AGU International Journal of Science and Technology
(AGUIJST) 2017, Vol. No. 5, Jul-Dec

http://www.aguijst.com
e-ISSN: 2455-0493; p-ISSN: 2455-3808

Comparative Analysis Classifiers for Abdominal Mass Detection

Shivshankar Sambhajirao Kore[1], Dr. Ankush B. Kadam[2]

[1]*Research Scholar, Pacific Academy of Higher Education and Research, University, Udaipur, (India)*
[2]*Assistant Professor, Jawahar Arts, Science and Commerce College, Andoor.Maharashtra.*

ABSTRACT

One of the practical ways of testing the internal organs is through Ultrasound image, however, the unprocessed image usually embedded with noises, which leads to be more tedious in granting clear view of the particular region that has affected. This paper aims in developing an advanced model for abdominal mass diagnosing with US images. The proposed detection model includes two phases: (i) Feature extraction and (ii)Classification. In the first phase, the texture features are extracted using Adaptive Gradient Location and Orientation histogram (AGLOH). Then, the second phase uses Linear Collaborative Discriminant Regression Classification (LCDRC) model to classify whether the given image is normal or abnormal. Further, this paper adopts an improved diagnosis precision while detecting mass that presents in abdominal regions. Furthermore, the proposed LCDRC classifier compares its performance over other conventional techniques include Support Vector Machine(SVM), Neural Network (NN), K-Nearest Neighbor (K-NN), Naive Bayes (NB) in terms of measures like Accuracy, Sensitivity, Specificity, Precision, False Positive rate (FPR), False Negative rate (FNR), Negative Prediction Value (NPV), False Discovery rate (FDR), Matthews correlation coefficient (MCC)and F_1Score, and the betterment of proposed method is proven.

Keywords—Abdominal Mass Detection; Ultrasound image; LCDRC; Traditional Classifiers

I.INTRODUCTION

The abdominal mass is normally the extension of any human anatomy that normally varies as per the position it located. Mostly the mass occurs by the hepatomegaly, splenomegaly, etc. The masses are unpredictable one, which could be found from the usual physical consideration or the prediction can be achieved from the US images of abdominals. The formation of an abdominal mass is clinically analyzed as the cyst formation in any of the human organ, which often causes by multicystic in kidneys. Usually, the size of cyst varies small to medium tissue.

Initially, the investigation of abdominal mass is achieved by an imaging module named Plain abdominal Radiographs. After some evaluation, radiography has become key criteria in determining the location of mass along its density. In any detection strategy, the classification [3] is the major role since it classifies the presenting tumors, and some of them are teratomas and lithiasis. More classifiers are there such as ANN[15], Bayes classifier, etc.which makes a possible way in detecting accurate diagnosis. Next, to this, more attention is payed in imaging modalities that used for detecting the mass.A number of imaging models are there like

Magnetic resonance imaging (MRI) [11], Computed Tomography' CT, x-rays. Further, US imaging [10]is the major imaging concept for analyzing the physical characteristics. More important fact is that number of contributions was already presented in cancer detection [1] [2] [4] [5] with US imaging, however; only restricted contributions are there in diagnosing abdominal masses. This fact motivates the researchers to develop the abdominal mass diagnosing model with US image.

This paper develops a new model of diagnosing normal and abnormal abdominal US images. As mentioned earlier, two phases are there: Feature Extraction and Classification. AGLOH method is adopted to extract the features andLCDRC method is used to achieve the classification process. Further, the performance of proposed classification approach is compared to existing models like SVM, NN, KNN and Naïve Bayes.

The rest of the paper is arranged as follows. Section II reviews the related works. Section III describes the proposed abdominal mass detection model. Section IV demonstrates the feature extraction and classification techniques. Section V explains the conventional classifier models. Section VI discusses the results obtained, and section VII concludes the paper.

