Grenze International Journal of Engineering and Technology, Jul 2019

An Efficient Approach for Detection and Classification of Breast Masses in Digital Mammograms

Aboli Umesh Vyawahare[1] and Dr.Vijaya Thool[2]
[1-2] Department of Instrumentation engineering, SGGSIE&T, Nanded, Maharshtra,India
Email: vyawahareaboli@gmail.com, vrthool@sggs.ac.in

***Abstract*—Breast Cancer is a fatal and common disease that affects thousands of women every year Worldwide. Therefore the early detection of the breast cancer plays crucial role in effective treatment and lowers the mortality rate. The breast examination is done with the help of mammograms with correct screening save lives of millions of peoples in the World. The Lobules (that produces milk) and glands that carries milk are the main two areas where breast cancer originates. The detection and classification of the masses in mammograms are still a difficult task and play a vital role to assist radiologists for accurate diagnosis. In this paper, the computer-aided approach that is based on deep learning technique is proposed. Even though most previous studies only deal with either detection or classification of masses, the proposed CAD system can handle detection and classification simultaneously in Single framework.**

***Index Terms*— Mammogram Image, Breast cancer, Cancer detection, Cancer Classification.**

I. INTRODUCTION

Breast cancer is the most common cancer among women in the world. It is the second leading cause of death after lung cancer. In 2016, 246,660 women were diagnosed with breast cancer and it is considered as the highest level of 29% among other kinds of cancers. Breast cancer alone contributes 14% among all other types of cancer [1]. Early detection is the only key to improve breast cancer prognosis and increase the survival rate. The breast cancer examination is done with Mammography that is the best diagnostic technique for screening. But, the interpretation of mammograms is not that easy because of minor differences in densities of different tissues within the image. This is especially true in case of dense breasts. The detection and diagnosis of breast cancer (i.e. Benign or malignant), the radiologist face so many challenges as they have a large number of breast images to examine daily. The difficulty of reading mammograms increases as the number of features of the tumor increases. (i.e., detecting the breast masses and correctly diagnosing them). Thus, CAD (computer-aided detection and diagnosis) is essential so that a second opinion can be provided to physicians to aid and support their decisions. The aim of our work is tumor identification in mammogram images and classification of the detected tumor. Most of the previous studies have been developed to build CAD systems utilizing conventional recognizers that are attempted to differentiate the breast lesions.

Almost all CAD systems developed before need manual detection of the masses before going to feature extraction where proper features are needed to be identified by a specialist. After this CAD system becomes manual or semi-automatic and could not support detection and classification issues in single framework. As a substitute to conventional classifiers that uses hand-crafted features, deep learning techniques can learn main

Grenze ID: 01.GIJET.5.2.501_1

features from the entire data [6,7] . For that purpose, now a days, deep learning is gaining a lot of attention in the field of machine learning. It has been also used in CAD for breast cancer to deal with some of the limitations of the conventional CAD systems mentioned above. The deep learning methods can learn a set of high-level characteristics and provide high recognition accuracy instead of handcrafted features.

In this research, a modern CAD system is developed for breast masses detection and classification by employing a new regional convolutional neural network using deep learning technique. In this model, DDSM database originally consisting of 600 mammographic images is used. These images were augmented by rotating the original mammograms in three different angles for training and testing purpose. The proposed method provides a powerful functionality so as the network can learn ROIs and their background at the same time. Thus, CAD system can carry out both detection and classification of breast lesions in only one framework. The proposed system exhibits an overall accuracy of detection and classification of 99.7 % and 97 %, respectively. This paper contains following main stages. Initially, the introduction of the overall system with the original and augmented databases for training and testing purpose is given. Next, the details of CAD system for detection and diagnosis of breast cancer masses is clarified. Then, we calculate the performance of our proposed CAD system throughout and get the five-fold cross validation result. At the end, discussion about the results comparing other CAD systems using classifiers such as DBN and CNN is done. Finally, the conclusion.

II. Materials and Methods

A. The proposed CAD system

Schematic diagram of the proposed CAD system is shown in Fig. 1. The model is an automated CAD system for simultaneous detection and classification of breast masses comprised of four main stages: Initially mammogram pre-processing,, feature extraction employing multilayer convolutional deep layers, mass detection and its confidence score and finally fully connected neural network (FC-NN) for classification of breast masses.

