International Journal for Modern Trends in Science and Technology, 7(02): 163-173, 2021

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0702028
Available online at: http://www.ijmtst.com/vol7issue02.html

Breast mass diagnosis in Magnetic Resonance Images using AlexNet

Manjula Devarakonda Venkata[1] | Sumalatha Lingamgunta[2]

[1]Research Scholar, JNTU Kakinada, Department of CSE, Amalapuram Institute Of Management sciences & College Of Engineering, Mummididvaram, India.
[2]Professor, University College of Engineering, JNTUK, Kakinada, Andhra Pradesh, India.

To Cite this Article
Manjula Devarakonda Venkata and Sumalatha Lingamgunta, "Breast mass diagnosis in Magnetic Resonance Images using AlexNet", *International Journal for Modern Trends in Science and Technology*, Vol. 07, Issue 02, February 2021, pp: 163-173.

Article Info
Received on 20-January-2021, Revised on 19-February-2021, Accepted on 21-February-2021, Published on 27-February-2021.

ABSTRACT

The earliest breast cancer diagnosis is critical. A new computer-aided detection technology (CADx) is proposed in this manuscript using the technique to identify benign and malignant mass tumours in breast Magnetic Resonance Images. Data normalization for image de-noise is used in this CADx scheme. For feature extraction, the Deep Convolution Neural Network (DCNN) is used. A well-known DCNN architecture called AlexNet is used to define two classes and is fine-tuned. The measurements are obtained using the following publicly available datasets:' QIN Breast DCE-MRI' and 'QIN-Breast' from the Cancer Image Database TCIA for 67 patient records. A high degree of accuracy is provided by practicing on a large amount of data. Nevertheless the biomedical datasets contain a relatively limited number of samples attributable to minimal patient volume. Data augmentation is often a means to maximize the size of the input data by extracting additional information from the original input data. There are many ways for data augmentation; the one used here is rotation and flipping. The precision of the newly-trained deep CNN architecture is 87.5 percent, 86.2 percent sensitivity and 87.7 percent specificity, 0.94 (94 percent) for malignant images, the highest Area Under the Curve (AUC) achieved.

KEYWORDS: *Deep Convolution Network(DCNN), Computer Aided Diagnosing (CADx), Data Augmentation*

INTRODUCTION

The most common type of cancer diagnosed in women around the world is breast cancer, which is also the main cause of harm in women. According to the World Health Organization, the number of cancer cases expected for 2025 will be 19.3 million cases. For countries with low and medium incomes, the death rates are comparatively high compared to developed nations. Cancer in India particularly breast cancer, is a growing concern.

Breast cancer may be treated safely with an early diagnosis. It is therefore important that appropriate techniques are given to screen for early signs of breast cancer. The research involves various imaging methods, the most common of which are thermography, ultrasound (US), mammography and MRI (magnetic resonance imaging). Mammography and ultrasound sonography are the most effective early detection tools. It is not especially fruitful except for situations where the breasts are dense. Mammography alone as well as mammography combined with breast ultrasound, seems ineffective for early detection of small breast cancer tumors in women. When MRI is used for cancer diagnosis, slightly higher sensitivity and a more

positive level can be achieved. MRI is a versatile instrument: it can detect cancer that is not visible in traditional imaging, it can be used as a problem-solving tool, and can be used to screen patients at high risk. When tracking the reaction to chemotherapy, breast MRI is still safer than other imaging modalities used today. It can change the treatment plan in 15-30% of patients with breast cancer. In MRI, enhancing lesions are classified into three major categories:1) focus/foci,2) mass, and3) non-mass enhancement areas.

Focus (or if many, foci) is a region of improvement that is too small to define, measuring less than 5 mm in diameter. A mass is a lesion in three dimensions that fills a space within the breast. We look at its shape, its margins and its internal features, much as in mammography and ultrasound. Non-mass like enhancement are areas of enhancement without a detectable three-dimensional mass shown in Fig 1.

A mass may be either malignant or benign. A mass can be circular, oval, or irregular, or lobulated. Lobulated masses have contours that are undulating. The unequal form of irregular masses can not be described as circular, oval, or lobulated. When a mass is formed irregularly, it has a higher chance of becoming malignant.

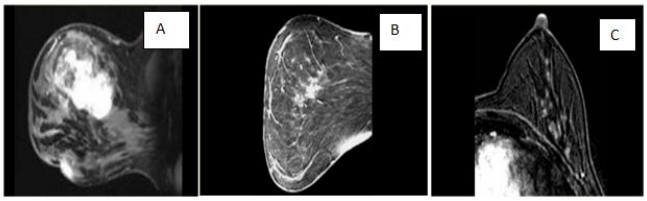


Fig1: Examples of MRI images A) Mass with an irregular shape B) Non-mass enhancement C) Little bright objects: Foci

Key Convolution Neural Network Aspects:

The use of CNNs is one of the most effective deep-learning techniques and they are common in the analysis and classification of images. Their architecture consists of the model of the perceptron, which means it has fully interconnected layers, with each neuron linking to the next layer with each other neuron. It is interesting to examine all the different types of layers that it contains.

First, there is a convolution layer, the name of which is identical to the neural network name. For CNNs, this particular type of layer is fundamental, because it creates activation maps. This means that the algorithm learns the patterns of the image and helps to classify any new image based on these patterns [38,39].

