ISSN: 0976-9102(ONLINE)
DOI: 10.21917/ijivp.2011.0046

CLASSIFYING BENIGN AND MALIGNANT MASSES USING STATISTICAL MEASURES

B. Surendiran[1] and A. Vadivel[2]
Department of Computer Applications, National Institute of Technology Tiruchirappalli, India
E-mail: [1]surendiran@gmail.com and [2]vadi@nitt.edu

Abstract
Breast cancer is the primary and most common disease found in women which causes second highest rate of death after lung cancer. The digital mammogram is the X-ray of breast captured for the analysis, interpretation and diagnosis. According to Breast Imaging Reporting and Data System (BIRADS) benign and malignant can be differentiated using its shape, size and density, which is how radiologist visualize the mammograms. According to BIRADS mass shape characteristics, benign masses tend to have round, oval, lobular in shape and malignant masses are lobular or irregular in shape. Measuring regular and irregular shapes mathematically is found to be a difficult task, since there is no single measure to differentiate various shapes. In this paper, the malignant and benign masses present in mammogram are classified using Hue, Saturation and Value (HSV) weight function based statistical measures. The weight function is robust against noise and captures the degree of gray content of the pixel. The statistical measures use gray weight value instead of gray pixel value to effectively discriminate masses. The 233 mammograms from the Digital Database for Screening Mammography (DDSM) benchmark dataset have been used. The PASW data mining modeler has been used for constructing Neural Network for identifying importance of statistical measures. Based on the obtained important statistical measure, the C5.0 tree has been constructed with 60-40 data split. The experimental results are found to be encouraging. Also, the results will agree to the standard specified by the American College of Radiology-BIRADS Systems.

Keywords:
Benign and Malignant Mammogram Masses, Feature Extraction, Weight Function, Statistical Measures, Neural Network, C5.0 Decision Tree Classifier

1. INTRODUCTION

The breast cancer is considered be one of the leading causes of death among women. It has been found that every 3 minutes a woman is diagnosed with breast cancer and every 13 minutes a woman dies from the disease. And one in 8 women may be diagnosed with breast cancer during her lifetime and 1 in 30 will die from it [1]. Various methods have been proposed for early detection and screening of breast cancers and the mammography is being considered as one of the most effective method [2, 3]. During diagnosis, microcalcifications and masses are being considered as two important early signs of the disease [4]. The masses are quite thin and often present in the dense areas of the breast tissue with smoother boundaries compared to microcalcification [5]. Based on the report by Breast Imaging Reporting and Data System (BI-RADS), the malignant masses are irregular in shape and benign masses are circular or oval in shape [6]. During the diagnosis phase, the masses are generally classified either as benign or malignant with a high biopsy yield ratio.

In medical viewpoint, reading visually and interpreting mammograms is considered as a complex task for radiologists. Their judgments are essentially depends on training, experience and other subjective parameters. The Computer Aided Diagnosis (CAD) systems have been developed to complement radiologists for the analysis of mammograms to identify and detect masses and calcifications. It is observed that 65-90 percent of the biopsies of suspected cancers turned out to be benign. Thus, it is essential to develop CAD that can distinguish benign and malignant lesions [7, 8]. The combination of CAD scheme and experts knowledge of radiologists would improve the rate of detection accuracy of masses. The detection rate without CAD is found to be 80 percent and with CAD the detection rate is around 90 percent [7].

Among various methods available, a very few methods has the reasonable classification rate for classifying mass region as either benign or malignant. This is achieved by introducing some constraints such as not considering irregular masses and by using large number of features. However, using large number of features will not really increase the classification rate much but increase the classification complexity. The shape properties of the masses can also be used to classify the region as either normal or abnormal by using shape properties [9, 10]. Most of the early works on mammograms are done based on histogram of mammograms [11, 12] and it is noticed that histogram based approach is not an effective method for classifying the masses. This is due to the fact that the histogram pattern changes due to noise and over-enhancement of mammogram.

Researchers have proposed various features for classifying masses in mammograms. The statistical features like uniformity, smoothness, and third moments etc which utilize gray value or histogram of masses are used for classifying the masses [10-13]. However, the gray values of mammogram will change in case of over-enhancement or in presence of noise. In this paper, benign and malignant masses are classified using a HSV grey weight function based statistical measures; instead of grey pixel value based statistical measures. The mass as region of interest present in mammogram is considered for analysis and classification. Various statistical measures such as Mean, Standard Deviation, Smoothness, Uniformity and Coefficient of Variation are used, which are based on grey weight value of the mass. Classification using these measures is found to be encouraging compared to other recently proposed methods.

The paper is organized as follows. In Section 2, presents the literature review. The gray weight function and the average thresholding method are discussed in Section 3. Gray weight value based statistical measures are discussed in Section 4. In Section 5, the experimental results are presented and conclude the paper at the end.

2. LITERATURE REVIEW

Computer Aided Detection (CAD) systems have been developed to complement radiologists for diagnosing breast cancer by analyzing and interpreting the mammograms. Several studies have proved that CAD improves breast cancer detection rate by 13.3 to 14.2% [14]. Previous approaches, which classify the abnormalities based on BI-RADS system, have been giving accurate results [15, 16]. The statistical methods yields less accuracy of maximum 70% compared to other decision and Neural Network (NN) based classifiers [17]. Various studies have been performed on shape characteristics of mammograms and they are able to classify masses either as normal or abnormal [10, 11].