II.LITERATURE REVIEW

2.1Related Works

In 2017, T.Tan*et al.* [1] has investigated the impact of Computer Aided Detection (CAD) software to detect the breast cancer automatically in automated breast ultrasound (ABUS). The examination was carried out by analyzing 90 patients. The results have shown that CAD-software for ABUS has made gradual enhancement in the rate of cancer detection. In 2017, JinzuJi *et al.* [2] have proposed a novel method that was on the basis of elastography for the purpose of identifying tumor. Finite-difference time-domain (FDTD) approach was utilized in simulation purpose. In 2015, S. Beura *et al.* [3] has developed an effective model for classifying mammograms to detect the breast cancer. The developed has utilized Two-dimensional Discrete Orthonormal S-transform (DOST) for extracting the feature from mammograms, and consequently, the classification was performed with the utility of chosen coefficients. The results have shown the superiority of proposed model.

In 2017, M. D. Hossain and A. S. Mohan [4]have proposed a new Coherent-Beam space-Time Reversal-Maximum Likelihood (C-B-TR-ML) model for obtaining accurate tumor locations with less computational problem. Then, the obtained results have reviewed the enhanced performance of developed model. In 2008, A. Mencattini *et al.* [5] have presented a new algorithm to de-noise the image de-noising to detect the mass that present, a novel segmentation method was developed that has the combination of dyadic wavelet information with mathematical morphology. Further, the approach was evaluated with huge clinical images and has proved its efficiency in early detection of breast cancer over the conventional models.

2.2 Review

Various cancer detection approaches were analyzed, where CAD [1] enhances the detection of cancer, and however, more impacts are there in individual factors. FDTD [2] is more reliable, and DOST [3] gives better performance in terms of significant level, but it higher the overhead by selecting additional features. C-B-TR-ML [4] is more accurate in detecting tumors, but the required computational cost is more. Dynamic wavelet transform [5] can enhance the features. However, more enhancements are needed in decision phase to get the accurate results. Hence, from all these discussions, it is clear that there is a need for an effective model in abdominal mass detection since all the contributions are mainly focused on cancer detection.

III.PROPOSED ABDOMINAL MASS DIAGNOSIS MODEL

The proposed method detects the mass that presents in abdominal region from the US image. Both the normal and abnormal images of abdominal region are given for processing. The proposed model has two stages: Feature extraction and classification. The architecture of the developed AGLOH-LCDRC method is described in Fig. 1.

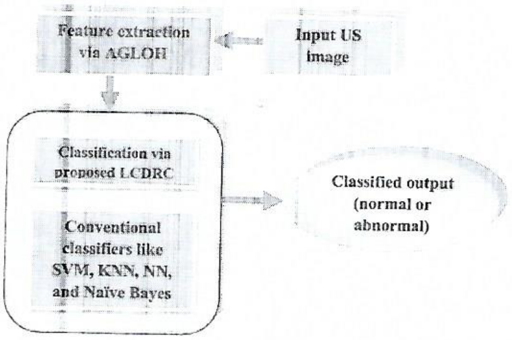


Fig. 1.Architecture of proposed model

PROPOSED FEATURE EXTRACTION AND CLASSIFICATION MODEL

3.1Feature extraction using Adaptive GLOH

This paper adopts adaptive GLOH [27] mechanism to extract the features from input US image, $U_I(x,y)$. This paper takes on adaptive bilinear filtering to smooth the texture of pre-processed image $U_I(x,y)$. The parameter $p(U_I(x,y))$ is given in Eq. (1).

$$p(U_I(x,y)) = z_{11}xy + z_{10}x + z_{01}y + z_{00} \qquad (1)$$

Hence, the evaluated filter image F_I is the resulting image, from which the features are extracted using GLOH descriptor. Further, GLOH descriptor grid comprises R circular ring, which centers the feature points. The rings include r regions that are consistently contributed with S directions. Let the region be RE_{uv}, $RE_{u,v} \in F_I$ with $u = 1,2,....,R$ and $v = 0,1,....,S-1$.