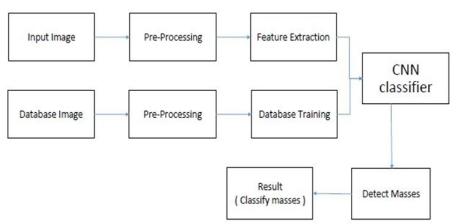


Fig1:Block diagram of the proposed CAD based system

B. Original database

In this study, a database from Digital Database for Screening Mammography (DDSM) to train and test our proposed CAD system is used. The DDSM database is developed by the University of South Florida and it has been used in breast research purposes worldwide. [6,11,17] . It consists of 2620 cases arranged in 43 volumes. Four mammograms are assembled for each case with two different views: Craniocaudal (CC) and Mediolateral oblique (MLO). Each mammogram is comprised of suspicious lesions similar to the information of the ground truth. This research consists of randomly picked out a set of 600 mammograms from DDSM database that are equally classified as benign and malignant cases.

C. Augmented database

Actually, deep learning technique needs a large amount of data for proper training of network. Owing to this, we have employed an augmentation technique so as the training data is increased. The generation of new instances from the original database utilizing different transformation methods such as rotation, translation, and scale is known as augmentation [6]. In order to minimize overfitting, we have augmented our database three times by rotating the mammographic images with angles of 90, 180, and 270 successfully [6,14]. Therefore, a total of 2400mammograms (i.e., the original Mammograms of 600 images along with their augmented database consisting of 1800 images) are used to train and test the proposed CAD system. Out of the total number of images, the half of mammograms represents the benign and the other half for the malignant case. All mammograms (original and augmented) are randomly mixed together in order to avoid any classification bias of CAD system.

D. Data pre-processing

In data pre-processing, the system must learn mammograms and their ROIs (i.e., masses). In the beginning, the mask comprised of breast lesions is obtained using the Otsu thresholding technique. Then, the mask image and the blurred image are multiplied which is generated by applying 2D Gaussian low pass filter to original mammogram. Then, the different threshold values are used to derive normalized thickness profile (NTP). These threshold values are derived with reference to the average of blurred image. Finally, the peripheral density correction image is carried out by dividing the original mammogram over the NTP image. This technique enhances the features of the mammograms by eliminating the background and irrelevant data. To obtain a high performance of CAD system, training and testing datasets are set in the range of [0, 1] as in [6] . In DDSM database, the mammograms found to be with different image sizes , hence training and testing datasets must be resized to a size of 448_448 as in to move further[11] .

E. Actual methodology of designing single framework

The proposed CAD system is one of the advanced deep learning techniques [11]. It can simultaneously detect and classify objects in the entire images. Unlike previous detection Techniques that applied the classifier to multiple regions of the image [8], this system employs a single convolutional neural network to the whole image. This proceeds towards dividing the input image into sub-regions and predicts multiple bounding boxes and their class probabilities for each region. Unlike traditional R-CNN that needs many networks for all the extracted regions, this system uses the entire mammograms so that the subjective information of the predictors with their Aspect are completely encoded with only one network in both training and testing time [11]. To encode the subjective information of predictors with their aspects in single framework, system uses the entire mammograms. Proposed CAD system has several benefits over other detection systems. This is because the system examines the image once and does not require a composite pipeline, it is extraordinarily fast and global context informs its prediction.

F. Architecture

The proposed CAD system is one of the advanced deep learning techniques [11]. It can simultaneously detect and classify objects in the entire images. Unlike previous detection techniques that applied the classifier to multiple regions of the image [8], this system employs a single convolutional neural network to the whole image. This proceeds towards dividing the input image into sub-regions and predicts multiple bounding boxes and their class probabilities for each region. Unlike traditional R-CNN that needs many networks for all the extracted regions, this system uses the entire mammograms so that the subjective information of the predictors with Their Aspect are completely encoded with only one network in both training and testing time [11]. To encode the subjective information of predictors with their aspects in single framework, system uses the entire mammograms. Proposed CAD system has several benefits over other detection systems. This is because the system examines the image once and does not require a composite pipeline, it is extraordinarily fast and global context informs its prediction Moreover, depending on the conditional class probability, the detected mass is identified as benign or malignant for the corresponding cell [11]. Then, the confidence score for each specific class is calculated.