Various texture features such as correlation, angle of second moment, inertia, inverse difference moment, sum average, sum entropy, difference entropy etc. for classifying masses [18]. These texture features are calculated using histogram constructed from mammogram, which varies considerably, if the mammogram is over-enhanced or if the mammogram contains noise. This approach classifies the masses as either abnormal or normal and considered only circumscribed masses, spiculated masses in abnormal mammogram category. However, this approach has not considered lobular, architectural distortion and other mammograms with complicated masses. Another ANN based approach has been proposed by Cascio et al [19] for classifying masses and extracted 12 features and out of 1236 masses considered for experiments only 10% of them are benign samples. In case, the benign samples are increased, malignant samples may get misclassified as benign. This is due to the fact that more number of benign masses will actually test the robustness of the system. In their work, obscured masses are not considered for classification as it has both benign and malignant mass characteristics. Retico et al [20] have extracted both shape and texture features for classification using NN. The experiments have used mammogram database with 109 malignant and 117 benign masses. However, the performance measures such as sensitivity and specificity are quite low. In addition, Mencattini et al [21] have used both texture and shape features for classifying the clusters of microcalcifications present in mammograms and the classification accuracy is not encouraging.

Similar to above approach, shape-based features has been used for classifying masses present in mammogram [22]. The training and learning phases has been incorporated and 60 training data sets and 25 test data sets have been used for training and learning the characteristics of the masses. However, it was not clear to note about the type of masses considered and the accuracy is also low. A Modular classifier has been proposed and used for classifying the masses using shape features of masses [23]. The age of patients is used as vital property for classifying malignant. This is due to the fact that the age of the patient has high discriminating power for malignancy. However, for the experiments, they have considered very less number of masses with complex shapes such as architectural distortion, microlobulated masses.

Descriptor based techniques has been proposed and extracted the encoded descriptor from DDSM databases for classifying masses with six encoded descriptors [16]. The age of the patient is also used to discriminate the benign from malignant. It is observed from the literature that the age has high discriminating power for malignancy. Zaiane et al [24] has proposed a scheme, which uses statistical measures for classifying the masses. This is an association rule based classifier with classification accuracy up to 80%. However, the statistical features tend to provide different results, in case the mammogram contains noise or if it is over-enhanced. The experiment consists of only medio-lateral oblique view mammograms. Rangayyan et al [25] have observed that various techniques and schemes tend to have good sensitivity, which is greater than 85% for the identification of masses. However, these methods have a high false positive rate. Statistical measures like mean, standard deviation, smoothness, entropy; uniformity has been used by [26] to classify breast tissues. The breast tissue is classified into 4 categories like fatty, dense etc. The accuracy obtained is 78%. However, the effects of these measures on masses are not studied. [27] used various shape descriptors along with statistical features like mean, standard deviation, smoothness, etc to classify masses and classification accuracy obtained was 80% for 57 mammograms.

In contrast to the above methods and schemes discussed, in this paper, HSV weight function based statistical features are used [28]. These statistical measures are considered as inputs the Neural Network and C5.0 decision tree classifier. The C5.0 classifier constructs rules in the form of if….then…else statements, which can be easily implemented.

3. GRAY WEIGHT FUNCTION AND AVERAGE THRESHOLDING

The content of a mammogram can be captured using only four levels of gray. While representing the masses in mammogram they are surrounded with pixels with slight variation in the gray values and forms smooth boundaries. Thus, it is imperative to capture the boundary information from pixel values for spatially segmenting the masses. The experiment is carried out on spatial domain processing using the properties of HSV color space and this color space is closely related to the human visual perception of color and gray pixels. For each pixel, a weighted value is calculated using a soft-decision function, which captures the degree of gray content of a pixel and is robust against noise and is given in Eq.(1).

$$W_{gray}(S,I) = 1 - S^{r_1(255/I)^{r_2}} \quad for\ !(R = G = B) \qquad (1)$$

The range of $W_{gray}(S, I)$ is [0-1] and it estimates the degree of gray content of a pixel using both the saturation and intensity values. From Eq. 1, it is noticed that the gray weight value holds for !($R = G = B$). For $R = G = B$, the saturation of a pixel is zero and the gray weight value will always be zero irrespective of the intensity value. However, in order to capture the degree of gray content of a pixel, slightly perturb either the value of R, or G or B, which influences the saturation value of a pixel. The function is smooth with saturation value and found to be continuous [29, 30] as shown in Fig.1. It is evident from Fig.1 that in Eq.(1), r_1 should take a value which is slightly higher than 0 and r_2 should take a value slightly less than 1 for having smooth variation gray level weighted values. From our earlier work [29, 30], it is found that the suitable values for r_1=0.1 and r_2=0.85.