Furthermore, the block histogram is defined as shown in Eq. (2), where, $\oplus$ refers to the concatenation operator, from which the final descriptor vector, VE is attained by concatenating histograms, which is shown in Eq. (3). Moreover, the descriptor length is defined in Eq. (4).

$$B_h(z) = \begin{cases} z & if \quad z < \frac{l}{2} \\ h - z & otherwise \end{cases} \qquad (2)$$

$$VE_{u,v} = \bigoplus_{r=0}^{S-1} b_{u,v}^{[v+r]S} \qquad (3)$$

$$h = R(RS + 1 + (S-1)\Psi(VE)) \qquad (4)$$

Principal
Jawahar Arts, Science & Commerce College,
Andur Tal. Tuljapur Dist, Osmanabad

The circulation of the descriptor δk is achieved by the factor achieves, which is given in Eq. (5), where $\delta k=\frac{2\pi}{S}$ consistently, where $VE_{0,k}$ is $VE_{0,k}=\bigoplus_{r=0}^{S-1} h_{0,v}^{[r+k+v]S}$.

$$VE_{\delta k}=\begin{Bmatrix} \bigoplus_{r=0}^{R}\bigoplus_{t=0}^{S-1} VE_{r,[k+t]S} & if\ \Psi(VE)=1 \\ VE_{0,k}\bigoplus_{r=1}^{R}\bigoplus_{t=0}^{S-1} VE_{r,[k+t]Q} & otherwise \end{Bmatrix} \quad (5)$$

The distance between features, VE and $\overline{VE}$ is defined in Eq. (6), where $\hat{DI}^T\left(VE,\overline{VE}\right)$ represents the assessment of usual distance. Hence, the extracted features from abdominal images is addressed as $GLOH^F=[Y_1,Y_2\cdots Y_K]\in\Re^{x\times y}$ where K denotes the total features.

$$\hat{DI}^T\left(VE,\overline{VE}\right)=\min_{k=0,1,\ldots S-1} DI^T\left(VE,\overline{VE_{\delta k}}\right) \quad (6)$$

3.2Classification using LCDRC

Consider $Y=[Y_1,Y_2\cdots Y_K]\in\Re^{x\times y}$ as the matrix of entire image with training image, where x denotes the dimension of all training images, $Y_i=[Y_{i1},Y_{i2}\cdots Y_{in_i}]\in\Re^{x\times y}$, the count of training image from i class is specified as y_i, and $y=\sum_{i=1}^{K} y_i$, $i=1,2,3\ldots K$ and K specifies the total classes. Consider a matrix $O\in\Re^{x\times d}$ with subspace and $d<x$. All u_{ij} should be subjected to learned subspace by $v_{ij}=O^T u_{ij}$, which lies between $1\le j\le y_i$.

The matrix with training image is allotted as defined in Eq. (7) and for all class; the same is defined in Eq. (8)

$$A=O^T Y\in\Re^{d\times y} \quad (7)$$

$$A_i=O^T Y\in\Re^{d\times y_i} \quad (8)$$

The CBCRE and WCRE are defined by Eq. (9) and Eq. (10),

$$CBCRE=\frac{1}{y}\sum_{i=1}^{K}\sum_{j=1}^{y_i}\left\|v_{ij}-\hat{v}_{ij}^{inter}\right\|_2^2 \quad (9)$$

$$WCRE=\frac{1}{y}\sum_{i=1}^{K}\sum_{j=1}^{y_i}\left\|v_{ij}-\hat{v}_{ij}^{intra}\right\|_2^2 \quad (10)$$

where $\hat{v}_{ij}^{inter}=A_{ij}^{inter}\alpha_{ij}^{inter}$ and $\hat{v}_{ij}^{intra}=A_{ij}^{intra}\alpha_{ij}^{intra}$. A_{ij}^{inter} is the A with A_i eliminated and A_{ij}^{intra} is the A_i with v_{ij} eliminated. Then α_{ij}^{inter} and α_{ij}^{intra} can be attained by Eq. (11).