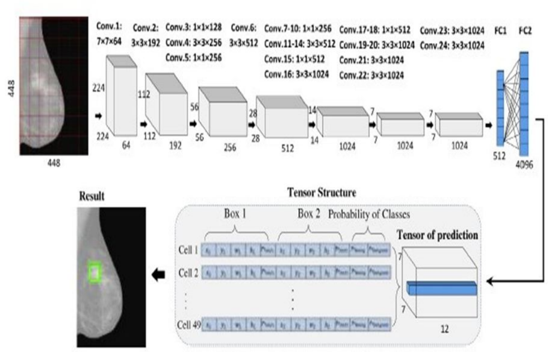


Fig2: Structure of proposed CAD based system

Confidence score = Probability Confidence = Probability IOU (ground truth predicted) (2)

For the representation of the mass that deals with the predicted box and to identify how accurate the mass is, we require to estimate the confidence score. If the grid cell does not consist of any lesion, then automatically the confidence score decreases to zero The system is trained to employ the entire breast image and its ROIs information as well. The training data with the ROI position and size information is prepared for training of data: the information of training data consists of the centre position (x, y), width (w), height (h), and class label of the masses. In this research, we are going to use 24 convolutional layers with different kernel sizes, size of 2 2 max-pooling layers, Activation functions and two fully connected Layers [11]. Deep-based feature maps are extracted from each convolutional layer by applying different filter types (i.e., different kernels W). Feature maps are generated when the convolutional filters extracts different features from the whole mammogram. Convolution operation of Y k L that shows the k th feature map of layer L is
Evaluated:

Y kL = wkLY kL1 + bkL (3)

Where
ϕ(_)=the activation function
?=convolution operator
b k L=bias
To minimize the dimensionality of the features and proper features selection at each layer of the network, down sampling by max-pooling method is applied. Therefore, the input for the next layer is only a maximum value from a 22 window. So, to make sure that the features information is not lost, max-pooling filter should not be large in size [6,15,27,28] . A pixel stride of two is used with all convolutional and pooling layers. The final aggregated deep features, which are generated by the convolutional and pooling layers, are passed to the fully-connected network (FC- NN). Linear leaky rectified activation function is used for all layers [11] and for final layer, the Rectified Linear Unit (ReLU) is used, (z) = max(0, z), [6]. The leaky rectified activation function represents the linear transformation model of an input z as shown below,
ϕ(z) = z, if z ¿ 0
0 . 1z , otherwise .
The fully connected neural network layers utilize the training set images so that the training weights are updated during the training phase. In this study, a learning rate of 0.001 and a bath size of 64 is used so that it can successfully be applied to the proposed CAD system [11]. Finally, tensor of prediction (ToP) with size of N N (5 M + C) is generated.
Where,
N N=number of grid cells
M=Bounding boxes
C= classes (i.e., benign and malignant)
C = 2 as there are two classes (i.e., benign and malignant) While, as each grid cell is responsible for detection and classification, it acts as an unit [11] . One is free to choose any size but the best performance is given by 7 _ 7 (i.e., N = 7) so we have chosen it.[15] . M = 2 to get the best predicted box of the object out of mammographic inner and outer boundary as well. Among the total bounding boxes, the one with the highest confidence score get selected, known aspredicted box. Thus, the final output of the network represents a 3D matrix of Tensor of Prediction with size of 7_7_12 (Fig2). The tensors are formed utilizing the grid information. These grids thus formed are expressed with 12 elements in the tensor. A set of elements are related to the particular bounding box prediction. Similarly, the first five elements are correlated with the predictions of the first bounding box, while the second five elements are with the second bounding box. In case of each,
these elements illustrate the prediction information of the mass locations which are x, y, w, h , with confidence probability. The last two elements (i.e., Pr Benign and Pr Malignant) in the Tensor of Prediction correspond to the confidence scores of the class probabilities for both benign and malignant cases, respectively. These class probabilities are observed to be the highest confidence probability among the bounding boxes. Therefore, only one bounding box is determined by the system for each grid cell which is can detect the mass location and assign its relevant class. Ultimately, the system selects those boxes with confidence score greater than a particular threshold among all of the potential predicted masses in each mammogram. The Darknet framework is used for all training and testing processes.

