A learning-based approach to breast cancer screening using mammography images

Khalid Shaikh, Sabitha Krishnan, Rohit Thanki
Prognica Labs, Health Technology Company, Dubai, United Arab Emirates (UAE)

Article Info

Article history:

Received Sep 7, 2021
Revised Dec 24, 2021
Accepted Aug 17, 2022

Keywords:

Breast cancer
Deep neural network
Double filtered
Gray level co-occurrence matrix
Mammogram

ABSTRACT

The current big challenge facing radiologists in healthcare is the automatic detection and classification of masses in breast mammogram images. In the last few years, many researchers have proposed various solutions to this problem. These solutions are effectively dependent and work on annotated breast image data. But these solutions fail when applied to unlabeled and non-annotated breast image data. Therefore, this paper provides the solution to this problem with the help of a neural network that considers any kind of unlabeled data for its procedure. In this solution, the algorithm automatically extracts tumors in images using a segmentation approach, and after that, the features of the tumor are extracted for further processing. This approach used a double thresholding-based segmentation technique to obtain a perfect location of the tumor region, which was not possible in existing techniques in the literature. The experimental results also show that the proposed algorithm provides better accuracy compared to the accuracy of existing algorithms in the literature.

Corresponding Author:

Khalid Shaikh
Prognica Labs, Health Technology Company
Dubai, United Arab Emirates (UAE)
Email: khalid@prognica.com

1. INTRODUCTION

Cancer, which leads to death, is caused by the changes that occur in cells which spread uncontrollably [1]. Mostly, cancer cells form a lump or mass which is called a tumor, and the tumor is named based on the body part where it originates [2]. This cancer produces no pain at its early stage [3], and this leads to the need for screening very often to ease early detection and thereby diagnosis. The majority of lumps discovered during early screening are non-cancerous, whereas 80% of breast cancers are invasive and classified as curable or incurable [4]. Breast cancer is usually referred to as a single disease, but there are several sub-categories [5] and chances of being cured completely among all other cancer types [6]. The initial stage of breast cancer diagnosis is manual screening, which is done by the physicians. If the physician notices any differences in the tissue of the breast, they will recommend computer aided screening, which is breast imaging. Now, once the imaging tells us the possibility of cancer existence, then there comes the need for biopsy, which returns the histopathological status of the tumor [7]. The different kinds of imaging technologies for breast cancer diagnosis are mammography, ultrasound, and magnetic resonance imaging (MRI). Among all these, mammography is gaining popularity because of its procedure, which includes projection of low-dose x-ray through which we can visualize the breast's internal structure [8]. To save the lives of humankind, it is necessary to develop a computer-aided diagnosis (CAD) system which can be used for the early detection of disease as early as possible. This led to the usage of artificial intelligence (AI) in medical science for fast and accurate diagnosis of cancer [9].

Generally, a mammogram will lead to four images, such as a cranio-caudal (CC) view and a mediolateral oblique (MLO) view of the right breast and left breast. Due to these varieties of images, it is very convenient for fast diagnosis [10]. These images are usually adulterated with noise, which can hinder the possibility of an accurate diagnosis. So, this leads to the need for a proper filtration technique that can filter the image for the proper diagnosis. So far, various filtration techniques which are based on multiresolution mathematical transforms have been developed [11]. This outdated the performance of traditional filters, which are based on convolution and arithmetic operations. But these multiresolution filters also suffer from data loss, and due to this limitation, thresholding-based convolutional filters came into existence, which can perform both denoising and segmentation simultaneously [12]. Once the filtration process is done, the system needs to get the features from the segmented region, which is achieved by certain feature extraction techniques. Features are the behavior of an image in terms of storage, efficiency, and time consumption [13]. Any feature extraction will collect the features based on the three broad categories such as color, shape, and texture. Then, it's important to make a machine learning algorithm that can use this data to learn how to classify things [14], [15].

Deep learning is garnering a lot of interest in the field of machine learning since it can learn a collection of high-level properties and deliver high identification accuracy. This is in contrast to traditional machine learning techniques, which use handcrafted features. A method that uses a cascade of deep learning and random forest classifiers was presented by Dhungel *et al.* [16] as a way to identify masses in mammograms. Following the initial step of the classifier, the potentially malicious areas are sent on to the second level of the cascade random forest. During this stage, the morphological and textural aspects are analyzed, and afterward, the surviving areas are merged using connected component analysis. Although this classifier has a high true positive detection rate, it is not successful when applied to big datasets [16]. Instead of designing descriptors to explain the content of mammography images, Arevalo *et al.* [17] utilized a hybrid approach that included the use of convolutional neural networks (CNN) to learn the representation in a supervised manner. This was done in place of the traditional approach of designing descriptors. This approach dispenses with the necessity of coming up with a one-of-a-kind solution for each and every type of data while also producing results that are very accurate. Despite all of these benefits, this method suffers from a significant problem that prevents it from handling huge datasets [17].