Secondly, the pooling layer is very important for the down-sampling of the image and the removal of any noise that may ambiguous the algorithm. Secondly, the pooling layer is very important for the down-sampling of the image and the removal of any noise that may blur the algorithm. It functions on a particular threshold where any pixel value is held higher than the threshold and any value smaller than the threshold is discarded.

The last layer of a CNN is a fully connected layer that flattens the output of the previous node, which by extension, means that the layer is transformed into a vector, so that it is ready for output. Following this, our algorithm gives each picture a name. As the first one does the vectorization, there are subcategories of the completely connected layer and the last one gives the probabilities for each of the given class [40–42]. The concept of deep convolutionary learning (DCNN) was developed to extract valuable data from the raw images and use it efficiently for the process of classifying tumors in order to solve the problems of traditional imaging methods as described above. This manuscript proposes a new CADx method to distinguish benign and malignant mass lesions from MRI samples.

Literature Analysis: Related Breast Cancer Diagnosis Study Using Convolutional Neural Networks (CNNs)

Previous research explored the inclusion of machine learning methods in breast cancer computer-aided diagnosis and image recognition with respect to the medical image area that deals with breast cancer diagnosis [8,16-20]. The state-of-the-art machine learning approaches for computer-aided detection of breast cancer deliver a broad variety of studies on the existing status of CAD structures, taking into account image modalities and machine learning classifiers [17,18,21,22]. Latest analysis articles in[23-26] have provided a systematic study of both conventional ML and DL literature for relevant implications in the diagnosis of breast cancer. In addition, the bibliographic analysis in the study [26] presented informative features of some well-known DL networks in the diagnosis of breast cancer.

There are several recent research studies exploring the field of computer-aided biomedical images used in the diagnosis of breast cancer that demonstrate the beneficial features of deep convolution networks [27-29]. In fact, CNN [30] was used for the classification of breast cancer histopathology photographs into two separate groups, benign and

malignant. Suzuki et al.(2016) introduced CNNs in computer-aided diagnosis for mass detection in mammographic images[31] Spanhol et al.(2016) used CNNs from histopathological images to identify breast cancer[32], while Wichakam and Vateekul [16], a mixture of deep convolution networks and SVMs for optical mammogram mass detection[33]. The use of deep learning and non-negative matrix factorization for breast cancer identification in mammograms[34] was investigated by Swiderski et al. Kallenberg and his colleagues investigated the strengths of unsupervised deep learning for segmentation of breast density and mammographic risk scoring in the same year[35]. Dealing with image detection for breast cancer, CAD In screening mammography, technologies are fundamentally used for mass detection [20,27,36,37]. In addition, CNNs have mostly been used for classification and segmentation case studies[17,18,25-28] in mammography images and computed tomography (CT) images. In addition, on the basis of US photos, deep CNNs and transfer-learning models were used for the prediction of lymph node metastasis in patients with primary breast cancer[38], as well as on the basis of histo- pathological images [39].

A recent research shows that deep learning strategies are deeply incorporated into the field of CAD medical image analysis; recognition, identification, segmentation, registration, recovery, image production and enhancement, while the effective application of deep learning to medical imaging tasks is extensively examined[[1-5]. Specifically, utilizing convolution neural networks (CNNs)[6-9], deep learning in medical imaging was accomplished. CNN is a deep, feed-forward artificial neural network capable of extracting an image's characteristics and classifying the image using this function. The objective of CNNs in all application fields is by means of a matrix, to transform the input into a feature vector and to align it with qualified feature vectors[10].

The AlexNet network (2012) [11], which has a rather similar architecture to LeNet, by Yann LeCun et al., but is broader, with more filters per layer and includes stacked convolution layers, are some big and common CNN models, implemented in published research papers in the context of image processing and deep learning, consisting of 16 convolution layers and its very uniform architecture, the VGGNet16(2014) makes it very attractive[12]. The ResNet model, which implements the residual learning building block for extremely deep convolution networks[13], the DenseNet (2017)[14], which provides the primary advantage of alleviating the issue of gradient disappearance with the direct link of all layers. Latest review studies have compiled all the major and most interesting applications of deep learning in medical image recognition and segmentation [1-3,8]. It is clear that much study on multi modality breast cancer photographs such as mammograms and ultrasound images has also been conducted in recent years for the identification and diagnosis of breast cancer. Table 1 summarizes several representative CNN software for breast cancer diagnosis. Our work, however is devoted to the investigation of deep-learning algorithms using SVM on MRI breast cancer images.

Table1: Overview of representative breast cancer (BC) diagnostic convolution neural networks (CNNs).

Modality	ML Operation	Reference
Histopathological	Classification	Kumar K. et al. [30]
Mammograms	Detection	Swiderski, B. et al. [34]
Histopathological	Classification	Spanhol, F.A. et al.[32]
Mammograms	Mass detection	Wichakam, I. et al. [33]
Mammograms	Mass detection	Suzuki, S. et al.[31]
Mammograms	Segmentation	Kallenberg, M. et al.[35]
Mammograms	Mass detection	Fenton J.J. et al. [37]
Mammograms	Mass detection & classification	Cheng H. et al. [20]
Mammograms	Classification	Chougrad, H et al.[36]
Ultrasound	Prediction	Zhou Li-Q. et al. [38]

Methodology

The use of CNNs is popular in medical analysis, since we can identify whether or not a woman has breast cancer using an automatic algorithm. Five processing steps are used in the CNN system for the classification of breast cancer (see Figure2): data collection, data pre-processing to normalize the image data obtained, CNN preparation, CNN validation and, finally, CNN checking, accompanied by the assessment of the effects of the classification as presented The five stages of the proposed methodology are thoroughly presented in the following.