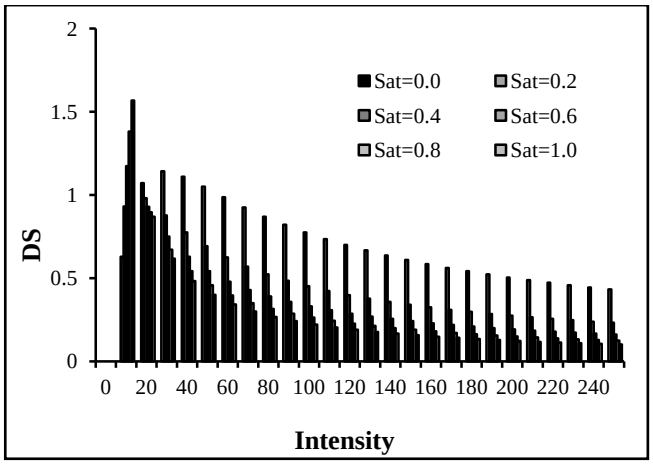


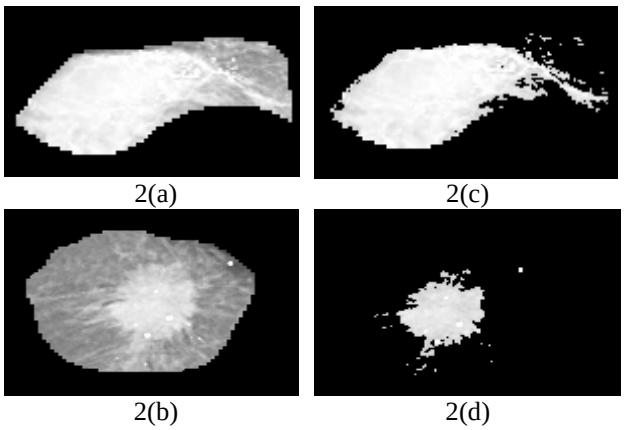
Fig.1. Variation of Partial Derivative of $W_{gray}(S, I)$ with S for different Values of Intensity with r_1=0.1, r_2=0.85

For each pixel of interested mass, the weighted value is calculated and the unwanted pixels are removed as a pre-processing step using average thresholding approach as given in Eq.(2). The average gray value of mass is calculated as $Gray_{Threshold}$ and pixels having lower value compared to $Gray_{Threshold}$ are removed from the mass. This pre-processing is considered as crucial step, as it removes noise and redundant background from the mass. The outcome of applying average threshold produces a precise mass, which can be used for further processing.

$$Gray_{Threshold} = \frac{\sum_{i=0}^{n} Mass_{gray}}{n} \quad (2)$$

where, n is total number of pixels in mass and $Mass_{gray}$ is the value at each pixel.

Fig.2 shows some of the sample mammograms and extracted mass after average thresholding. The mass after average thresholding is shown in Fig.2(c,d).



2(a) 2(c)

2(b) 2(d)

Fig.2. Sample Mammogram Masses 2(a,b) Extracted Mass. 2(c,d)After Average Thresholding

From Fig.2(c,d), it is observed that the preprocessed mass has clear boundary and features can be extracted for classification.

4. FEATURES OF MASSES USING STATISTICAL MEASURES

In our approach, the features of mass extracted using statistical measures are presented in Table.1. The effectivenesss of these statistical measure is evident from the Fig.3 mean measure.

Table.1. Statistical Measures

Measure	Formula
Area	n = Number of pixels in the mass
Mean	$m = \left(\frac{\sqrt{\sum (W_{gray}(S, I))}}{n} \right)$
Std Deviation	$\sigma_x = \left(\sqrt{\sum (W_{gray}(S, I) - m)^2 * P(W_{gray}(S, I))} \right)$
Smoothness	$S_x = \left(1 - \frac{1}{1 + \sqrt{\sum (W_{gray}(S, I) - m)^2 * P(W_{gray}(S, I))}} \right)$
Uniformity	$U_x = \left(\sum P^2 (W_{gray}(S, I)) \right)$
Coefficient of Variation	$CV = \left(\frac{\sigma}{m} \right)$

where, *n* is total number of pixels in mass, m is the mean gray weight value of mass and $P(W_{gray}(S, I))$ is the histogram of the $W_{gray}(S, I)$ values present in the mass.

The standard deviation measure is used to calculate the deviation of $W_{gray}(S, I)$ values of the pixels of the mass. For benign masses, the deviation is high compared to malignant.

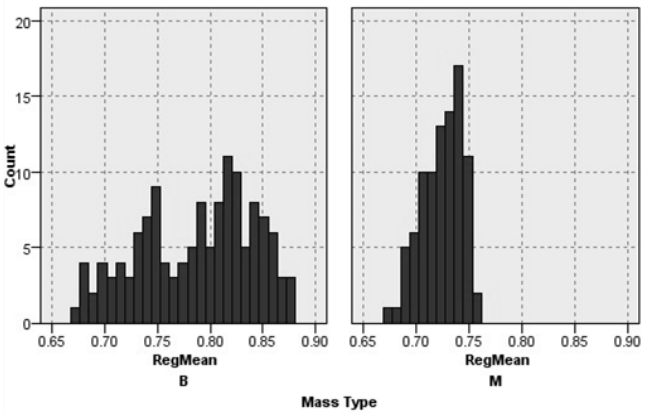


Fig.3. $W_{gray}(S, I)$ Mean for Benign and Malignant masses

Fig.3 shows that the value of statistical measures for benign mass is clearly distinct against the value of statistical measure of malignant mass. Thus, it can easily discriminate the benign masses from malignant masses. From Fig.3, it can be observed that mean of malignant masses ranges from 0.68 to 0.76. Similarly, for benign masses it ranges from 0.68 to 0.89, while most of them are near 0.88. Mean is considered as one of the best measures for discriminating masses. This is due to the fact

value of mean is calculated using weight function has high discriminating characteristics.

In Fig.4, the standard deviation of malignant masses and benign masses are shown.