Gustvo *et al.* [18] illustrated an automated algorithm for detailed examination of CC and MLO mammography with the use of deep learning models for the problem of jointly classifying unregistered mammogram views and respective segmentation maps of breast lesions. This paper reduces the disadvantage of dealing with large datasets, but this has the disadvantage of relying upon manual labeling for training the dataset [18]. Dubrovina *et al.* [19] CNN to learn discriminative features automatically. This approach solves the problem of difficulty involved in a medium-sized database by training the CNN in an overlapping patch-wise manner, and this approach is faster and maintains classification accuracy. In spite of all these advantages, this algorithm suffers from the issue of instability in the classification process [19]. Hai *et al.* [20] aimed to collect high-end semantic features for training a convolutional neural network and this algorithm then targets optimizing the CNN. They achieved this by combining the extracted multi-level features into one new CNN. This optimization makes the network pay different kinds of attention to different levels of features. Though this seems to be good, this approach again suffers from the issue of large datasets [20]. The main aim of this paper is to develop an algorithm that can utilize the deep neural network (DNN) for the diagnosis of breast cancer for its variety of categories without any supervision or annotation. Also, this proposed algorithm provides better accuracy compared to existing algorithms in the literature [21]–[24]. The rest of the paper is organized such that the working flow of the proposed algorithm along with technical theories is covered in section 2. Section 3 discusses the obtained results by the proposed algorithm and its discussion, and section 4 discusses the work's conclusion.

2. PROPOSED ALGORITHM

This proposed algorithm relieves radiologists of the burden of accurately diagnosing a patient's image in order to determine the status of cancer. This algorithm refines the network twice using the following important process, hence the name double distilled DNN (triple D neural network). The name "double distillation" comes from the fact that it involves refining tumor extract twice. This framework's neural network strategy employs fewer dense layers with proper feature selection, which may result in greater accuracy in breast cancer diagnosis. Figure 1 depicts the entire architecture of the proposed algorithm.

Mammography is a type of medical imaging that uses a low-dose x-ray system to examine the insides of the breasts. A mammography exam, also known as a mammogram, helps women detect and diagnose breast diseases early. This mammogram, which yields four images, screens two breasts for diagnosis. Two of these images are MLO views, while the others are CC views of each breast. One of the standard mammographic views is the MLO view. It is the most important projection because it depicts the majority of breast tissue. The entire breast parenchyma is depicted in the CC view, and the fatty tissue closest to the chest wall appears as a dark strip on the mammogram. The pectoral muscle is shown in this view, and the nipple is shown in profile. Figure 2 depicts various mammogram image views (a) left CC, (b) left MLO, (c) right CC, and (d) right MLO.