$$\hat{\alpha}_i=\left(X_i^T X_i\right)^{-1} X_i^T y,\quad i=1,2\cdots K \quad (11)$$

At last, the value of CBCRE and WCRE is defined as Eq. (12) and Eq. (13).

$$CBCRE=tr\left(O^T e_b O\right) \quad (12)$$

$$WCRE=tr\left(O^T e_W O\right) \quad (13)$$

Principal
Jawahar Arts, Science & Commerce College,
Andur Tal. Tuljapur Dist. Osmanabad

 http://www.aguijst.com
e-ISSN: 2455-0493; p-ISSN: 2455-3808

To increase the value of CBCRE and to minimize the value of WCRE simultaneously, a Maximum Margin Criterion (MMC)is employed as in Eq. (14), which is took to a solution by spotting the maximized DI Eigenvalues and therespectiveEigenvectorsis in Eq. (15), where $\lambda_1 \geq \cdots \geq \lambda_k \cdots \lambda_{DI}$ and $O = [o_1, o_2, \cdots o_k \cdots o_{DI}]$.

$$\max_O J(O) = \max_O (CBCRE - WCRE) = \max_O \left(tr\left(O^T (e_b - e_w) O\right)\right) \quad (14)$$

$$(e_b - e_w) o_k = \lambda_k o_k, \quad k = 1, 2 \cdots DI \quad (15)$$

Further, MMC resolves the problem termed as Small Sample Size Problem (SSSP) where the measurement of abdominal image is higher than the amount of training images.

IV.CONVENTIONAL CLASSIFIERS

In this section, certain conventional classifiers like SVM[6] [14], NN, KNN [8] and Naïve Bayes [9]are explained. The proposed model, LCDRC compares its performance over these mentioned classifiers.

4.1SVM

SVM [6]is a machine learning approach that has the ability of deciphering subtle patterns either in noisy as well as complex datasets. Initially, SVM is designed for binary classification.

The primal equation of a soft margin SVM is defined in Eq. (16), where w denotes the normal vector of the splitting hyperplane in feature space and P specifies the regularization parameter that controls the penalty for misclassification.

$$\min_w \left\{ \frac{1}{2} \| w \|^2 + P \sum_i \xi_i \right\} \quad (16)$$

4.2NN classifier

NN [7]model is a multilayer feedforward network, which is trained to do the classification process. The network model is determined in Eq. (17), (18) and (19), where i denotes the hidden neuron, $WH^{(HI)}_{(Bli)}$ denotes the bias weight to i^{th} hidden neuron, nu_i specifies the count of input neurons, nu_{HI} represents the count of hidden neurons, $WH^{(o)}_{(B1m)}$ is the output bias weight to m^{th} layer, $WH^{(o)}_{(im)}$ denotes the output weight from i^{th} hidden neuron to m^{th} layer. The network output $\hat{N}_m$ is given in Eq. (34), where N_m is the actual output.

$$e^{(HI)} = NF\left(WH^{(HI)}_{(Bli)} + \sum_{j=1}^{nu_i} WH^{(HI)}_{(ji)} Input \right) \quad (17)$$

$$\hat{N}_m = NF\left(WH^{(o)}_{(B1m)} + \sum_{i=1}^{nu_{HI}} WH^{(o)}_{(im)} e_i^{(HI)} \right) \quad (18)$$

$$WH^* = \underset{\left\{WH^{(HI)}_{(Bi)}, WH^{(HI)}_{(ji)}, WH^{(o)}_{(B1m)}, WH^{(o)}_{(im)}\right\}}{\arg\min} \sum_{m=1}^{nu_o} N_m - \hat{N}_m \quad (19)$$

Principal
Jawahar Arts, Science & Commerce College,
Andur Tal. Tuljapur Dist, Osmanabad