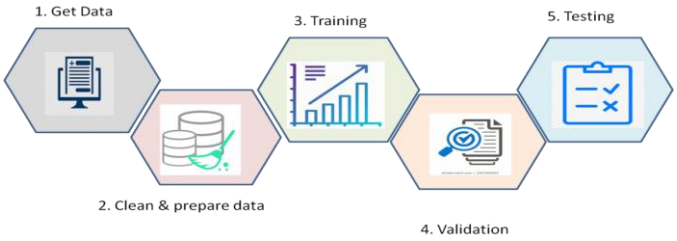


Fig 2: Five steps of Proposed Methodology

1. Having Data: Demography for Dataset:

'QIN Breast DCE-MRI' and 'QIN-Breast' were included in two public datasets:
Mean SD Age: 49.6 ± 11.9 years

Range: Years 21-79

Number of Patients: 67

For this work, benign/malignant prevalence used: Benign: 141, Malignant: 167

Benign Scale (mm) of lesion: Mean: 8.86, Median: 7.33, Range: 3.38-42.88

Malignant Scale (mm) of lesion: Mean: 17.9, Median: 14.9, Range: 3.37-73.77

You can download the public data collection at "https://wiki.cancerimagingarchive.net/display/Public/QIN+ Breast+DCE-MRI".
"https://wiki.cancerimagingarchive.net/display/Public/QIN-Breast ".
The photos were in RGB mode (red, green, blue) and stored in the memory of the PC. We gave them the equivalent prefix according to their class, whether they were mild or malignant. Inside the deep-learning algorithm, this technique allowed us to classify them and pass the right output to them. We provided '0' directly to malignant images as production and '1' to benign images.

2. DATA CLEAN & PLAN

Phase 2.1: Normalization of data

In a machine learning method, Data Normalization is used. This technique rescales the data values within the 0 and 1 ranges, although reliably knowing or estimating the minimum and maximum measurable values is essential. As for the efficiency of the systems being tested, the Min-Max normalization approach has been chosen as the most common method. In addition, by examining the related literature, studies have shown that the accuracy of the findings from this process is higher than that provided by other methods of normalization [38]. The degree to which all images are below zero and one often removes outliers that annoy the algorithm.

Phase2.2: The Swap of Data

The probability of selecting incorrect samples for our algorithm comes from separating the dataset into training and research, so the solution to this issue is given by the implementation of the shuffling process. That gives the data a random order.

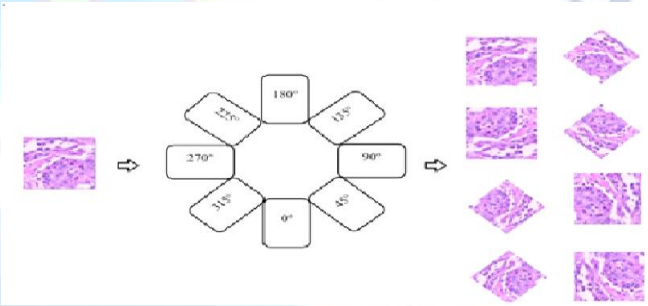
Phase2.3: Data Splitting

We divide the image knowledge available into three sections, which are training, validation and checking. Every partition has a particular percentage taken from the entire dataset, which means that 20 percent of the dataset has been used for testing and 80 percent is divided into 75 percent training and 5 percent for validation for two other sections. Three distinct datasets are then generated: Training, Validation and Testing, respectively.

3. Training

Phase3.1: Augmentation of Data

When we have a limited volume of data, the data augmentation approach is commonly used because we do not want to construct a model that does not have generalization capability. For eg, with rescale rotation range, zoom range and flip techniques, we generate variant images and avoid overfitting. In this process, rotation and flips were applied: rotation by 0◦, 45◦, 90◦ , 135◦, 180◦ , 225◦, 270◦, 315◦ for each one. This increased the amount of training data by a factor of eight.



MRI Patch Rotation angles Augmented patches
Fig 3: Rotation Process

Phase 3.2: Identify architecture for CNN

An experimentation approach is adopted to determine the correct design of a CNN for image classification. The most significant parameters that can contribute to successful network design have been discussed in the experimental and trial phase; The number and shape of convolution layers, the number of pooling layers, the number of iterations, the rate of learning, the function of activation(ReLu) and initial weights and biases. We are using the architecture of AlexNet for the suggested technique here.

From AlexNet

DCNN has found progress in topics relating to image recognition, including image processing as in (Han et al., 2015; Zabalza et al., 2016). A convolution neural network (CNN) consists of several training stages, followed by a supervised classifier and collections of arrays called feature maps, stacked on top of each other(LeCun, Kavukcuoglu & Farabet, 2010). For the development of CNN architectures, three major types of layers are used;(1) convolution layer,(2) pooling layer, and(3) fully connected (fc) layer (Spanhol, 2016).

