DOI Number: 10.37506/v20/i1/2020/mlu/194726

An Accurate Breast Cancer Detection System Based on Deep Learning CNN

Kian Raheem Qasim[1], Alia Jumaa Ouda[1]

[1]*University of Information Technology and Communications, Baghdad, Iraq*

Abstract

Deep learning of multilayered computational models allowed processing to recognize the representation of data at multiple levels of abstraction. These technologies have significantly improved to identify breast cancer. Convolutional Neural Network (CNN) is a type of radiological basis for classification problems and is based on the Bayes decision-making base, which reduces the expected error in classification. In this paper,it is proposed to detect breast cancer using CNN of the mammography system to classify the mammogram to noncancerous abnormality and cancerous abnormality. The goal of Breast Cancer Detection based on CNN for speeding up diagnosis by helping a particular diagnose and classify breast cancer.A set of images of a mammogram is used to perform the pre-processing of the histogram equalization and adjust the appropriate parameters for the CNN work. Then, the whole changed images are set in CNN to classify as a training source. A CNN classifies can be produced as a model foridentifying the mammogram. the BCDCNN method with the mammography classification utilizing MCCANN comparison with BCDCNN improved the classificationaccuracy on mammographic. Therefore, the results showed that the proposed system has a higher resolution than the other recently existing systems, and the only mass containing all was maximized from 0.84 to 0.88 and 0.70 to 0.82 accuracy.

***Keywords**: Breast Cancer, Convolutional Neural Network, Artificial Neural Network, Mammography Classification, Histogram Equalization.*

Introduction

All over the world, the most common causes of many women deaths are breast cancer, the widespread type of cancer impacting women. In 2017, the estimates indicated that there were more than seventeen thousand five hundred women and one hundred and forty-four men with breast cancer. As an average, forty-eight of people with breast cancer were diagnosed daily. In Australia, the number of men and women affecting breast cancer has been increasing; as yet, the number of dead peoplebecause breast cancer is decreasing. The widespread type of cancer among women is breast cancer, which accounts for about 28.3% of all new cancers in women in 2017. The risk of this type of cancer at the age of 85 is 1 in 8 for women and 1 in 631 for men. If breast cancer is known to occur of younger women. In 2017, it was also estimated that 841 women aged between (twenty and thirty-nine) years were diagnosed with breast cancer. This represents 4.7% of all types of breast cancer diagnosed in Australia [1].

According to the Breast Cancer Study, we found that 50% of NHS say they do not have the particular staff to recruit individuals with a limited nurse specializing in breast cancer [2]. This is the most significant cause that may result a low-rate of survival of breast cancer worldwide. because of the lack of a nurse or breast cancer doctor, this will result in delayed diagnosis of breast cancer, lack of compliance with detection or optimal treatment, and unequal access to optimal treatment. Thus,detection of breast cancerhas been developed to perform effectively both an anomaly and breast screening classification. This is to help and diagnose breast cancer.New proposal for image recognition system withdetection of breast cancer. Image recognition utilized NN (Neural Network), a type of ANN that is designed and utilized successfully to determine visual images.This frameworkutilized a neural network, a kind of artificial neural network constructed and used effectively to identify visual images. This system is capable of classifying and detecting abnormalities in the image of mammograms. In general, the image of a mammogram may be categorized as normal or cancerous abnormality (malignant), non-cancerous abnormality (benign) [3] . In this paragraph, some previous researchers will be reviewed thathave used various types of the neural network to depict mammograms into breast cancer as shown below:

In [4] (2009),Bozeket al. presented an indicator of breast cancer, such as architectural distortion but these are less important. The mass may be each non-benign. The variation between malignant and benign tumors is that tumors of benign have an oval or round shapes, whereas malignant tumors have a partially round shape of irregular outlines. Therefore, the malignant mass can seem whiter than other surrounding tissue.

In [5],(2011),Sharkas, Al-Sharkawy and Ragab presented discrete wavelet transformations (DWT), conversion of contour, and principal component analysis methods (PCA) to extract the feature. The system was the capability of detection and classification of abnormal and normal tissues. In addition, MC tumors were classified as malignant and benign. The rate of classification was nearly 98%.