4.3K-NN

K-NN [8] evaluates the distance among training as well as testing samples that present in dataset. It gives the k closest samples. The projection function wf is defined as in Eq. (20), where zo_j is the j^{th} column vector of ZO, $ZO_{(i)}$ specifies the submatrix of ZO and $G^{(ba)}(\cdot)$ refers to the kernel function with ba bandwidth.

$$wf = \frac{G^{(ba)}\left(l_i, zo_j\right)}{\sum_{j' \in ZO} G^{(ba)}\left(l_i, zo_{j'}\right)}, j \in ZO_{(i)} \tag{20}$$

1.1 Naive Bayes

The NB classifier [9]is a probabilistic approach for classification. For an unclassified object $OB = (ob_1, ob_2, \ldots\ldots ob_n)$, this classifier predicts the category of OB that it falls. Particularly, this classifies OB object into CA_i category only if it satisfies the posterior probability condition that defined in Eq. (21).

$$PO(CA_i \mid OB) > PO(CA_j \mid OB) \text{for all } j \neq i \tag{21}$$

With the aid of Bayes theorem, Eq. (21) is expressed as defined in Eq. (22).

$$PO(CA_j \mid OB) = \frac{PO(OB \mid CA_j)PO(CA_j)}{PO(OB)} \tag{22}$$

V.RESULTS AND DISCUSSIONS

5.1Simulation procedure

The developed abdominal mass detection approach was developed in MATLAB. Here a database with 20 sets was taken that has 12 normal images and 10 abnormal images. The proposed LCDRC was compared with SVM, NN, K-NN and NB classifiers. The analysis was done in terms of measures like accuracy, sensitivity, specificity, precision, FPR, FNR, NPV, FDR, F1-score, and MCC and the results that obtained was more promising in terms of diagnosing whether the image is normal or abnormal.

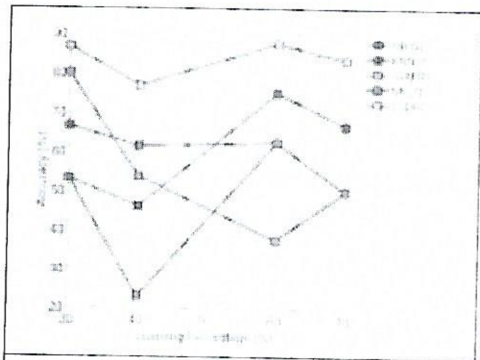
5.2Performance analysis

The performance of proposed LCDRC is compared to other conventional classifiers like SVM, NN, K-NN, and NB by varying the learning percentage by 30%, 40%, 50%, 60% and 70%, which is shown in Fig 2. From the analysis is observed that for 30%, the accuracy measure of proposed LCDRC is 66.10%, 10.23%, 27.80% better than the conventional classifiers like NN, SVM, and NB, respectively. Further, for 40%, the accuracy of proposed model is 70.56%, 44.88%, 62.31% and 28.25% superior to NN, SVM, K-NN, and NB, respectively. For 60% learning, the proposed model is 40.00%, 56.68%, and 20.81% better from the methods like NN, SVM and KNN, respectively. The sensitivity of proposed model for 40% is 30.83% and 26.165 better than SVM and NB classifiers, respectively.

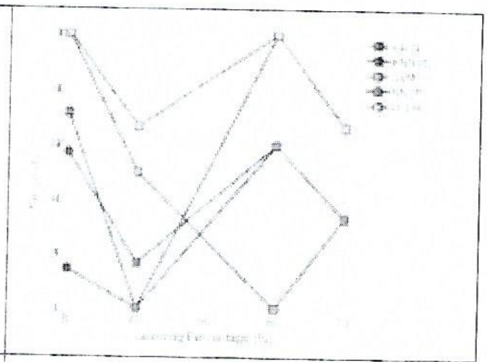
Similarly, the analysis is made of all the remaining measures, and the results have reviewed that proposed LCDRC is superior to other models in terms of accurate classification.