CiFarNet (Krizhevsky,2009, Roth et al.,2016), AlexNet (Krizhevsky, Sutskever & Hinton,2012), GoogLeNet (Szegedy et al.,2015), ResNet (Sun,2016), VGG16, and VGG19. There are many CNN architectures, such as The most widely used architectures, however are AlexNet, CiFarNet and Inception V1(GoogleNet), Compared to the other deep learning approaches, AlexNet has five convolution layers, three pooling layers and two entirely linked layers with approximately 60 million free parameters, the AlexNet architecture achieved substantially better performance((Krizhevsky, Sutskever & Hinton, 2012). The conv 1-5 layer in Fig. 4 are the layers of convolution. Each neuron in the convolution layers calculates a dot product related to the input volume between their weights and the local area (Krizhevsky, Sutskever & Hinton, 2012). The layers for pooling are pool1, pool2, and pool5 as seen in Fig. 4. In order to decrease the amount of measurement and increase the robustness, these layers conduct a down sampling procedure along the spatial dimensions. They execute a kind of lateral inhibition in the brain that is seen (Krizhevsky, Sutskever & Hinton, 2012). In addition, as seen in Fig. 4, the completely connected layers are fc6, fc7, and fc8. As in ordinary feed forward neural networks, neurons in the entirely linked layer have complete similarities to all neurons in the previous layer. The architecture of Alexnet is illustrated in figure 4.

Fig 4. Fine tuning the AlexNet Architecture

Phase 3.3: The functions of the CNN architecture are described.

The key functions of the CNN architecture are the activation function, which is the function that determines the output of the layer, and the loss function, which is the function used to maximize the weights of the network, following a search in the related literature. ReLU (rectified linear unit) is a triggering function that adjusts all negative values of the input to zero. The ReLU layer boosts and simplifies training and calculations, and helps avoiding the problem of vanishing gradient. It is mathematically defined as: f(x) = max(0, x). The neuron input is represented by x. Conversely, other triggered functions involve parametric ReLU, randomized ReLU, leaky ReLU, tanh and the sigmoid functions. In the decision layer Softmax-Regression technique is utilized. The

equation for softmax function

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad (1)$$

where σ = Softmax , Z-> = Input Vector, e^{zi} = Standard exponential function for input vector, k= number of classes in multi-class classifier, e^{zj} = standard exponential function for output vector.

Phase3.4: CNN's Training

Training a CNN, as the name implies, involves training the model with photos that provide the ability to understand patterns, as the more distinct the training is the better the model becomes. When the algorithm attempts to construct a function that represents the desired relationship, based on the training data, the training process starts. Afterwards, based on this function, it makes predictions (it is an error function) and moves to the validation stage.

4. VALIDATION

For the validation process, the validation dataset is used exclusively. The weights are normalised during validation, and the algorithm makes assumptions on data that is already known. The primary goal during the validation process is to minimise the error that shows the estimation efficacy of the system proposed.

5. TESTING/EVALUATION

The best trained model is deployed for the testing/evaluation process after the completion of the preparation and validation phases [39,40]. Using the test results, as initially broken, which is totally unfamiliar to the model, the validation process for the CNN model is completed. During the testing point, we verify if all recognizable images can be correctly labeled by our classifier, thereby determining the predictive potential of the current model. The classifier makes assumptions about the class of each image and eventually compares the expected class measured to the actual class. Next, some well-known and common performance metrics are computed for classifier evaluation, such as the model's testing precision, accuracy, recall, sensitivity, specificity and f1-score, and followed by the use of an error/confusion matrix for further model evaluation[39,40,41].

The confusion matrix

A special table visualizing the classifier's output is the uncertainty matrix. Typically, the confusion matrix is known as the error matrix in the field of machine learning. An picture area is said to be positive or negative depending on the data type. In addition, either correct (true) or incorrect (fasle) may be an option for the detected result. The judgement would then fall into one of four possible categories: true positive (TP), true negative (TN), false positive (FP) and false negative (FN), respectively. The diagonal of the matrix of uncertainty is the right judgement. An example of the uncertainty matrix for the classification of the two groups is given in Table 2.

Table 2 :Example for confusion matrix

	class label Predicted	
Class label	Normal	Abnormal
Normal	TN	FN
Abnormal	FP	TP

Accuracy

Accuracy is the indicator of a successful prediction made by the classifier. This provides the output potential of the entire classifier. The accuracy is defined as

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (2)$$

The TPR and the FPR are also called sensitivity (recall) and specificity, respectively. They are defined as

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

Receiver operating characteristic (ROC)

ROC analysis is a well-known method of assessment for mission identification. Second, a medical decision-making ROC analysis was used; therefore, it was used in medical imaging. The ROC curve is a graph of operating points that can be viewed as a true positive rate (TPR) graph as a function of the false positive rate.

Area under the ROC curve (AUC)

The AUC is used in the medical diagnostic method and it offers an approach for testing models based on the average of each point on the ROC curve. The AUC score should always be between'0 'and' 1 'for a classifier result, the model with a higher AUC rating provides a better performance of the classifier.

Precision

Precision is the percentage of positive observations predicted accurately to the positive observations predicted overall. Strong accuracy corresponds to a low FPR. The accuracy is determined using the equation below.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

F1 score is the weighted average of precision and recall. It is used as a statistical measure to rate the performance of the classifier. Therefore, this score

takes both false positives and false negatives into account. F1 score is defined as in Equation

$$F1\ score = \frac{2*Recall\ *Precision}{Recall\ +Preccesion} \quad (6)$$

Experimental Configuration

The AlexNet is used to classify the MRI breast images, offering the potential for each image to be either benign or malignant in one of the two classes. This data was downloaded from TCAI (The Cancer archive imaging). Using the MATLAB® 2018b environment, the proposed model is created and educated.