G. Training and testing

The data is already trained and tested on the proposed system using original and augmented datasets. After training, the system with augmented dataset got the results of both detection and classification of the breast abnormalities except one case of comparing the effect of data augmentation against the case of using the original dataset explained later in Section 3.4. First step is to standardize the parameters of the proposed CAD system using only the training (i.e., 80 % of the data) dataset to avoid any bias in training and testing Then, the final system performance was analysed using only the testing dataset (i.e., 20% of the data). The transfer-learning concept is effective in training a deep net [11,14]. As this transfer learning was applied to DDSM in [6,14] , the CAD system with the pre-trained weights with a large computer vision ImageNet dataset is trained. Subsequently, it was finely tuned with the training augmented mammograms. The main requirement is to ensure that every image in the dataset appears to be in test exactly one time and to reduce the bias error that may occur during the classification process. Thus, a k fold cross validation is performed. (In our case, k=5). So, the dataset is randomly divided into five subsets where each subset is made by 10 % benign and 10 % malignant cases. One of the subsets (i.e., 80 % of dataset) is utilized as a training set and the other subsets (i.e., 20 % of dataset) are considered together as a testing set. This means that the CAD system is trained five times (No of folds, k=5) to get the performance of the CAD system. For each k - fold, the computation time took almost eight days to perform the training stage. However, the decoding (i.e., detection and classification) for each mammogram takes only less than ten seconds. Thus, the proposed CAD system seems to be feasible and reliable to apply in the future for clinical applications. This wok was conducted on a PC Intel Core(TM) i3□6006U CPU @ 2.00GHz with 4 GB RAM. We utilized Matlab 2018b as programming language on operating system of windows10. The detection and classification results are evaluated as an average of the 5-fold cross validation results.

H. Performance evaluation measures

In this research, to evaluate the performance of our CAD, some objective measures have been utilized Fig. 3 shows the evaluation process during the testing phase to detect the mass location in the mammogram and categorize into benign or malignant. The confidence probability scores may vary, and if the confidence scores of the detected boxes are less than a particular threshold, the corresponding predicted ROIs or masses are concluded as undetectable .The predicted masses which have confidence probability scores equal or greater than this threshold are considered for the next classification and detection assessment processes excepting those undetectable cases as they have been eliminated before further process. For detection assessment, the mass location is properly detected if and only if IOU is equal to or greater than 50 % comparing with its ground truth. In the case of false mass detection, as long as the condition of threshold probability is satisfied, the final decision to distinguish the type of these masses is achieved utilizing the capability of the classifier. To quantitatively represent the ability of the proposed technique, confusion matrix is employed to indicate how the proposed CAD can classify between benign and malignant classes [17]. While, to classify and evaluate the overall result, the curve of receiver operator characteristic (ROC) with its area under curve (AUC) are used. AUC of ROC exhibits the functioning and performance of the classifiers. In this case, if the AUC value is close to 1.0 means the highly accurate diagnostic rate whereas AUC value close to 0.5 represents the unreliable Performance.

Sensitivity = TP=(TP + FN) (5)

Specificity = TN=(TN + FP) (6)

Where,

TP= True positive

FN= false negative

TN =True negative

FP= false positive.

The overall classification accuracy of the system is

Overall classification accuracy = TP +TN=(TP +FN +TN + FP) (7)

Besides this, free response operating characteristic (FROC) curve is used to measure the accuracy of the detected masses in mammograms compared to the ground truth is quantitatively evaluated using.

III. RESULTS

A. Class probability threshold

As mentioned earlier, the confidence score for each grid cell corresponds to the class probability of the highest confidence between the two bounding boxes. Thus, the proposed CAD system creates many potential ROIs for each testing mammogram. The attempt was to find the correct threshold that avoids the unnecessary ROIs with very small class probabilities. All potential ROIs are shown on the mammogram in the case of zero thresholds as shown in Fig. 4 (a). While, to maintain the testing data, threshold should not be larger.