Principal
Jawahar Arts, Science & Commerce College,
Andur Tal. Tuljapur Dist. Osmanabad

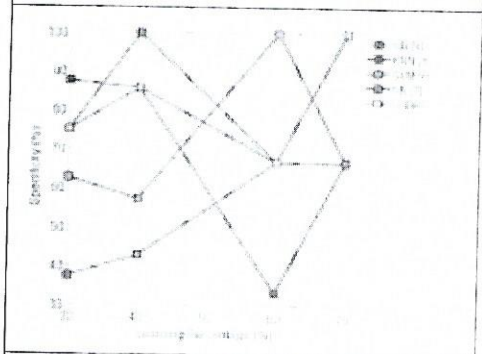
http://www.aguijst.com
e-ISSN: 2455-0493; p-ISSN: 2455-3808



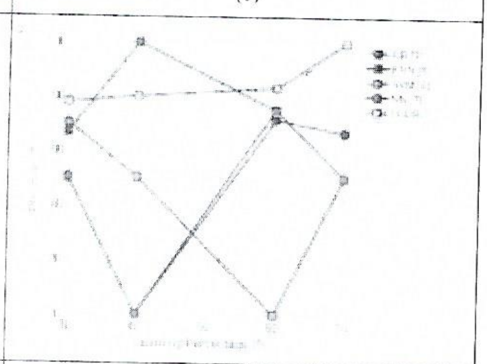
(a)



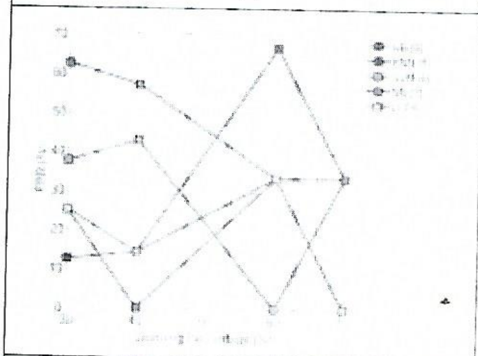
(b)



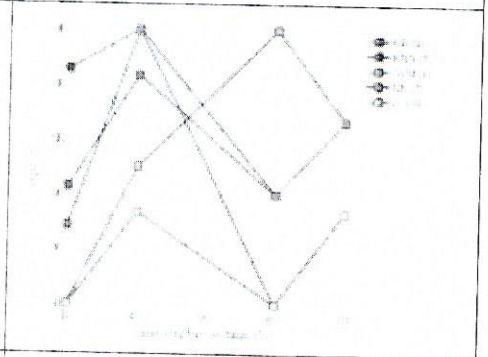
(c)



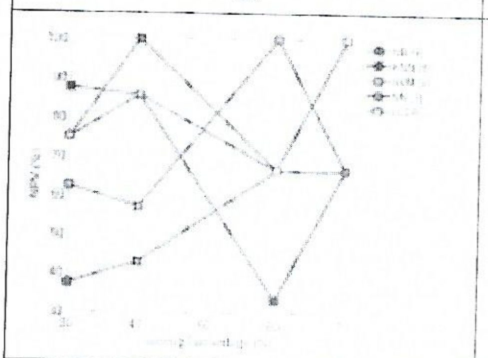
(d)



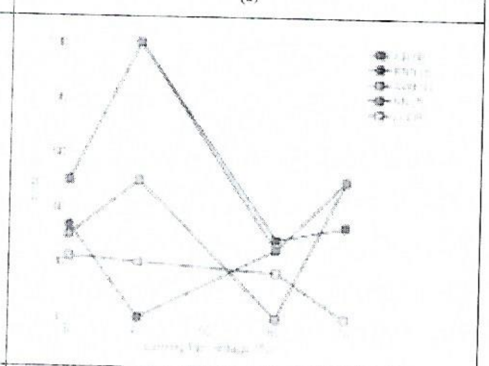
(e)



(f)