The output of convolution layer and Pool layer are calculated using Equation (7) & (8).

The output size of the convolution layer = $\left(\frac{input\ -filter\ size}{stride}\right)+1$ (7)

The output size of the pool layer = $\left(\frac{output\ of\ conv\ -pool\ size}{stride}\right)+\ 1$ (8)

Results Presentation

To train the AlexNet, the maximum number of Epochs was set to 20. The input layer of the AlexNet architecture requires that the size of the image is 227x227x3. Therefore, there is a pre-processing step to convert all the input images regardless of their sizes to the size required by the AlexNet.

As mentioned in the publication of the architecture used in this work(6), one of Deep Learning's key issues, there are two strategies that this approach applies to face overfitting. The first is data augmentation, since the introduction of rotations and reflections on the input images means that different copies of the same samples are supplied, so that the model knows more general characteristics. The other is the architectural application of dropout layers. Dropout considered as 0.1 introduces regularization within the network, which ultimately improves generalization by randomly skipping some units or connections with a certain probability. In NNs, multiple connections that learn a non-linear relation are sometimes co-adapted, which causes overfitting. The principle is to "turn off some of the neurons randomly, so local dependencies are prevented in the network, making the model more stable. In order to maximize the training samples, data augmentation was added to all the mass samples used from the data collection. Using the rotation and flipping method, the samples were increased to eight frames by rotation. The detailed representation of AlexNet architectures and values obtained subsequently in each layer has been shown in Fig 6.

