.1137/060649781.
[13] Shichong Zhou, Jun Shi, Jie Zhu, Yin Cai, Ruiling Wang, Shearlet based texture feature extraction for classification of breast tumor in ultrasound image", *Biomedical Signal processing and control ,Elsevier*, Vol 8, No. 6, (2013), pp 688-696.
[14] Lina Pedraza , Carlos Vargas, Fabian Narvaez , Oscar Duran , Emma Munoz, Eduardo Romero, "An open access thyroid ultrasound-Image Database", *Proceedings of SPIE*, Vol. 9287 (2015).
[15] Easley, G.R.; Labate, D.; Colonna, F, "Shearlet-based total variation diffusion for denoising", *IEEE Transactions on Image Processing*, Vol 18, No. 2, (2009), pp. 260–268. https://doi.org/10.1109/TIP.2008.2008070.
[16] Easley, G.; Labate, D.; Lim, W.Q, "Sparse directional image representations using the discrete shearlet transform", *Apllied and Computational Harmonic Analysis*, Vol 25, No.1, (2008), pp. 25–46 , https://doi.org/10.1016/j.acha.2007.09.003.
[17] Hadi Rezaeilouyeh, Mohammad H Mahoor, "Automatic Gleason Grading of Prostate Cancer using Shearlet Transform and Multiple Kernel Learning", *Journal of Imaging* , Vol. 2, No.3, (2016), 25. https://doi.org/10.3390/jimaging2030025.
[18] Corinna Cortes, Vladimir Vapnik. "Support-Vector Networks", *Machine Learning*, Vol. 20, No 3, (1995), pp. 273-297. https://doi.org/10.1007/BF00994018.
[19] Burges CJS, "A tutorial on support vector machines for pattern recognition, data mining and knowledge", Discovery 2, (1998), pp 121-167.
[20] Muller K R, Mika S, Raisch G, Tsuda K Scholkopf B , "An introduction to kernel-based learning algorithms", *IEEE Transactions on Neural Networks*, Vol. 12, No.2, (2001), pp 181-202. https://doi.org/10.1109/72.914517.
[21] Larode D T KNN. In: Discovering Knowledge in Data: An introduction to data mining, Wiley Interscience, (2004), pp 90-106.