In [6] ,(2013),Jawad Nagi presented Detection Breast Cancer (DBC) in Mammogram Images, achieved in four stages. The first phase was preprocessing and improved mammogram using global morphological techniques to reduce the image of the marker. The second phase was the process of segmentation using two techniques. Seed and growing the region and multiple OTSU threshold techniques. The third phase was the extraction of texture-based features and extracted fluorescence features. The last phase is passed by training and testing classification, and for the process of the training, the artificial neural network (ANN) was proposed.

In [7] , 2015, W. Peng, R.V. et al. presented seeded region growing (SRG) and median filter using the 2D algorithm, improved image contrast, noise removal, deletion of radiology artifacts and elimination of chest muscle reduction from digital mammography, and ANN technology to classification the mammogram as follows: Normal Indicates a benign tumor or a malignancy. Here, 222 random images from the Breast Image Analysis Society database.

In [8] , (2016), Jain. A,LevyD.presented Convolutional neural networks will be utilized to classify breast-segmented mammalian directly into mammography as benign or malignant, utilizing a combination of transport learning, pre-processing and increased data to overcome limited training data. We deliver the state-of-the-art result on the DDSM data set, surpassing the performance of human, and demonstrate the viability of our model.

· In [9] ,(2018),SairaCharan et al. presentedneural networks to classifyabnormal and normal detection of breast cancer. Convolutional neural networks (CNN)can be utilized in this detection. The dataset of mammograms-miss is utilized for this purpose system, containing 322 mammogram images of about 189 natural images and 133 abnormal breast images. A promising trial has been obtained that demonstrates the effectiveness of deep learning to detect breast cancer in mammograms and encourages the using of new methods of extracting and classifying features based on deep learning in different applications of medical imaging, essentially in the detection of breast cancer.

Convolutional Neural Networks (CNN)

AConvolutional neural networks (CNN) is a special case of the neural network and a main tool of deep learning. CNN is perfect for pattern recognition for images. It includes more than convolutional layers, overwhelmingly in a subsampling layer, followed by more than connected layers just as in a Standard Neural Network(SNN) [10] . The designing of CNN is derived behind the discovery of an optical mechanism, the visual cortex, into the brain. The visual cortex contains many cells in charge oflightdetecting in the small, overlapping subfields of the visible field, called the receiving fields. These cells act as local filters on the input area, while more complex cells contain more receptive fields. The torsion layer in CNN performs the function performed through cells into the visual cortex [11] .

The architecture of CNN

The primary focus of CNN is on the input basis whichis comprised of images. This will give a focus on the architecture that will be set up in the best way that is required for handling data of specific type [12] . CNN consists of three types of classes. These are convolutional layers, pooling layers, and fully-connected layers. When these layers are stacked, CNN architecture has been formed.The simple architecture of CNN forclassification of Breast cancer is shown in Figure (1).

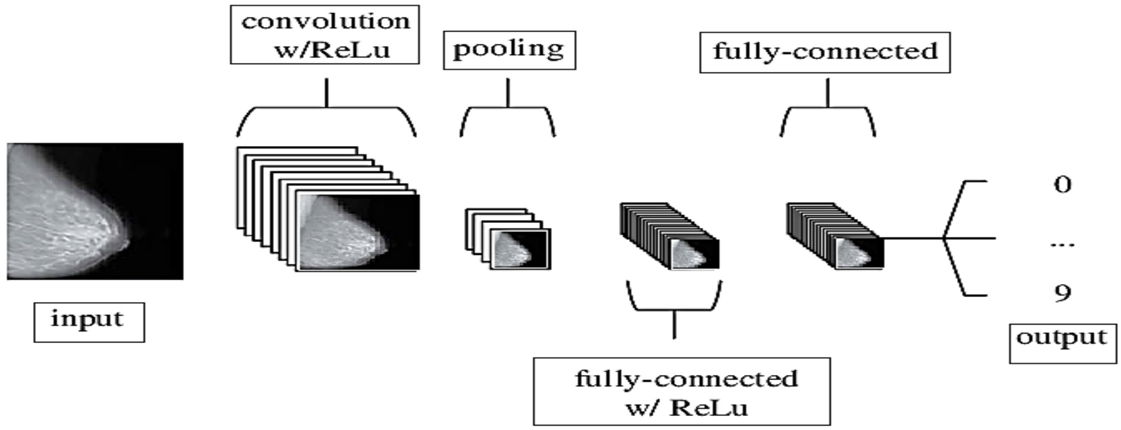


Figure (1): A simple architecture of CNN [12] .