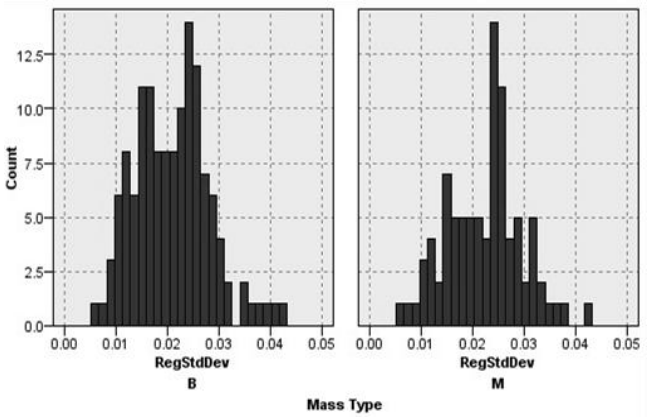


Fig.4. Standard Deviation for Benign and Malignant masses

The smoothness measure of mass is shown in Fig.5. While measuring the smoothness, the values for benign and malignant masses are more or less similar, and it may not form the cluster like mean. The similar effect is shown in Fig.6 and Fig.7 for uniformity and coefficient of variation of the benign and malignant masses respectively.

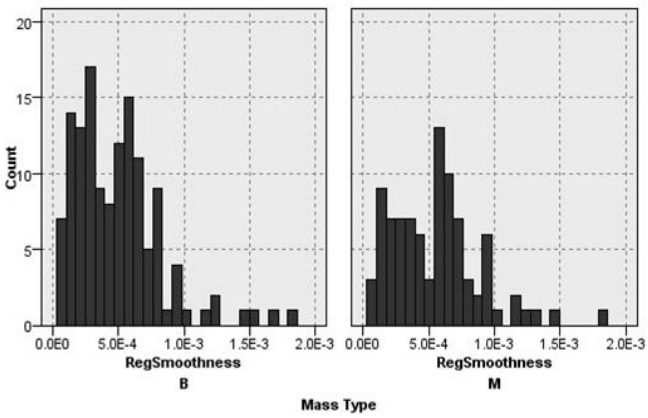


Fig.5. Smoothness for Benign and Malignant masses

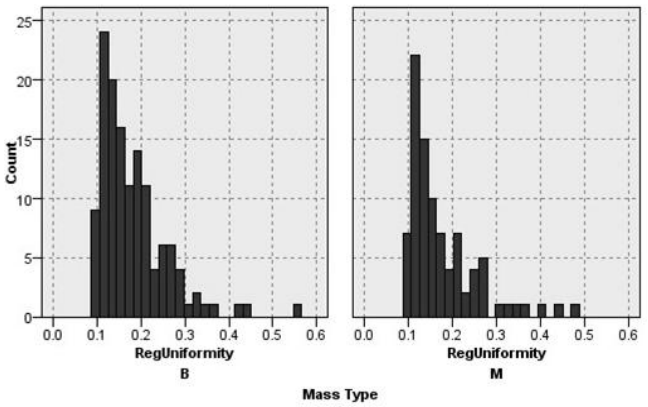


Fig.6. Uniformity for Benign and Malignant masses

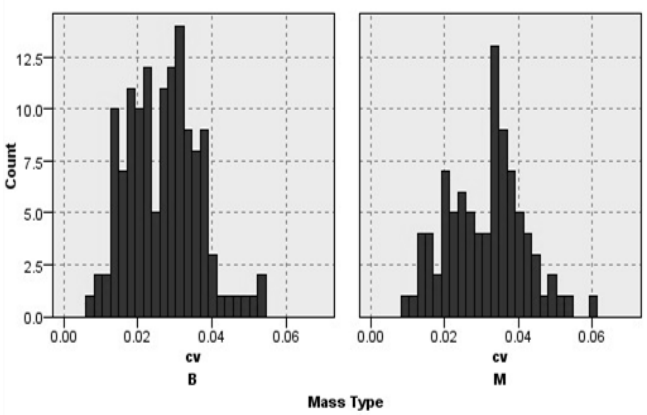


Fig.7. Coefficient of Variation of Benign and Malignant masses

It is observed from all the above results that the weight function based statistical measures such as mean, standard deviation, smoothness, uniformity and coefficient of variation can be used for classifying benign and malignant masses effectively.

5. EXPERIMENTAL RESULTS

Mammogram images from the Digital Database for Screening Mammography database (DDSM) (http://marathon.csee.usf.edu/Mammography/DDSM) are used for carrying out experiments [31]. For experiments 223 sample mammograms from the DDSM database has been considered and out of which 90 are malignant masses and 133 are benign masses. The statistical measures such as mean, standard deviation, smoothness, uniformity and coefficient of variation are shown in Table.2 for four sample cases.