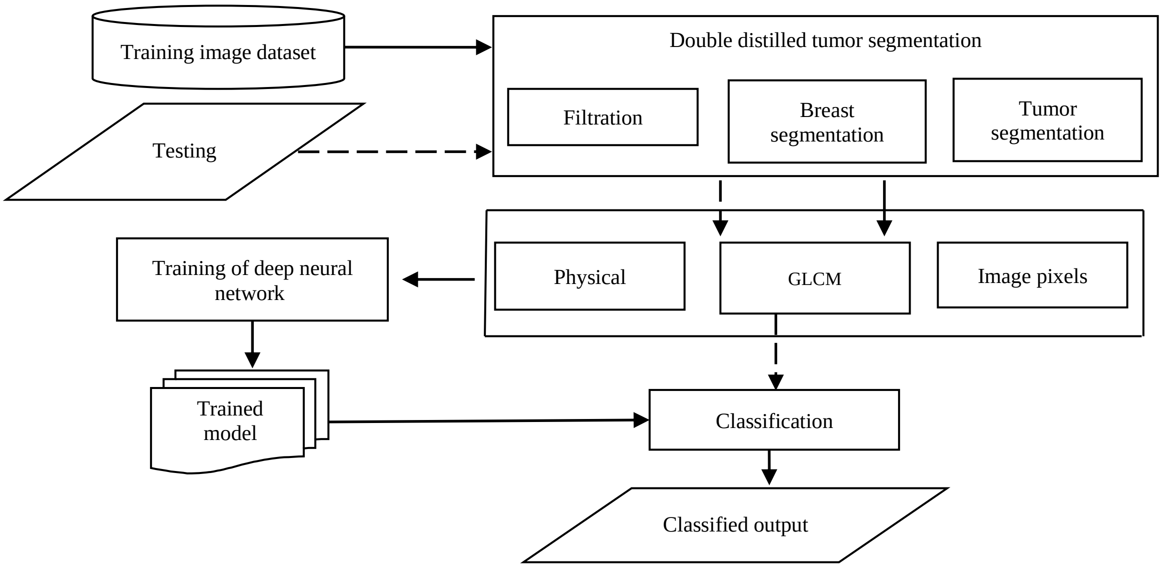


Figure 1. Working flow of proposed algorithm

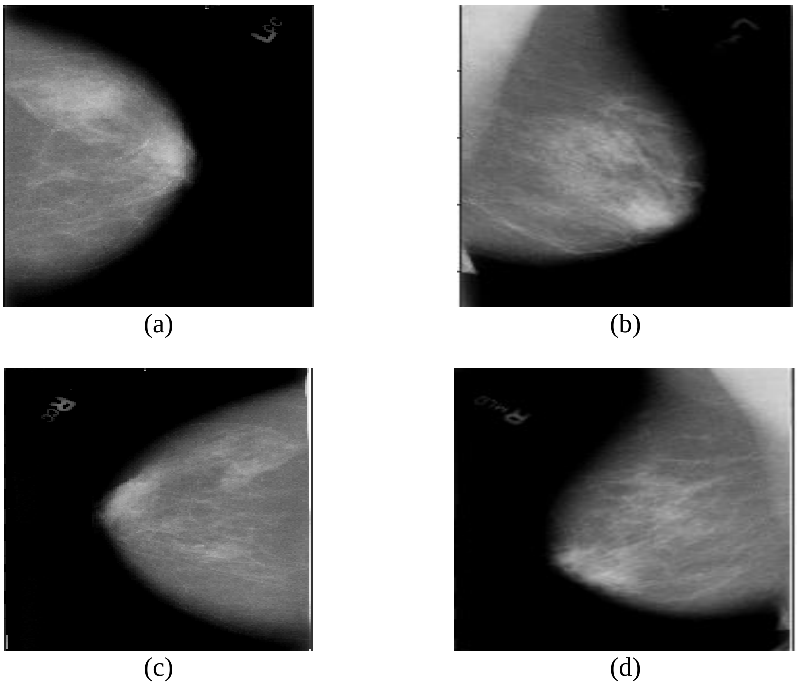


Figure 2. Four different views of mammogram from two breasts (a) left CC, (b) left MLO, (c) right CC, and (d) right MLO

2.1. Double distilled tumor segmentation

The mammogram image is accumulated with lot of noise as it is achieved by contacting the human. So, there is a need for filtration which is carried out by multifiltered and thresholded peripheral equalization.

This algorithm not only filters the image but also it aids in segmentation of breast part completely. So now it is clear that first part of double distillation completes over here since it distills the image for the extraction of breast region. The next turn is to extract the tumor from the region, and it is done by adaptive morphological segmentation. Here the second distillation of double distillation happens, and it extracts the entire tumor region without distortion.

2.1.1. Multifiltered and thresholded peripheral equalization for preprocessing and breast segmentation

Breast deformation is an unavoidable limitation while scanning process of mammography undergoes. Due to this limitation peripheral area of the breast is affected which in turn affects the grey level values of breast tissue. This always results in lesser intensity in peripheral areas than at central area. Physician will adopt for adjusting window settings which is a time eating process. So, this leads to the necessity for image enhancement for proper breast segmentation where the first distillation of triple D framework takes place. Multi-threshold peripheral equalization algorithm is applied over images for image enhancement and automatic segmentation of breast region. This algorithm enhances and eliminates irrelevant information from mammograms. The main necessity of this method is to enhance the contrast of the peripheral area of the mammogram by utilizing multiple thresholds. This process creates multiple images and then averages them to produce the smooth transitions between the central and peripheral areas of the mammogram. Thus, physicians can view and inspect the lesions through one window level setting. Results of breast extraction from mammogram breast images as shown in Figure 3. Figure 3 shows the resultant images of each stage of proposed breast segmentation (a) thresholded image, (b) Gaussian filtered, (c) thresholded multiplied with gaussian filtered, and (d) extracted breast region using peripheral equalization.