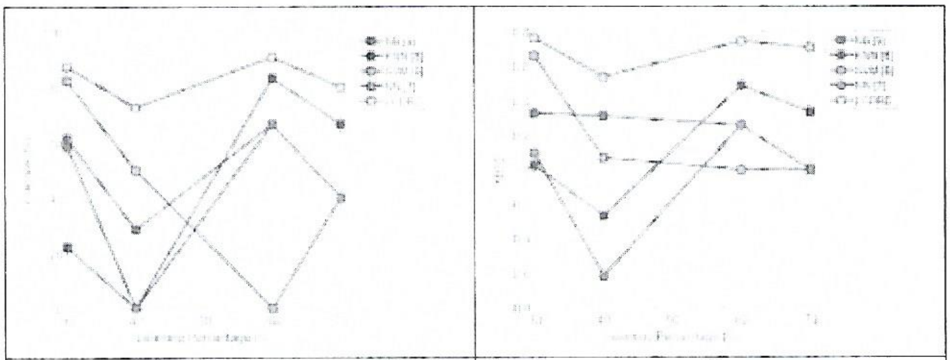


(g)



(h)

Principal
Jawahar Arts, Science & Commerce College,
Andur Tal. Tuljapur Dist. Osmanabad



(i)	(j)

Fig. 2 Comparison of proposed model over other classifiers by varying the learning percentage by 30%, 40%, 50%, 60% and 70% (a) Accuracy (b) Sensitivity (c) Specificity (d) Precision (e) FPR (f) FNR (g) NPV (h) FDR (i)F_1Score (j) MCC

TABLE 1 OVERALL PERFORMANCE OF SEVERAL CLASSIFIERS

Measures	NB [9]	KNN [8]	SVM [6]	NN [7]	LCDRC
Accuracy	0.666667	0.533333	0.8	0.533333	0.866667
Sensitivity	0.571429	0.142857	1	0.714286	1
Specificity	0.75	0.875	0.625	0.375	0.75
Precision	0.666667	0.5	0.7	0.5	0.777778
FPR	0.25	0.125	0.375	0.625	0.25
FNR	0.428571	0.857143	0	0.285714	0
NPV	0.75	0.875	0.625	0.375	0.75
FDR	0.333333	0.5	0.3	0.5	0.222222
F1-score	0.615385	0.222222	0.823529	0.588235	0.875
MCC	0.327327	0.026207	0.661438	0.094491	0.763763

Table I summarizes the overall performance of proposed model over other classifiers. Here, the proposed model has attained better accuracy rate over other methods, which is 62.50%, 8.33%, 62.50%, and 29.99% better from the methods like NN, SVM, K-NN, and NB, respectively. The sensitivity of proposed model is high than the other approaches, and the proposed model is 39.99%, 85.71%, and 42.85% better from NN, K-NN, and NB, respectively. The specificity of proposed model is 50%, 20%, and 14.28% better from NN, SVM, and K-NN, respectively. The precision of proposed LCDRC is 55.55%, 11.11%, 55.56%, 16.66% better than NN, SVM, K-NN, and NB, respectively. Thus, the overall performance shows that proposed LCDRC has achieved better performance than the conventional classifiers in terms of better diagnosing of mass.

VI.CONCLUSION

This paper has offered an enhanced approach for diagnosing abdominal masses from US images. In this method, two phases were there: Feature extraction and Classification processes. Texture features were attained from the image in feature extraction phase with the use of AGLOH method. Moreover, in the classification phase, a new classification model named LCDRC approach was adopted for predicting the existence of mass in abdomen from US image. The proposed LCDRC approach was compared to other classifiers like NN, SVM, K-NN, and NB. From the results, it was found that the proposed model has attained better accuracy rate over other methods, which is 62.50%, 8.33%, 62.50%, and 29.99% better from the methods like NN, SVM, K-NN, and NB, respectively. The sensitivity of proposed model is high than the other approaches, and the proposed model is 39.99%, 85.71%, and 42.85% better from NN, K-NN and NB, respectively. Hence, the overall performance have shown that the developed LCDRC has attained better performance over the existing classifiers in terms of better diagnosing of mass.