Table.2. Statistical Measures for various mammogram masses

Images	**Area (*n*)**	**Mean (m)**	**(*σ*)**	**(*S*)**	**(*U*)**	**(*CV*)**
Benign Masses						
B1	215	0.81	0.03	0.0008	0.11	0.03
B2	224	0.83	0.04	0.0013	0.13	0.04
B3	467	0.81	0.03	0.0008	0.13	0.04
B4	517	0.78	0.03	0.0006	0.13	0.03
Malignant Masses						
M1	272	0.69	0.02	0.0004	0.14	0.03
M2	325	0.73	0.02	0.0005	0.14	0.03
M3	1592	0.67	0.01	0.0001	0.35	0.01
M4	1480	0.69	0.01	0.0001	0.31	0.01

The standard deviation is a measure, which gives the amount of deviation from the mean value. Higher the standard deviation more volatile or the data is spread about the mean. The smoothness is the contour smoothness of mass, where malignant will not be having smooth margin compared to benign. This is evident from the Table.2 that smoothness value of the benign masses is smoother than malignant. The uniformity measures the

uniformity of $W_{gray}(S, I)$ values in the mass and Coefficient of Variation is the ratio standard deviation to mean. From the above table, it is clearly evident that these statistical measures can be used to discriminate or classify the benign and malignant masses effectively with less error.

5.1 CLASSIFICATION OF MASSES

For classification, the dataset consists of 90 malignant and 133 benign samples and its distribution is shown in Fig.8. The Neural Network and C5.0 decision tree algorithm is used as classification methods.

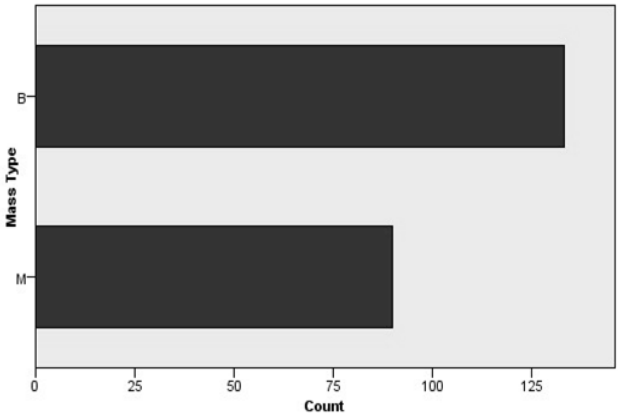


Fig.8. Benign and Malignant masses

C5.0 ALGORITHM:

C5.0 builds decision trees from a set of training data in the same way as ID3, using the concept of information entropy. The training data is a set $S = S_1,S_2,...S_n$ of already classified samples. Each sample $S_i = x_1,x_2,... x_k$ is a vector where $x_1,x_2,... x_k$ represent attributes or features of the sample. The training data is augmented with a vector $C = c_1,c_2,... c_m$ where $c_1,c_2,... c_m$ represent the class to which each sample belongs.

At each node of the tree, C5.0 chooses one attribute of the data that most effectively splits the set of samples into subsets enriched in one class or the other. Its criterion is the normalized information gain, which is the difference in entropy that results from choosing an attribute for splitting the data. The attribute with the highest normalized information gain is chosen to make the decision. The C5.0 algorithm then recurses on the smaller sub lists.

This algorithm has a few base cases as given below,

- All the samples in the list belong to the same class. When this happens, it simply creates a leaf node for the decision tree saying to choose that class.
- None of the features provide any information gain. In this case, C5.0 creates a decision node higher up the tree using the expected value of the class.
- Instance of previously-unseen class encountered. Again, C5.0 creates a decision node higher up the tree using the expected value.

5.2 DESIGNED SYSTEM MODEL

For experimental results, PASW modeler data mining software is used for designing Neural Network and C5.0 decision tree classifier with various training – testing data splits. The system model designed is shown in Fig.9. The settings used for various classifiers are discussed in the next section.