The basic function of CNN above will be divided into four major regions.

1. As found in other forms of artificial neural network (ANNs), the input layer can contain pixel value into image.

2. ***Convolutional Layer*** will be determined the yield of the neurons that are connected to the local input areas by calculating theproduct of scalar between its weights and the region related to the input volume. The aims of the rectified linear unit (ReLu) apply the "initial" activation function, like the sine, to the activation outputs produced through the previous layer.

3. ***Pooling Layer*** simply reduces the samples with the spatial dimension of the selected inputs, minimizing the number of parameters in that activation.

4.***Fully Connected Layers*** can perform the same duties as the standard artificial neural network (ANNs) and try to produce the activation grades for classification. Also,it is suggested to use ReLu between these layers, for performance improving.

V. Evaluation performance of CNN

In any field of science, it is almost a necessary requirement that the performance of classifier must be evaluated to understand and measure the suitability of classifying a given problem. The classification of binary predicts all instances of data for the test data set either negative or positive.This classification or prediction results in four outcomes - true or false (positive and negative).

1. ***True positive (TP)***: Correct of Positive Prediction

2. ***False positive (FP):***Incorrect ofPositive Prediction

3. ***True negative (TN):***Correct of Negative Prediction

4. ***False negative (FN):***Incorrect ofNegative Prediction

One of the most common measures is the confusion matrix it will be explained in the next section.

I. Confusion Matrix

The Confusion matrix is consisting of a two by two (2 x2)Table that contains four outcomes produced by a classification, these are essential performance measures, like accuracy, specificity, and sensitivity which are derived from the confusion matrix. The confusion matrix is utilized to represent the test of the result prediction model. Each of rows stands for the predicted class which means the (Output Class), and the columns stand for the true class which means the (Target Class). In Table (1),The matrix of confusion is displayed, which is described as the different values and equations associated with them. Few of these equations are closely related to performance analysis.

Table (1): A Typical Confusion matrix [13] .

Confusion matrix		**Predicted**	
		Negative	**Positive**
Actual	Negative	TN	FP
	Positive	FN	TP

The inputs of the confusion matrix have meaning in the context of the problem of data mining:

1. TN is the number of the prediction of true that in the case is negative,

2. FN is the number of predictions of false that a case positive,

3. TP is the number of predictions of true in which a positive instance,

4. FP is the number of predictions of false of negative instance.

The following are theBasic measures derived from the confusion matrix:

1) Accuracy

Accuracy (ACC) is determined as a number for all predictions of correct (TP + TN) divided by the total number of data sets (P + N). The best of accuracy equal to 1.0, while the worst equal 0.0. It can likewise be determined by 1 - error (ERR) as shown in equation (1).

$$ACC = \frac{TP + TN}{TP + TN + FN + FP} = \frac{TP + TN}{P + N} \quad (1)$$

2) Sensitivity (Recall or True Positive Rate)

Determined the number of predictions of true positive (TP) divided by a total number of the positives (P) this method called Sensitivity (SN) or likewise Recall or the True Positive Rate (TPR) (REC). The sensitivity equal to 1.0 is best, whereas the worst equal 0.0 as shown in equation (2).

$$SN = \frac{TP}{TP + FN} = \frac{TP}{P} \quad (2)$$

3)*Specificity (True Negative Rate)*

Determined the number of predictions of True Negative (TN) divided by the total the number of negatives (N) this method called Specificity or True Negative Rate (TNR). The specificity equal to 1.0 is best, whereas the worst equal 0.0 as shown in equation (3).

$$SN = \frac{TN}{TN + FP} = \frac{TN}{N} \quad (3)$$

VI. The Proposed Method

The proposed method consists of three main stages, in the first stage is the acquisition of the image, second stage extraction features from the mammograms, selecting more optimal features, classifier to identify an appropriate classification of mammogram utilized CNN classification accuracy mainly depend on careful selection of features. In the finally step mammograms are classified, for this Convolutional neural network is used as a classifier to distinguish mammogram and classified it into malignant and normal classas shown in this Figure (2).