P(Condition Positive	20
N(Condition negative	20
TP	27
TN	30
FP	3
FN	0
TPR(SENSITIVITY	1
TNR(SPECIFICITY)	1
TP+FP	27
Precision	1
TP+TN	57
P+N	57
ACCURACY	1
2TP	54
2TP+FP+FN	54
F1-SCORE	1

Data 1: Sensitivity and specificity of Normal Vs benign

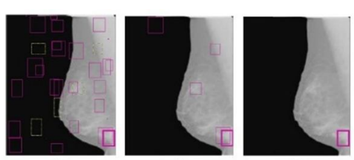
Thus, the right threshold that accomplishes the minimum number of undetectable data was analysed. Fig. 4 (b) and (c) show the potential ROIs and class probability threshold is 0.01 and 0.2, respectively. Boxes having higher confidence are denoted using thicker border. In this study, a probability threshold to be 0.2 was evaluated experimentally. As explained before, the numbers of potential ROIs are controlled by this threshold through each testing mammogram. With the probability threshold of 0.2, at least one potential ROI was provided while ignoring all undesirable ROIs with too small class probability. A proper threshold must provide enough detected ROIs for further classification. Fig A shows the GUI for breast mass detection and classification that will be displayed on MATLAB window after successful complication of earlier stages.

P(Condition Positive	20
N(Condition negative	20
TP	27
TN	30
FP	3
FN	1
TPR(SENSITIVITY	1
TNR(SPECIFICITY)	1
TP+FP	27
Precision	1
TP+TN	57
P+N	57
ACCURACY	1
2TP	54
2TP+FP+FN	54
F1-SCORE	1

Data 2: Sensitivity and specificity of Normal Vs malignant

B. Mass detection using CAD

The results of the proposed CAD system on the ability of detecting the location of the benign and malignant masses are shown in Fig. 5; Fig. 5 (a) and (c) show the ground truth and Fig. 5 (b) and (d) show the masses detected by our CAD. The abnormalities (i.e., masses) detection performance throughout the 5-fold test of benign and malignant is reported in Table I.



(a) (b) (c)

Fig 4: Potential ROIs with the threshold probability of (a) zero (b) 0.01, (c) 0.2 The ROI with the highest confidence has the thickest border

Dataset consists of 480 mammograms. At every fold test that are equally divided into two cases (benign and malignant cases). The results show the robustness of our CAD when detected the mass position in mammograms with an overall accuracy of 99.7 %. As mentioned previously, this system generates the confidence probability for every potential ROI that shows the mass position. In Table 1, false detection cases represent those have the detected boxes with confidence probability less than 0.2 or IOU less than 0.5. These detection results are achieved by comparing the IOU of each detected boxes With the ground truth. Fig. 7 (a) shows the detection performance of the proposed CAD system through 2nd fold test in term of FROC curve. The potential detected boxes are considered as successfully detected if the overlap equals or exceeds 50 % comparing with their corresponding ground truth masses. The proposed method in Fig. 7 (a) produces a true positive rate of 99.17 % at false positive per image equals to 0.22.

C. Mass classification using CAD

The classification is done considering an overall accuracy and is evaluated using all test data. Except those undetectable mammograms in detection stage throughout each k fold subset all other images are considered at this stage. Some representative solutions of the CAD system in terms of detection and classification for benign and malignant cases are shown in fig 5.

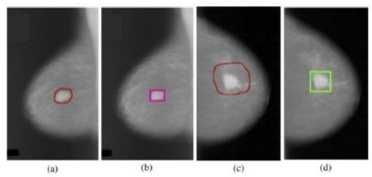


Fig 5: Mass detection (a) and (b) show the ground truth mass and detected from new proposed method for benign case, while(c) and (d) for a malignant case

The Undetectable breast tissues are Labelled as normal tissue.

TABLE I: 5 FOLDS CROSS VALIDATION PERFORMANCE OF THE MASS DETECTION VIA THE PROPOSED CAD SYSTEM

Fold test	Benign		Malignant		Total	
	True	False	True	False	True	False
1st fold	**240**	**0**	**240**	**0**	**480**	**0**
	100%	0.0%	100%	0.0%	100%	0.0%
2nd fold	**237**	**3**	**239**	**1**	**476**	**4**
	98.75%	1.25%	99.58%	0.42%	99.17%	0.83%
3rd fold	**238**	**2**	**240**	**0**	**478**	**2**
	99.17%	0.83%	100%	0.0%	99.58%	0.42%
4th fold	**239**	**1**	**240**	**0**	**479**	**1**
	99.58%	0.42%	100%	0.0%	99.79%	0.21%
5th fold	**240**	**0**	**240**	**0**	**480**	**0**
	100%	0.0%	100%	0.0%	100%	0.0%
Average (%)	**99.50**	**0.50**	**99.92**	**0.08**	**99.71**	**0.29**