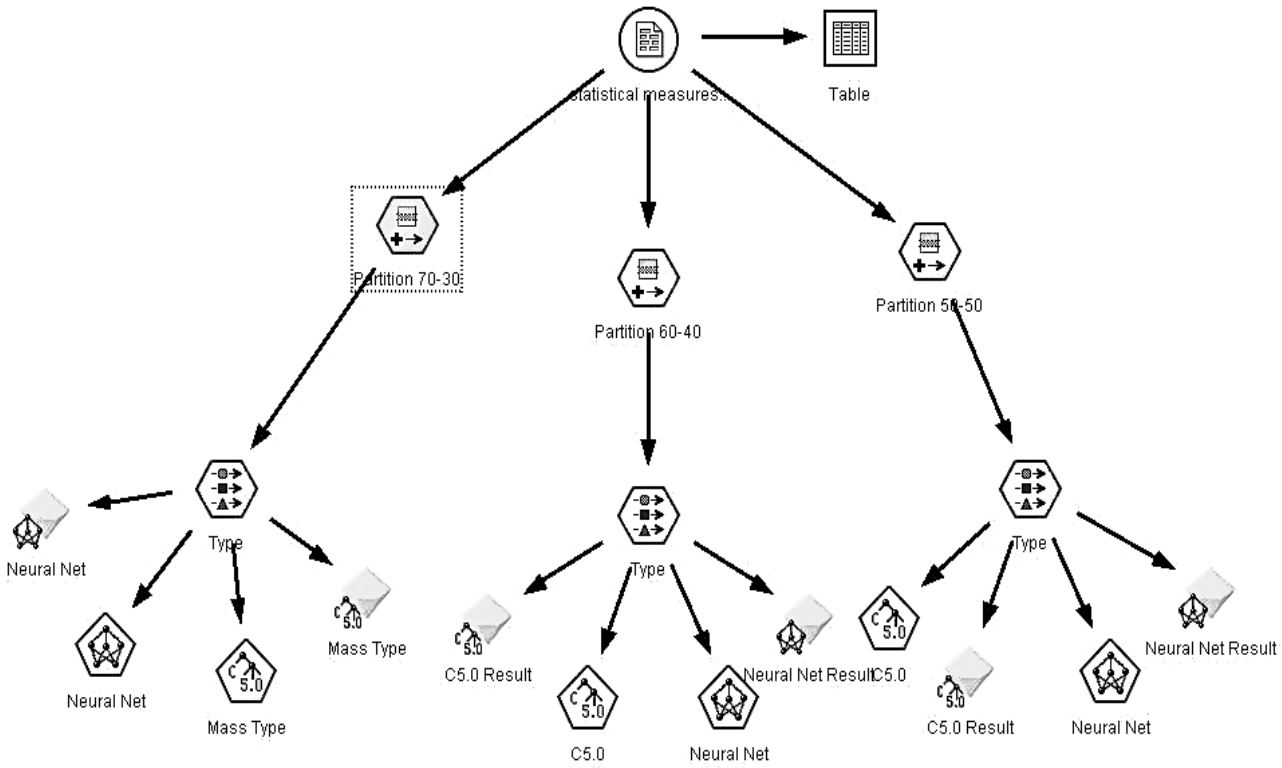


Fig.9. Designed System Model

The design model shown above consists of various nodes for different tasks and the functionality of each node is explained below.

Partition node: Partition node is used to generate a partition field which splits the dataset into separate subsets for the training and testing the models. Here, the dataset was partitioned by the ratio 50:50%, 60:40% and 70:30% for training and testing subsets respectively.

C5.0 node: C5.0 node is Decision Tree (DT) model which is trained without boosting to test the model accuracy. 10 fold cross validation is done. The C5.0 node builds either a decision tree or a rule set. The model works by splitting the sample based on the field that provides the maximum information gain at each level. The target field must be categorical. Multiple splits into more than two subgroups are allowed. Fig.11 illustrates that only 1 attribute is required to predict the diagnosis with this degree of accuracy. This is due to the fact that the HSV based statistical measures combined the basic shape descriptor area was able to classify the masses with high degree of accuracy.

NN node: NN node is used to train the NN model with quick method optimized for memory. The default settings of the model are used for NN classifier. Details about NN model is described below.

5.3 NEURAL NETWORK

The neural network model used for classification is shown in Fig.10. The model consists of 6 input neurons, 2 hidden layers and 1 output neuron. The same model is used for various training – testing data splits also. Default number of cycles is 200. The accuracy obtained is 95.238% for classifying the masses as benign or malignant. This is due to the fact that weight function statistical measures have good discriminating characteristics for benign and malignant masses (presented in Table.1).

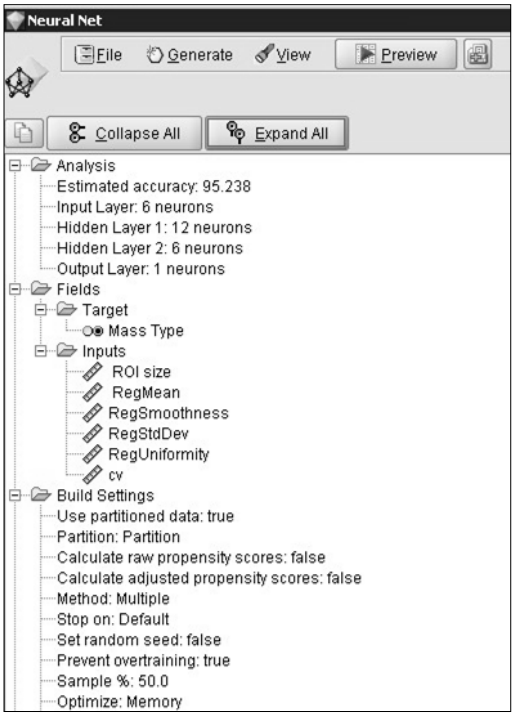


Fig.10. Neural network model used for 70-30 data split

The variable importance obtained as a result of neural network classification is shown in Fig.11. It is found that mean, smoothness and uniformity has higher importance compared to other measures. Also, the obtained result is better than mass size characteristic described in BI-RADS system. The effect of gray weight based statistical measures is evident from Fig.11 and also the mean is found to be the best discriminating feature for classifying benign from malignant masses.

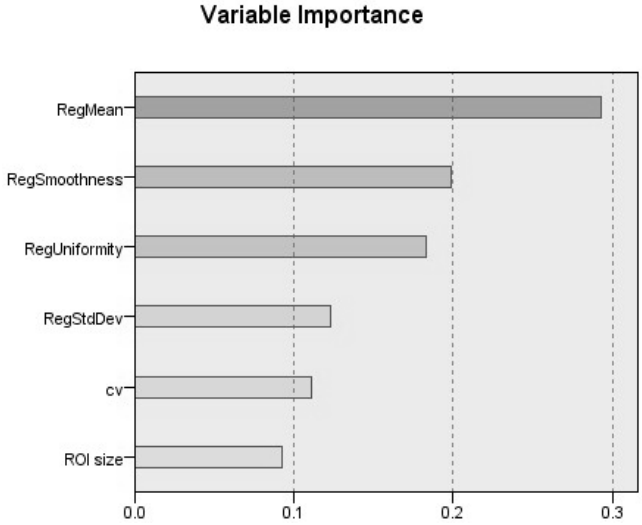


Fig.11. Variable importance by NN for 70-30 data splits

The decision tree constructed using C5.0 algorithm is shown in Fig.12 below. The root node identifies 86 mass as benign and 59 as malignant using mean of the mass. Based on these values, the nodes below take decision with the value of 0.754. The classification result for various training-testing data split is shown in Table.3. It is found that mean alone can discriminate the masses effectively.