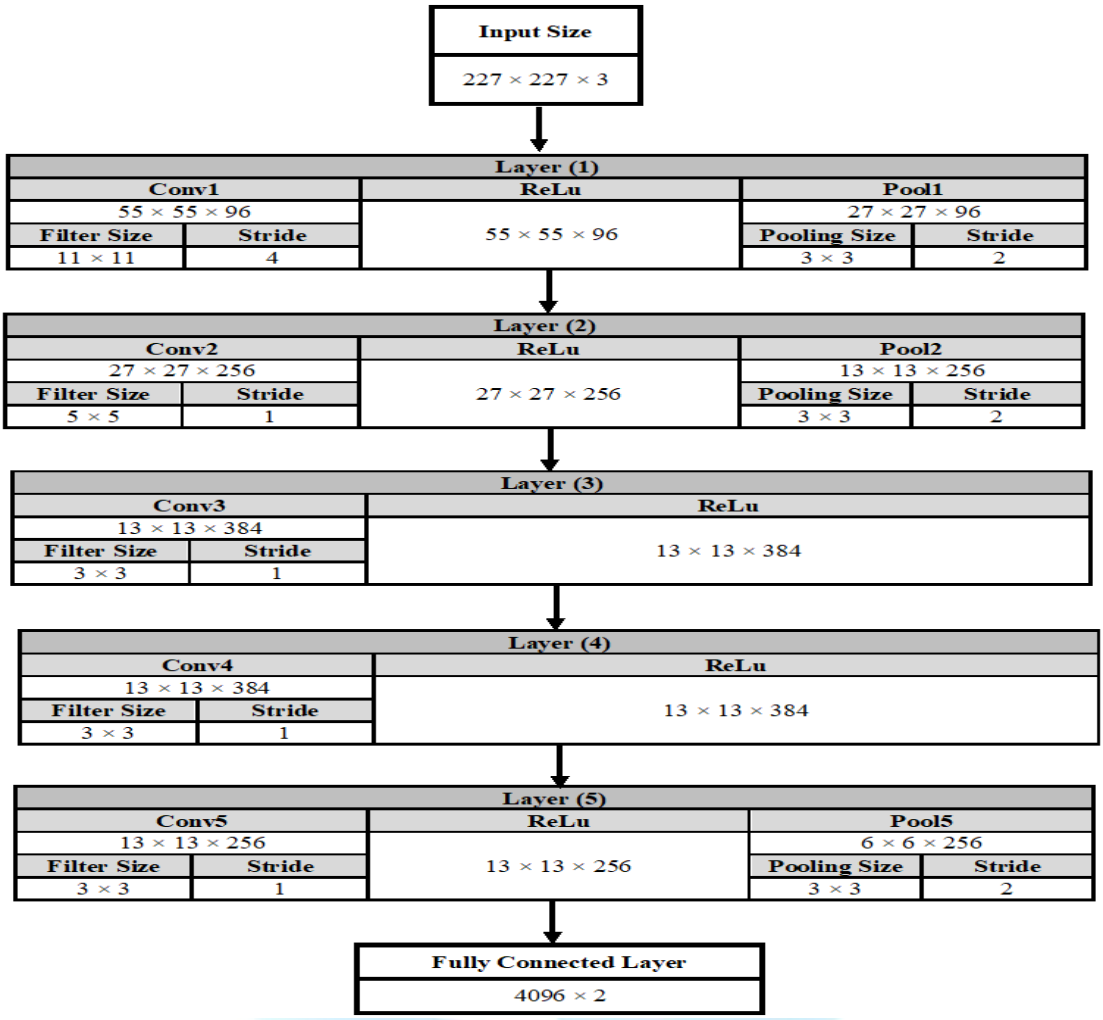


Fig 6: Detailed representation of AlexNet Architrcture

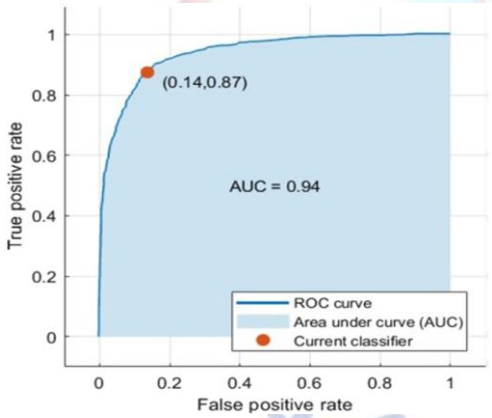


FIg 7: The ROC curve for malignant samples

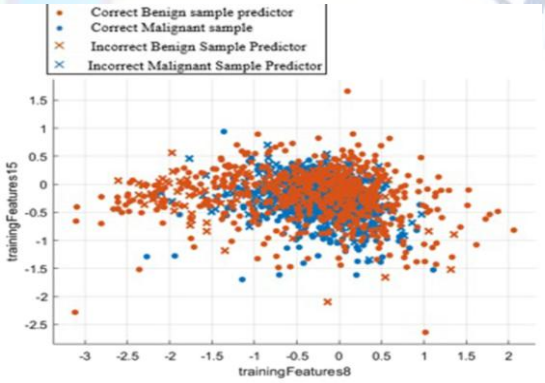


Fig 8: AlexNet classification between benign and malignant masses

Table 6 : Best confusion matrix for proposed CNN

Title	Malignant	Benign
Malignant	35	1
Benign	6	25

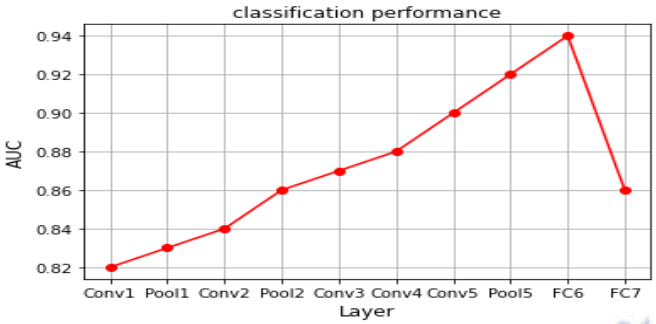


Fig 9 :Classification performance mri malignant lesions, for classifiers based on features from each layer of AlexNet. Fully connected layer 6 ("FC6" in the figure) was selected as the optimal layer for feature extraction, due to its high AUC performance and reduced computational cost.

Table 7: A Comparative view of mass detection methods based in QIN Breast DEC-MRI dataset, including the newly proposed method.

.

Reference	Contribution	Database	AUC
Amit G, Ben-Ari R, Hadad O, Monovich E, Granot N, Hashoul S (2017) [43]	DCNN	DEC-MRI	0.91
P. Herenta, B. Schmaucha(2019) [44]	DCNN _AlexNet classify benign and malignant masses	QIN Breast DEC-MRI	0.89
Roberta Fusco1*, Vincenza Granata1(2020) [45]	DCNN _AlexNet classify benign and malignant masses	DEC-MRI	0.93
Qiyuan Hu,Heather M, whitney (2020) [46]	DCNN	DEC-MRI	0.93
Proposed work	DCNN _AlexNet classify benign and malignant masses	QIN Breast DEC-MRI	0.94

Completion

The purpose of this research was to distinguish the masses and to differentiate benign and malignant tissues from the public dataset of QIN Breast DEC-MRI in magnetic resonance breast cancer images. A new framework for CADx has been suggested. The DCNN was used in the function extraction point. In order to differentiate between two genders, the AlexNet was retrained and its criteria to identify breast images were changed.). The precision of the newly-trained deep CNN architecture is 87.5 percent, 86.2 percent sensitivity and 87.7 percent specificity, 0.94 (94 percent) for malignant images, the highest Area Under the Curve (AUC) achieved. In comparison, the sensitivity, precision, accuracy and F1 score hit 0.871(87.1%) respectively. It was possible to use the proposed CADx method to diagnose other breast anomalies, such as MRI foci and NMEs. Other networks, including the very deep convolution network (VGG) and the residual (ResNet) architecture, will be proposed for future work.

REFERENCES:

[1] Litjens, G.; Kooi, T.; Bejnordi, B.; Setio, A.; Ciompi, F.; Ghafoorian, M.; Laak, J.; Ginneken, B.; Sánchez, C. "A survey on deep learning in medical image analysis". Med. Image Anal. 2017, 42, 60–88.

[2] Biswas, M.; Kuppili, V.; Saba, L.; Edla, D.R.; Suri, H.S.; Cuadrado-Godia, E.; Laird, J.R.; Marinhoe, R.T.; Sanches, J.M.; Nicolaides, A.; et al." State-of-the-art review on deep learning in medical imaging. Front". Biosci. 2019, 24, 392–426.

[3] Shen, D.; Wu, G.; Suk, H.I. "Deep Learning in Medical Image Analysis". Annu. Rev. Biomed. Eng. 2017, 19, 221–248.]

[4] Sahiner, B.; Pezeshk, A.; Hadjiiski, L.M.; Wang, X.; Drukker, K.; Cha, K.H.; Summers, R.M.; Giger, M.L. "Deep learning in medical imaging and radiation therapy". Med. Phys. 2019, 46.