Confusion matrices of the CAD for each k -fold subset system are presented in Table 2. It is clearly shown that the benign cases are correctly classified with 100 % in all k-fold subsets, while the malignant cases are in between 92.5 % and 95.8 %.The false Positive in the 1st fold subset represents the 18 malignant cases that are in-correctly classified and negatively affect the specificity. It is clear from Table 2 that the results of AUCs and accuracies of all k -fold are similar to each other. This shows the efficiency and feasibility of the proposed CAD system in diagnosis of the breast masses. The performance of the proposed CAD system as an average of the 5-fold cross validation results in terms of sensitivity, specificity, AUC, and overall detection and classification accuracies is evaluated. The CAD system achieved an overall detection and classification accuracies of 99.70 % and 97.00 %, respectively. The statistical measures of true positive and true negative rates are reflected by the sensitivity and specificity values, respectively. The CAD system concluded 100 % sensitivity for benign cases and 94 %. Specificity for malignant cases.

The challenging task for mass detection and classification is in case of the masses that exist over the pectoral muscle or surrounded by the dense region due to variation in shape, texture and size of different masses [4]. This CAD system is capable of detecting and predicting the mass in these two challenging cases as shown in Fig. 6 (a) and (c).The proposed CAD system has overcome these challenges and results are shown in Fig. 6 (b) and (d). Fig. 6 (b) and (d) illustrate the capability of this CAD.

D. The effect of the size of data sets

This section shows the effect of different training dataset sizes (i.e., original vs. augmented datasets) that are used in training of CAD system. Though the original dataset comprises of only 600 mammograms, the augmented dataset that adds all the original and augmented mammograms, total 2400 cases is used. Usually, the performance of deep learning techniques improves when the size of the training dataset is increased. Fig. 7 (b) illustrates the effect of enlarged dataset on the proposed CAD system. It indicates the improvement in respect of ROC curves in the case of the augmented dataset against the original ones with AUCs of 96.45 % and 87.74%, respectively. Importantly, the augmented data also improves the specificity with 16% as shown in Table 3 which apparently increases the overall classification accuracy from 85.52% to 97%. Note: Results in all sections are based on the training of the augmented dataset.

IV. DISCUSSION

In this research, the detection and classification of the mammograms via deep learning technique is done. According to Survey, the recent deep learning CAD systems only could complete the diagnosis task of the extracted patches from mammograms [6, 7, 10].

TABLE II: CONFUSION MATRICES AND PERFORMANCE OF CAD SYSTEM THROUGH 5 FOLD CROSS VALIDATION

Fold test	Actual classes	Predicted classes		Sensitivity (%)	Specificity (%)	AUC (%)	Accuracy (%)
		Benign	Malignant				
1st fold	**Benign**	**240** 100%	**0** 0.0%	100	92.5	95.73	96.25
	Malignant	**18** 7.5%	**222** 92.5%				
2nd fold	**Benign**	**239** 100%	**0** 0.0%	100	95.82	97.16	97.91
	Malignant	**10** 4.2%	**229** 95.8%				
3rd fold	**Benign**	**240** 100%	**0** 0.0%	100	94.58	96.83	97.29
	Malignant	**13** 5.4%	**227** 94.6%				
4th fold	**Benign**	**239** 100%	**0** 0.0%	100	94.17	95.85	97.08
	Malignant	**14** 5.8%	**226** 94.2%				
5th fold	**Benign**	**240** 100%	**0** 0.0%	100	92.92	96.66	96.46
	Malignant	**17** 7.1%	**223** 92.9%				
Average	**Benign**	**100%**	**0.0%**	**100**	**94.00**	**96.45**	**97.00**
	Malignant	**6.00%**	**94.00%**				

A CAD system based on the deep learning technique that detects the locations of potential masses on mammograms and classifies them into benign or malignant simultaneously in single framework is developed. Figure 5 and 6 indicates the ability of the proposed CAD system in detecting the potential breast lesions and Providing the proper diagnosis for each mammogram (i.e., benign or malignant). Moreover, this CAD system can overcome two important challenges faced in the clinical mammogram. First, it could detect the breast masses, existed over the pectoral muscle as shown in Fig. 6 (b) and the proposed methodology successfully identified breast masses in the dense tissues as shown in Fig. 6 (d). When these two regions are examined and compared with the normal breast tissues some high intensity (more bright) issue occurs and due to that radiologist faces challenges in detecting the existing mass in those regions.