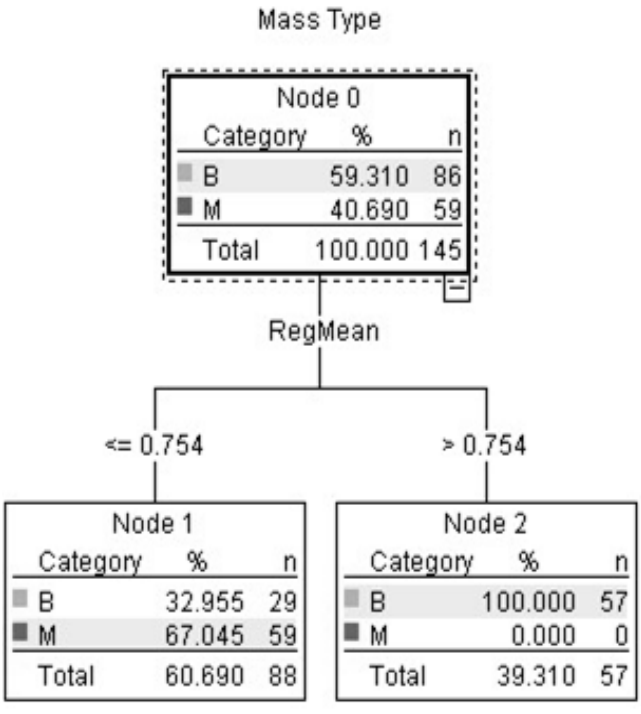


Fig.12. C5.0 Decision Tree for 60-40% data splits

Table.3. Performance comparison

Classifier	Training-Testing (50-50)	Training–Testing (60 - 40)	Training-Testing (70 - 30)
C5.0	1 rule If mean<0.775 Then Malignant Elseif mean>0.775 Then Benign	1 rule If mean<0.754 Then Malignant Elseif mean>0.754 Then Benign	3 rules
C5.0 Variable Importance	Mean	Mean	Mean, Mass size, uniformity
Neural Network	93.33%	94.28%	95.238%
NN Variable Importance	Uniformity, mean, CV, smoothness, stddev, Mass size	mean, smoothness, stddev, CV, uniformity, Mass size	mean, smoothness, uniformity , stddev, CV, Mass size

6. CONCLUSION

Classifying mammogram masses either as malignant or benign is considered to be a difficult task. A HSV based grey weight function has been used for measuring statistical features of masses in mammograms such as like mean, standard deviation, smoothness, uniformity and coefficient of variation. It shows that statistical measures are highly discriminative for benign and malignant the classification has been carried by constructing Neural Network and C5.0 models. The variable importance has been identified using neural Network and it is found to be mean. Using the mean, C5.0 has been constructed for 60-40 data split. The classification results are found to be encouraging. In future, we will use these measures to construct a fuzzy membership function for classifying masses as either benign or malignant.

ACKNOWLEDGEMENT

The work done by Dr. A.Vadivel is supported by the research grant from the Department of Science and Technology India, Under Grant SR/FTP/ETA – 46/07 dated 25th October, 2007 and DST/TSG/ICT/2009/27, dated: 03-09-2010.

REFERENCES

[1] N. C. I., "Cancer stat fact sheets: Cancer of the breast", http://seer.cancer.gov/statfacts/html/breast.html, (accessed on Mar 2011), 2011.

[2] ACS, "Mammograms Remain Best Way to Spot Breast Cancer", American Cancer Society (ACS), 2008.

[3] ACS, "Learn about breast cancer", http://www.cancer.org, 2011.

[4] Cheng H.D, X.P. Cai, X.W. Chen, L.M. Hu and X.L. Lou, "Computer-aided detection and classification of microcalcifications in mammograms: a survey", *Pattern Recognition*, Vol. 36, pp. 2967–2991, 2003.

[5] Wolfe J.N. "Breast patterns as an index of risk for developing breast cancer", *American Journal of Roentgen*, Vol. 126, pp. 1130–1139, 1976.

[6] American College of Radiology (ACR), "Breast imaging reporting and data system (BI-RADS), breast imaging atlas", 4th ed., Reston, Via: American College of Radiology, from http://www.acr.org/, 2010.

[7] Doi K., "Computer-aided diagnosis: potential usefulness in diagnostic radiology and telemedicine", *Proceedings of National Forum Military Telemedicine On-Line Today, Research, Practice and Opportunities*, pp. 9–13, 1995.

[8] Egan R.L and R.C. Mosteller, "Breast cancer mammography patterns", *Cancer*, Vol. 40, pp. 2087–2090, 1977.

[9] Flores B.A and J. A. Gonzalez, "Data Mining with Decision Trees and Neural Networks for Calcification Detection in Mammograms", *Third Mexican International Conference on Artificial Intelligence, Proceedings –LNCS*, pp. 232-241, 2004.

[10] Sun Y, Babbs C and Delp E, "Normal Mammogram Classification Based on Regional Analysis", *Proceedings of the IEEE Midwest Symposium on Circuits and Systems*, Vol.45, pp.375-378, 2002.