[5] Lundervold, A.S.; Lundervold, A. "An overview of deep learning in medical imaging focusing on MRI". Zeitschrift für Medizinische Physik 2019, 29, 102–127.

[6] Yang, L.; Xie, X.; Li, P.; Zhang, D.; Zhang, L. "Part-based convolutional neural network for visual recognition". In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017

[7] Komeda, Y.; Handa, H.; Watanabe, T.; Nomura, T.; Kitahashi, M.; Sakurai, T.; Okamoto, A.; Minami, T.; Kono, M.; Arizumi, T.; et al. "Computer-Aided Diagnosis Based on Convolutional Neural Network System for Colorectal Polyp Classification: Preliminary Experience". Oncology 2017, 93 (Suppl. S1), 30–34.

[8] R Anwar, S.M.; Majid, M.; Qayyum, A.; Awais, M.; Alnowami, M.; Khan, M.K. "Medical Image Analysis using Convolutional Neural Networks: A Review". J. Med. Syst. 2018, 42, 226.

[9] Gao, J.; Jiang, Q.; Zhou, B.; Chen, D." Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: An overview". Math. Biosci. Eng. 2019, 16, 6536–6561.

[10] Zeiler, M.D.; Fergus, R. "Visualizing and understanding convolutional networks". In European Conference on Computer Vision; Springer: Berlin, Germany, 2014; pp. 818–833.
[11] Krizhevsky, A.; Sutskever, I.; Hinton, G.E." ImageNet Classification with Deep Convolutional Neural Networks; Advances in Neural Information Processing Systems "; Pereira, F., Burges, C.J.C., Bottou, L.,Weinberger, K.Q., Eds.; Curran Associates, Inc.: Dutchess County, NY, USA, 2012; pp. 1097–1105.
[12] Simonyan, K.; Zisserman, A. "Very deep convolutional networks for large-scale image recognition". arXiv 2014, arXiv:1409.1556.
[13] He, K.; Zhang, X.; Ren, S.; Sun, J. "Deep residual learning for image recognition". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
[14] Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. "Densely connected convolutional networks:. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, Hawaii, 21–26 July 2016.
[15] Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: "Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs". arXiv 2016, arXiv:1606.00915.
[16] Nahid, A.; Kong, Y." Involvement of Machine Learning for Breast Cancer Image Classification: A Survey. Comput. Math. Methods" Med. 2017, 2017, 29.
[17] Cheng, H.; Shi, X.; Min, R.; Hu, L.; Cai, X.; Du, H." Approaches for automated detection and classification of masses in mammograms". Pattern Recognit. 2006, 39, 646–668.
[18] Ponraj, D.N.; Jenifer, M.E.; Poongodi, D.P.; Manoharan, J.S. "A survey on the preprocessing techniques of mammogram for the detection of breast cancer". J. Emerg. Trends Comput. Inf. Sci. 2011, 2, 656–664.
[19] Jiang, Y.; Chen, L.; Zhang, H.; Xiao, X. "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module". PLoS ONE. 2019, 14, e0214587.
[20] Sert, E.; Ertekin, S.; Halici, U. "Ensemble of convolutional neural networks for classification of breast microcalcification from mammograms". Conf. Proc. IEEE Eng. Med. Biol. Soc. 2017, 2017, 689–692.
[21] Rangayyan, R.M.; Ayres, F.J.; Desautels, J.L. "A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs". J. Frankl. Inst. 2007, 344, 312–348.
[22] Magna, G.; Casti, P.; Jayaraman, S.V.; Salmeri, M.; Mencattini, A.; Martinelli, E.; Natale, C.D. "Identification of mammography anomalies for breast cancer detection by an ensemble of classification models based on artificial immune system". Knowl. Based Syst. 2016, 101, 60–70.
[23] Yassin, N.I.R.; Omran, S.; El Houby, E.M.F.; Allam, H. "Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review". Comput. Methods Programs Biomed. 2018, 156, 25–45.
[24] Gardezi, S.J.S.; Elazab, A.; Lei, B.;Wang, T." Breast Cancer Detection and Diagnosis Using Mammographic Data: Systematic Review". J. Med. Internet Res. 2019, 21, 14464.
[25] Munir, K.; Elahi, H.; Ayub, A.; Frezza, F.; Rizzi, "A. Cancer Diagnosis Using Deep Learning: A Bibliographic Review. Cancers 2019, 11, 1235.
[26] Chougrad, H.; Zouaki, H.; Alheyane, O. Deep Convolutional Neural Networks for breast cancer screening". Comput. Methods Programs Biomed. 2018, 157, 19–30.
[27] Abdelhafiz, D.; Yang, C.; Ammar, R.; Nabavi, S. "Deep convolutional neural networks for mammography: Advances, challenges and applications". BMC Bioinform. 2019, 20 (Suppl. S11), 481.
[28] "CNNs Applied in Breast Cancer Classification". Available online: https://towardsdatascience.com/convolutional-neural-network-for-breast-cancer- classification-52f1213dcc9 (accessed on 10 November 2019).
[29] Kumar, K.; Chandra Sekhara Rao, A. "Breast cancer classification of image using convolutional neural network". In Proceedings of the 2018 4th International Conference on Recent Advances in Information Technology (RAIT), Dhanbad, India, 15–17March2018: Availableonline: https:// ieeexplore.ieee.org/ abstract/ document/ 8389034 (accessed on 6 January 2020).
[30] Suzuki, S.; Zhang, X.; Homma, N.; Ichiji, K.; Sugita, N.; Kawasumi, Y.; Ishibashi, T.; Yoshizawa, M. "Mass detection using deep convolutional neural networks for mammoghraphic computer-aided diagnosis". In Proceedings of the 55th Annual Conference of the Society of Intruments and Control Engineers of Japan (SICE), Tsukuba, Japan, 20–23 September 2016; pp. 1382–1386.
[31] Spanhol, F.A.; Oliveira, L.S.; Petitjean, C.; Heutte, L. "Breast cancer histopathological image classification using convolutional neural networks". In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 2560–2567.
[32] Wichakam, I.; Vateekul, P. "Combining deep convolutional networks and SVMs for mass detection on digital mammograms". In Proceedings of the 8th International Conference on Knowledge and Smart Technology (KST), Bangkok, Thailand, 3–6 February 2016; pp. 239–244.
[33] Swiderski, B.; Kurek, J.; Osowski, S.; Kruk, M.; Barhoumi, W. "Deep learning and non- negative matrix factorization in recognition of mammograms". In Proceedings of the Eighth International Conference on Graphic and Image Processing, International Society of Optics and Photonics, Tokyo, Japan, 8 February 2017; Volume 10225, p. 102250B.
[34] Kallenberg, M.; Petersen, K.; Nielsen, M.; Ng, A.Y.; Diao, P.; Igel, C.; Vachon, C.M.; Holland, K.; Winkel, R.R.; Karssemeijer, N.; et al. "Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring". IEEE Trans. Med. Imaging 2016, 35, 1322–1331.
[35] Giger, M.L.; Vybomy, C.L.; Huo, Z.; Kupinski, M.A. "Computer-aided diagnosis in mammography". In Handbook of Medical Imaging, 2nd ed.; Breast Cancer Detection and Diagnosis Using Mammographic Data: Systematic Review; SPIE Digital Library: Cardi_, Wales, 2000; pp. 915–1004.
[36] Fenton, J.J.; Taplin, S.H.; Carney, P.A.; Abraham, L.; Sickles, E.A.; D'Orsi, C.; Berns, E.A.; Cutter, G.; Hendrick, R.E.; Barlow, W.E.; et al. "Influence of computer-aided detection on performance of screening mammography". N. Engl. J. Med. 2017, 356, 1399–1409.
[37] Moustakidis, S.; Christodoulo, E.; Papageorgiou, E.; Kokkotis, C.; Papandrianos, N.; Tsaopoulos, D. "Application of machine intelligence for osteoarthritis classification: A classical implementation and a quantum perspective". Quantum Mach. Intell. 2019.