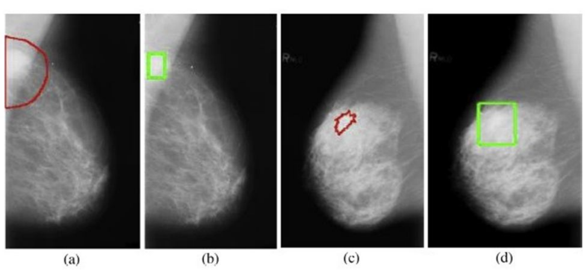


Fig 6 detection and classification (a) and(c) show the ground truth mass over the pectoral Muscle and the detected by new proposed method. (b) and (d) present the ground truth mass Surrounding by dense tissue and the detected by new proposed CAD

The results of Tables 1 and 2 show the analysis of both detection and classification of the breast masses throughout 5-fold cross validation. As discussed earlier the training using augmented data improves the performance of breast masses detection and classification [6,11].
The overall accuracy performance increased from 85.5% to 97% with the 600 original mammograms and the 2400 augmented mammograms respectively as shown in Fig. 7 (b) and Table 3.
These results demonstrate that the CAD system is efficient approach to achieve high accuracy in simultaneous detection and classification of the abnormalities in single framework. In order to present how powerful the proposed CAD system is, the results were compared with the latest studies employing DBN and CNN. Comparison with the conventional classifiers that do not utilize deep learning is also provided to present the efficiency of the deep learning algorithms.

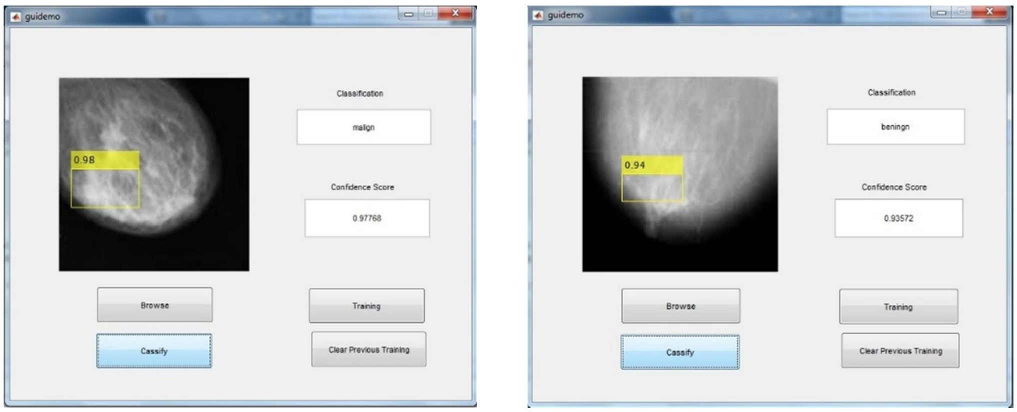


Fig 7: Results of mass detection and classification with confidence score

In the previous work that used same kind of DDSM mammograms, a DBN-based CAD system was applied and compared its outcomes against the conventional linear discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and Neural Network (NN) classifiers. Statistical handcrafted features are excerpted from the extracted masses. That study observed the effect of features dimensionality reduction through different feature selection methods. The overall classification accuracies are shown by employing the Sequential Floating Forward (SFFS) algorithm as shown in Table 4. Conversely, For DBN technique, all the extracted features are used, without features reduction, to train and test the CAD system. It clearly shows that how the DBN can overcome the conventional classifiers with overall accuracy of 90.48%. Random forest (RF) classifier against CNN was compared. Different types of features set are manually extracted from the mass patches to train RF classifier. Each feature set are trained individually and then they are applied separately to the test set. Their results concluded that the features group of candidate detector, contrast, texture, geometry, location, context, and patient information got AUCs of 85.8%, 78.7%, 71.8%, 75.3%, 68.6%, 81.6%, and 65.1%, respectively. While the entire feature sets together obtained AUC of 90.6% against 92.0% in the case of CNN and a framework for CAD system utilizing CNN technique was presented in [6]. Our proposed CAD system that used the whole mammograms for the convolution layers, but they have used only ROIs of the cropped masses. The middle level and high-level feature combination is used to train and test these CAD system based CNNs. The performance ability of CNN to classify the masses into benign or malignant with overall accuracy is 96.7% this system can be highly compared with these results.
Additionally, this CAD system is the only system that can detect the masses in mammograms apart from predicting their types compared with the conventional CNN. Finally, we show a comparison of the effect of utilizing augmented data instead of original ones. Rotation, Translation and scaling
are the three transformations applied on original dataset to get the augmented dataset. We apply CNN to the mass patches with a size of 250 250 to classify it in normal and malignant cases. Their AUC results of the CNN on original dataset reached 87.5% whereas with augmented dataset it achieved the accuracy of 92.9%. This improvement rate is comparable with our AUC results from 87.74% to 96.45%.