REFRENCES

[1] J.C.M.van Zelst, T.Tan, B.Platel, M.de Jong,A.Steenbakkers, M.Mourits, A.Grivegnee, C.Borelli, N.Karssemeijer and R.M.Mann, " Improved cancer detection in automated breast ultrasound by radiologists using Computer Aided Detection", European Journal of Radiology, vol. 89, pp. 54-59, 2017.

[2] JinzuJi, Kin-FaiTongbAllannAl-Armaghany and Terence S.Leung, " A feasibility study of elastography based confocal microwave imaging technique for breast cancer detection", Optik - International Journal for Light and Electron Optics, vol. 144, pp. 108-114, 2017.

[3] S. Beura, B. Majhi, R. Dash and S. Roy, "Classification of mammogram using two-dimensional discrete orthonormal S-transform for breast cancer detection," in Healthcare Technology Letters, vol. 2, no. 2, pp. 46-51, 4 2015.

[4] M. D. Hossain and A. S. Mohan, "Cancer Detection in Highly Dense Breasts Using Coherently Focused Time-Reversal Microwave Imaging," in IEEE Transactions on Computational Imaging, vol. 3, no. 4, pp. 928-939, Dec. 2017.

[5] A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio and F. Caselli, "Mammographic Images Enhancement and Denoising for Breast Cancer Detection Using Dyadic Wavelet Processing," in IEEE Transactions on Instrumentation and Measurement, vol. 57, no. 7, pp. 1422-1430, July 2008.

[6] E.A.Zanaty, " Support Vector Machines (SVMs) versus Multilayer Perception (MLP) in data classification", Egyptian Informatics Journal, vol. 13, no. 3, pp. 177-183, 2012.

[7] Y. Mohan, S. S. Chee, D. K. P. Xin and L. P. Foong, "Artificial neural network for classification of depressive and normal in EEG," 2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, 2016, pp. 286-290.

[8] ZhenyunDeng, XiaoshuZhu, DeboCheng, MingZong and ShichaoZhang, " Efficient kNN classification algorithm for big data", Neurocomputing, vol. 195, pp. 143-148, 2016.

[9] Ronald R.Yager, " An extension of the naive Bayesian classifier", Information Sciences, vol. 176, no. 5, pp. 577-588, 2006.

[10] Yi Le, XianzeXu, Li Zha, Wencheng Zhao and Yanyan Zhu, " Tumor boundary detection in ultrasound imagery using multi-scale generalized gradient vector flow", Journal of Medical Ultrasonics", vol. 42, no. 1, pp. 25–38, 2015.

[11] Curiel L, Souchon R, Rouvière O, Gelet A and Chapelon JY, " Elastography for the follow-up of high-intensity focused ultrasound prostate cancer treatment: initial comparison with MRI", Ultrasound Med Biol, vol. 31, no. 11, pp. 1461-8, 2005.

Principal
Jawahar Arts, Science & Commerce College,
Andur Tal. Tuljapur Dist, Osmanabad

(01)

[12] Y. Liang, L. Liu, Y. Xu, Y. Xiang and B. Zou, "Multi-task GLOH feature selection for human age estimation," 2011 18th IEEE International Conference on Image Processing, Brussels, pp. 565-568, 2011.

[13] Punal M. Arabi, Gayatri Joshi, N. Vamsha Deepa, " Performance evaluation of GLCM and pixel intensity matrix for skin texture analysis",Perspectives in Science, vol. 8, pp. 203-206, September 2016.

[14] David Meyer, FriedrichLeisch, Kurt Hornik, "The support vector machine under test",Neurocomputing, vol. 55, pp.169 – 186, 2003.

[15] Francisco Guzmán, Shafiq Joty, Lluís Màrquez, Preslav Nakov, "Machine translation evaluation with neural networks", Computer Speech & Language, vol. 45, pp. 180-200, September 2017.