Figure (2). the main proposed system.

A. Dataset

The dataset used for this paper is a known Mammographic Database. Convolutional neural networks require a large amount of training data to realize height accuracy. Because of less accessibility of thedataset of enormous,testing and training were done from the most accessible dataset on the web. In this paper were utilized322 full images. The actual size of the images was 1024- 1024. It contained about 133 images that were abnormal and 189 of the regular category. The images of abnormal are included an inconsistency in which one breast increased density, 21 of which were included. The architectural deformity was the second types of abnormal, showing the arrangement of abnormal tissue on the breasts. This was included about 22 images.The small calcium deposits in breasts can be developed, where 24 of these images were found to contain calcification, which is irregularly shaped in the breast and can be malignant and some have not been ascertained if they are malignant. Data were randomly training by 70% and 30% were tested by CNN technology.

B.Equalize Histogram

In order to eliminate the effect of different illumination conditions, where images are taken in, the color histogram of the grayscale image is equalized, so that, the features are more distinguishable by the classifier, the overall effect of the lighting in the environment is removed.

Results

In this paragraph will illustrate the results obtained using the proposed method for a range of breast cancer images. The results of the proposed algorithm CNN are the best for detection of breast cancer.

Dataset:Mammographic Database is used in this proposed system.breast images Include both sides arranged in pairs, where each pair represents the right and left mammograms for 161 patients that mean 322 images as shown in this Figure (3).

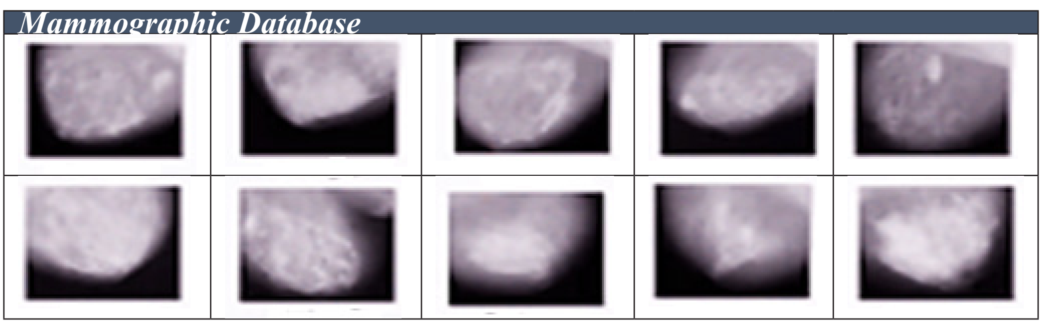
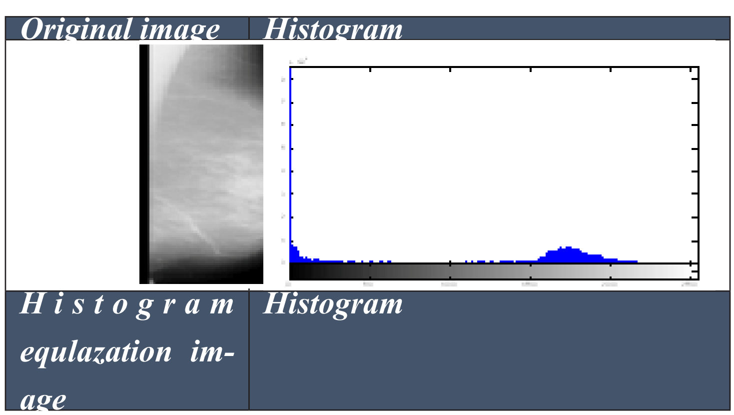


Figure (3): Mammographic Database

Histogram Equalization:Real life images are taken under different illumination conditions, which may affect the distinguishing ability of the patterns and features in the image. Thus, it is important to adjust the intensity of each pixel image through histogram equalization as shown in this Figure (4).