V. CONCLUSION

In this paper, CAD system for detection and classification breast mass is presented. In the system, a ROI-based CNN approach included two layers starting with the convolutional layers followed by fully connected neural networks to detect the proper location of the mass and to distinguish the tumor types as either benign or malignant. The results show feasible and promising results in case of detecting the location of benign and malignant masses and recognizing their proper classes as well. Moreover, the CAD system detects most challenging cases of breast cancer, mass existing over the pectoral muscle or surrounding by the dense tissue in the mammograms further step of the presented CAD system is to be tested in practice for its real validity.

REFERENCES

[1] R.L. Siegel , K.D. Miller , A. Jemal , Cancer statistics,2016, CA Cancer J. Clin. 66 (1) (2016) 730

[2] J. Virmani , N. Dey , V. Kumar , in: PCA-PNN and PCASVMBased CAD Systems For Breast Density Classification, Springer International Publishing, Warsaw, Poland, 2016, pp. 159180 .

[3] C. Muramatsu , T. Hara , T. Endo , H. Fujita , Breast mass classification on mammograms using radial local ternary patterns, Comput. Biol. Med. 72 (1) (2016) 4353 .

[4] H. Li , X. Meng , T. Wang , Y. Tang , Y. Yin , Breast masses in mammography classification with local contour features, BioMed. Eng. OnLine 16 (1) (2017) 4454 .

[5] S.A. Taghanaki , J. Kawahara , B. Miles , G. Hamarneh, Pareto-optimal multi-ob- jective dimensionality reduction deep auto-encoder for mammography classification, Comput. Methods Progr. Biomed. 145 (2017)

[6] Z. Jiao , X. Gao , Y. Wang , J. Li , A deep feature based framework for breast masses classification, Neurocomputing 197 (C) (2016) 221231 .

[7] J. Arevalo , F.A. Gonzlez , R. Ramos-Polln , J.L. Oliveira , M.A.G. Lopezd , Repre- sentation learning for mammography mass lesion classification with convolutional neural networks, Comput. Methods Progr. Biomed. 127 (2016) 248257 .

[8] N. Dhungel , G. Carneiro , A.P. Bradley , Automated mass detection in mammo- grams using cascaded deep learning and random forests, 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA), 2015 .

[9] N. Dhungel , G. Carneiro , A.P. Bradley , A deep learning approach for the analysis of masses in mammograms with minimal user intervention, Med. Image Anal. 37 (2017) 114128 .

[10] T. Kooi , A. Gubern-Merida , J.-J. Mordang , R. Mann , R. Pijnappel , K. Schuur , A. d. Heeten , N. Karssemeije , A comparison between a deep convolutional neural network and radiologists for classifying regions of interest in mammography, International Workshop on Digital Mammography, 2016 .

[11] J. Redmon , S. Divvala , R. Girshick , A. Farhadi , You only look once: unified, realtime object detection, IEEE Conference on Computer Vision and Pattern Recog- nition,2016 .

[12] M. Dong , X. Lu , Y. Ma , Y. Guo , Y. Ma , K. Wang , An efficient approach for automated mass segmentation and classification in mammograms, J. Digit. Imag. 28 (5) (2015) 613625..

[13] Y. Qiu , S. Yan , R.R. Gundreddy , Y. Wang , S. Cheng , H. Liu , B. Zheng , A new approach to develop computeraided diagnosis scheme of breast mass classi- fication using deep learning technology, J. X-Ray Sci. Technol. 25 (5) (2017) 751763.

[14] G. Carneiro , J. Nascimento , A.P. Bradley , Automated analysis of unregistered multi-view mammograms with deep learning, IEEE Trans. Med. Imag. 36 (11) (2017) 23552365.