The sub steps for this procedure are defined as per below:

– Otsu for breast segmentation (Iseg): Otsu is a global thresholding technique which will select only the breast region for filtering.

$$I_{seg} = otsu(MI) \tag{1}$$

Where, *MI* is a mammogram image, I_{seg} is a segmented breast image.

– Gaussian filtering (I_{filt}): gaussian filter is a filter whose impulse response is gaussian function. Gaussian filters are designed to give no overshoot to a step function input while minimizing the rise and fall time. This behaviour is connected to the fact that the gaussian filter has the minimum possible group delay.

$$I_{filt} = gaussian(MI, sigma) \tag{2}$$

Here sigma denotes the standard deviation of the filter, which is given as 0.1, 0.2, 0.3, 0.4 and 0.5 randomly.

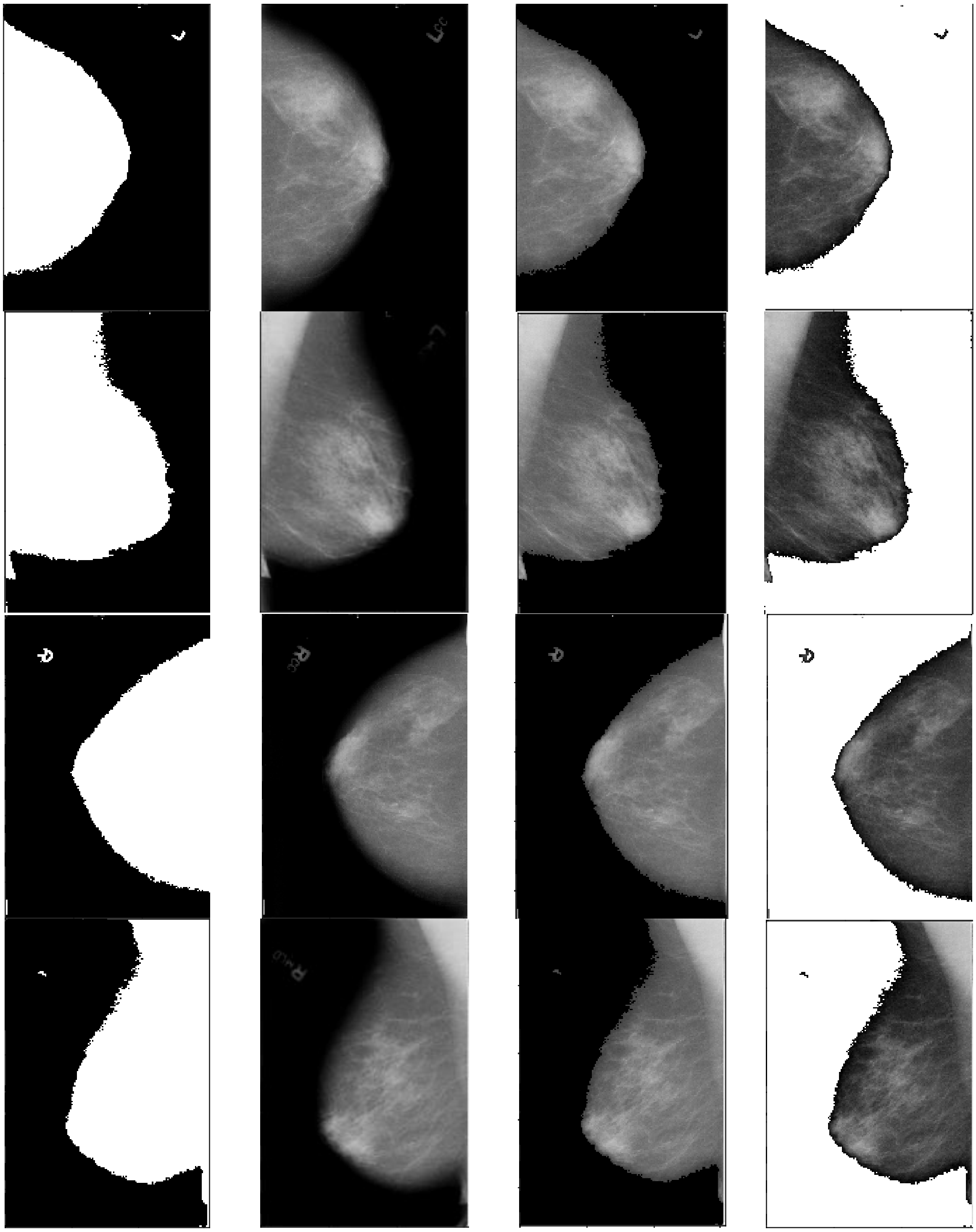
– Multiplication of I_{seg} and I_{filt}: this is done to eliminate the information which lies outside the breast portion of the image.