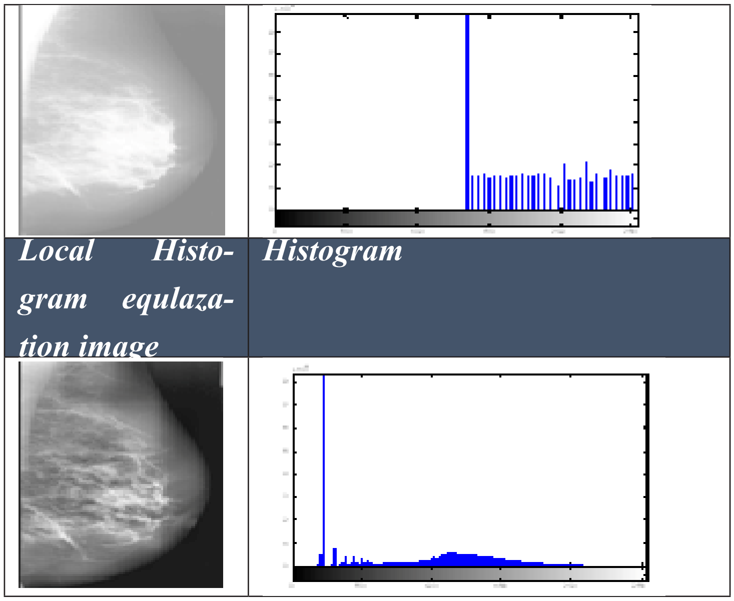


Figure (4):Histogram Equalization

- ***CNN Classification***:The result of CNN detection of breast cancer divided three phases malignant, benign and normalas shown in this Figure (5) and Table (2) as shown Evaluation performance of CNN.

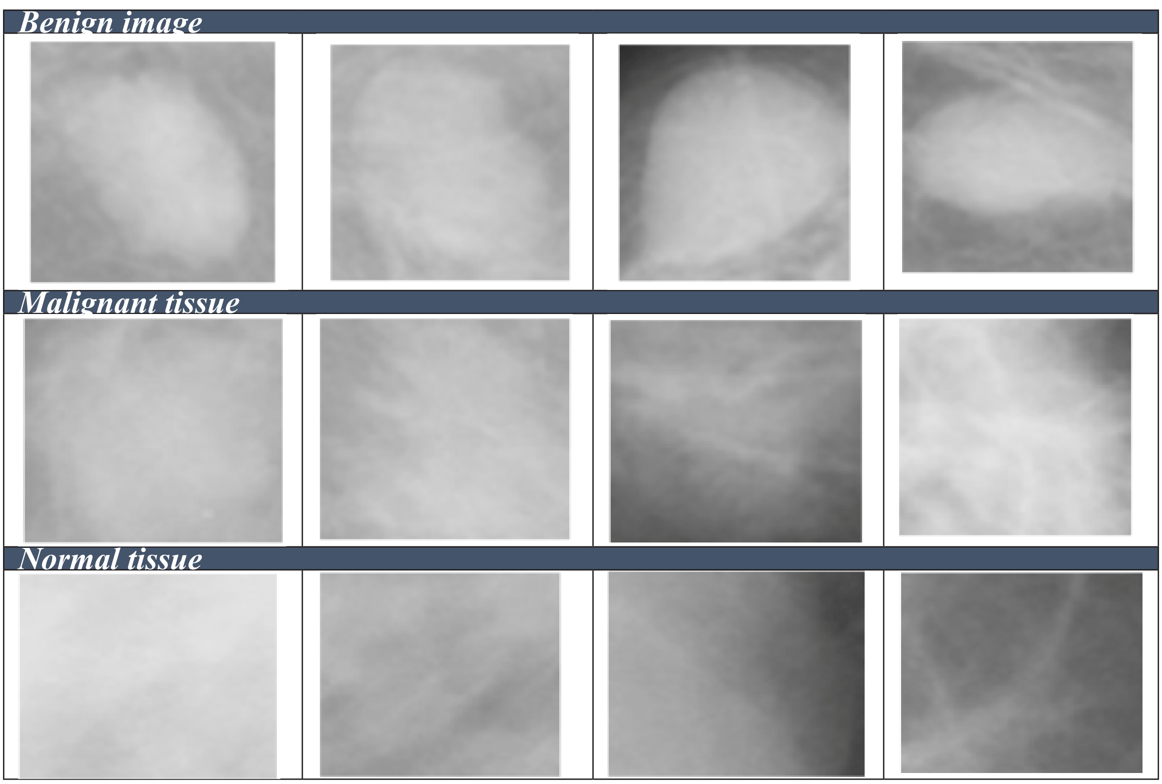


Figure (5): Breast cancer (malignant, benign and normal)