[11] Vibha L, Harshavardhan G. M, Pranaw K, Deepa Shenoy P, Venugopal K. R and Patnaik L. M, "Classification of Mammograms Using Decision Trees", *10th International Database Engineering and Applications Symposium (IDEAS'06). IEEE*, pp. 263-266, 2006.

[12] Sheshadri H.S and A. Kandaswamy, "Experimental investigation on breast tissue classification based on statistical feature extraction of mammograms", *Computerized Medical Imaging and Graphics 31*, pp. 46–48, 2007.

[13] Surendiran B, Vadivel A and Selvaraj H, "A Soft-Decision Approach for Microcalcification Mass Identification from Digital Mammogram", *Proceedings of world academy of Science, Engineering and Technology*, Vol 36, pp.1236-1240, 2008.

[14] "Computer-aided Detection Improves Early Breast Cancer Identification", American Journal of Roentgenology, http://www.ajronline.org. [accessed on Mar 2011], 2011.

[15] Markey M. K., Lo J. Y, Tourassi G. D and Floyd C. E, "Cluster analysis of BI-RADS descriptions of biopsy-proven breast lesions", In: *Medical Imaging: Image Processing, Proceedings of SPIE*, Vol. 4684,pp. 363-370, 2002.

[16] Sampat M.P, Alan C, Bovik B and Markey M.K, "Classification of mammographic lesions into BI-RADS shape categories using the Beamlet Transform", *Medical Imaging: Image Processing, Proc. of the SPIE*, Vol. 5747, pp.16-25, 2005.

[17] Zhang P, Kumar K and Verma B, "A Hybrid Classifier for Mass Classification with Different Kinds of Features in Mammography. FSKD" Vol. 2, pp. 316-319, 2005.

[18] Bovis K and S. Singh, "Detection of masses in mammograms using texture features", *Proceedings of the 15th International Conference on Pattern Recognition, ICPR00*, pp. 267–269, 2000.

[19] Cascio D, Fauci F, Magro R, Raso G, Bellotti R, De Carlo F, Tangaro S, De Nunzio G, Quarta M, Forni G, Lauria A,

Fantacci ME, Retico A, Masala GL, Oliva P, Bagnasco S, Cheran SC and Torres EL, "Mammogram segmentation by contour searching and mass lesions classification with neural network", *IEEE Transactions on Nuclear Science*, Vol. 53, No. 5, pp. 2827–2833, 2006.

[20] A. Retico, P. Delogu, M.E, Fantacci and P. Kasae, "An automatic system to discriminate malignant from benign massive lesions on mammograms", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Vol. 569, No.2, pp. 596-600, 2006.

[21] Mencattini A, Rabottino G, Salmeri M, Caselli F, Lojacono R and Frondizi G, "Features Extraction to Classify Microcalcification Clusters in Mammography", *16th IMEKO TC4 Symposium, Exploring New Frontiers of Instrumentation and Methods for Electrical and Electronic Measurements*, pp. 22–24, 2008.

[22] Ertas, G, Guelcur H. O, Aribal E and Semiz A, "Feature Extraction from Mammographic mass Shapes and Development of a Mammogram Database", *Proceedings of 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 3, pp. 2752 – 2755, 2001.

[23] Shah V., L. Bruce and N. Younan, "Applying Modular Classifiers To Mammographic Mass Classification", *Proceedings of the IEEE Eng. in Medicine and Biology Society*, Vol. 3, pp. 1585 – 1588, 2004.

[24] Zaiane O., Maria-Luiza A and Alexandru C., "Mammography classification by an association rule-based classifier". *Proceedings of the Third International Workshop on Multimedia Data Mining*, pp. 62-69, 2002.

[25] Rangayyan R.M., F.J. Ayres and J.E.L. Desautels, "A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs", *Journal of the Franklin Institute*, Vol. 344, pp. 312-348, 2007.

[26] Sheshadri H. S and Kandaswamy A, "Breast Tissue Classification Using Statistical Feature Extraction Of Mammograms". Medical Imaging and Information Sciences, Vol. 23, No.3, pp. 105–107, 2006.

[27] Rojas.A and A K Nandi, "Detection of masses in mammograms via statistically based enhancement, multilevel-thresholding segmentation, and region selection", *Computerized Medical Imaging and Graphics*, Vol. 32, No.4, pp. 304-315, 2008.

[28] Surendiran .B and A.Vadivel, "An Automated Classification of Mammogram Masses using Statistical Measures", In *Proc. of 4th Indian International Conference on Artificial Intelligence (IICAI-09)*, pp. 1473-1485, 2009.

[29] Vadivel.A, Shamik Sural and A.K. Majumdar, "An Integrated Color and Intensity Co-Occurrence Matrix", *Pattern Recognition Letters, Elsevier Science*, Vol. 28, No.8, pp. 974-983, 2007.

[30] Vadivel A., Shamik Sural and A.K. Majumdar, "Robust Histogram Generation from the HSV Space based on Visual Colour Perception". *International Journal of Signal and Imaging Systems Engineering*, Vol. 1, pp. 245 – 254, 2008.

[31] Rose C., D. Turi, A. Williams, K. Wolstencroft, and C. Taylor, "Web services for the DDSM and digital mammography research", International Workshop on Digital Mammography, Vol. 4046, pp. 376–383, 2006.