[38] Springenberg, J.T.; Dosovitskiy, A.; Brox, T.; Riedmiller, M. "Striving for simplicity: The all convolutional net", Proceedings of ICLR-2015. arXiv 2014, arXiv:1412.6806.
[39] Theodoridis, S.; Koutroumbas, K.; Stork, D.G. Pattern Recognition; Academic Press: Cambridge, MA, USA, 2009.
[40] Labatut, V.; Cherifi, H. "Accuracy measures for the comparison of classifiers". In Proceedings of the 5th International Conference on Information Technology, Amman, Jordan, 11–13 May 2011.
[41] Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: "A simple way to prevent neural networks from overfitting". J. Mach. Learn. Res. 2014, 15, 1929–1958.
[42] Loe, S.; Szegedy, C. "Batch normalization: Accelerating deep network training by reducing internal covariate shift". In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456.
[43] G Amit, R Ben-Ari, O Hadad, E Monovich, N Granot, S Hashoul : "Classification of breast MRI lesions using small-size training sets: comparison of deep learning approaches" , Medical Imaging 2017: Computer- Aided Diagnosis 10134, 101341H
[44] P.Herent B.Schmauch P.Jehanno O.Dehaene C.Saillard C.Balleyguier J.Arfi-Rouche S.Jégou : "Detection and characterization of MRI breast lesions using deep learning in Diagnostic and Interventional Imaging" Volume 100, Issue 4, April 2019, Pages 219-225
[45] Roberta Fusco , Vincenza Granata , Francesca Maio , Mario Sansone , Antonella Petrillo : "Textural radiomic features and time-intensity curve data analysis by dynamic contrast-enhanced MRI for early prediction ofbreast cancer therapy response": preliminary data Eur Radiol Exp. 2020 Feb 5;4(1):8. doi: 10.1186/s41747-019-01412.
[46] Qiyuan Hu, Heather M. Whitney, & Maryellen L. Giger: "A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI", Sci Rep. 2020; 10: 10536. Published online 2020 June 29. doi: 10.1038/s41598-020-67441-4
[47] Suzuki, K." Machine Learning in Computer-Aided Diagnosis: Medical Imaging Intelligence and Analysis": 9781466600591: Medicine & Healthcare Books; University of Chicago: Chicago, IL, USA, 2012.
[48] Gunn SR"Support vector machines for classification and regression". Available at http:// svms.org/ tutorials/ Gunn1998.pdf , . 1998.
[49] El-naqa I, Member S, Yang Y, Wernick MN, Member S, Galatsanos NP, Member S, Nishikawa RM."A support vector machine approach for detection of microcalcifications". IEEE Transactions on Medical Imaging , 2002. 21:1552_1563.