Table (2) Evaluation performance of CNN

breast cancer (malignant, benign and normal)

Sensitivity	45.32%	63.32%	95.1%
Specificity	45.43%	63.35%	99.3%
Accuracy	48.45%	63.40%	99.4%

Conclusion

To reduce the mortality rate because of breast cancer, it is very important to identify cancer in the initial phase. Mammography from the mini-MIAS database is used in this paper. This database consists of 322 mammograms images, of which 52 are cancerous and 270 are normal. Ten texture features from CNN were calculated along 0 °, and the feature space is reduced to six features using the Classification Features method. Results show 100% accuracy for validation and test data, and total accuracy achieved using the proposed method is 99.4%.

Financial Disclosure: There is no financial disclosure.

Conflict of Interest: None to declare.

Ethical Clearance: All experimental protocols were approved under the niversity of information technology and communications, Baghdad, Iraq and all experiments were carried out in accordance with approved guidelines.

References

1. Breast cancer network Australia," Current Breast Cancer Statistics in Australia" Australian Institute of Health and Welfare. Cancer in Australia,2017
2. Faozia AS. Alsarori " Automatic Detectionof Breast Cancer in Mammogram Images", thesis, M.S. Department of Computer Engineering, Cankaya ,university , 2013.
3. Henry J, Mankin MD, et al " Diagnosis, Classification, and Management of Soft Tissue Sarcomas", paper, Dorothy Fox. Fifth Avenue Springtime. Watercolor, 2005; 12(1**).**
4. Bozek J ,Mustra M ,Delac K ,Grgic M. "A survey of image processing algorithms in digital mammography" .In:GrgicM,DelacK,GhanbariM,eds.Recent advances in multimedia signal processing and communications. Berlin:Springer, 2009;631–657
5. Bozek J ,Mustra M ,Delac K ,Grgic M. "A survey of image processing algorithms in digital mammography" . In: GrgicM, DelacK, GhanbariM,eds.Recent advances in multimedia signal processing and communications. Berlin:Springer, 2009;631–657.
6. Sharkas M, Al-Sharkawy M, Ragab DA ."Detection of micro calcifications in mammograms using support vector machine". In: UKSim 5th European symposium on computer modeling and simulation. DOI10.1109/EMS.2011; 23**.**
7. Jawad N. The Application of Image Processing and Machine Learning Techniques for Detection and Classification of Cancerous Tissues in Digital Mammograms", thesis, Faculity of computer science and information technology university of Malaya Kuala Limpur, 2013**.**
8. W Peng , RV, et al. An Automated Confirmatory System for Analysis of Mammograms", thesis, Computer Methods and Programs in Biomedicine.2011.
9. Jain A , Levy D," Breast mass classification using deep convolutional neural networks".In:30 th conference on neural information processing systems(NIPS2016). Barcelona,Spain. 2016;1–6.
10. A Gibson , J Patterson, "Deep Learning", O'Reilly Media, Inc. 2017**;**
11. F Lie , J Johnson. Convolutional Neural Networks for Visual Recognition. Stanford University, 2017.
12. J Bouvrie. Notes on Convolutional Neural Networks", Cambridge, MA. 2006**.**
13. S Hijazi, R Kumar, C Rowen. Using Convolutional Neural Networks for Image Recognition", IP Group, Cadence, 2015**.**
14. S Hijazi, R Kumar, C Rowen. Using Convolutional Neural Networks for Image Recognition. IP Group, Cadence, 2015.
15. V Hlavác. Classifier performance evaluation", Czech Technical University in Prague.2015**.**