$$I_{mult} = I_{seg} * I_{filt} \tag{3}$$

– Finding normalized thickness profile (NTP): the steps for finding an NTP are given as per below.
a. Rescale the I_{filt} with different scaling parameter to get I_{mult}(n)
b. Find average of all the filtered images
c. Get the threshold value from the average image
d. Find NTP value using (4)

$$NTP = \frac{1}{5}\sum_{i=1}^{n} I_{mult}(n) \tag{4}$$

– Peripheral equalized image (I_{PE}) using original image (*I*) and NTP: an image with suppressed noise and clearly defined edges is obtained at this stage with the help of NTP.



(a) (b) (c) (d)

Figure 3. Results of multiple stages thresholding and preprocessing of mammogram image (a) thresholded image, (b) Gaussian filtered, (c) thresholded multiplied with gaussian filtered, and (d) extracted breast region using peripheral equalization

2.1.2. Adaptive morphological operation for breast cancer tumor segmentation

Once everything is done for breast segmentation, now it is the turn to segment only the tumor portion which is done by a set of morphological operation where the second distillation of the triple D framework takes place. Figure 4 clearly portrays the overall process of tumor segmentation.

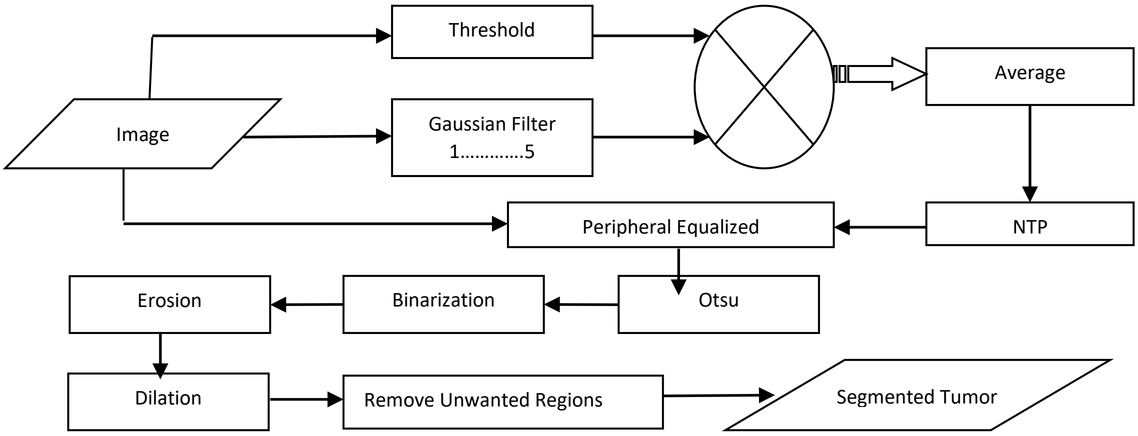


Figure 4. Working flowchart of tumor segmentation from mammogram image

The sub steps of this segmentation procedure are shown in:

– Otsu thresholding: here Otsu thresholding is used again as now it has new image values and it can be applied to any image globally.

$$I_{TH} = otsu(I_{PE}) \tag{5}$$

– Image binarization: image *binarization* is the conversion of gray scale images to black-and-white and dividing into constituent objects. It completely dependent on content of image and it is mainly used to extract an object from an image. By this process, the image will have two divisions namely foreground and background.

$$I_B = \begin{cases} 1, I_P > I_{TH} \\ 0, else \end{cases} \tag{6}$$

Where, I_B is a binary image and I_P is a pixel value of image.

– Erosion process: erosion is one of the two basic operators in mathematical morphology where the basic effect of the operator on a binary image is to erode the boundaries of regions of foreground pixels (i.e., white pixels, typically). Here the binarization yields an image with minute hole which are not needed for the process. So, this will close those holes by a structural element.

$$I_E = I_B \Theta B = \left\{ z \in \frac{E}{B_z} \subseteq I_B \right\} \tag{7}$$

Where I_E is an eroded image, I_B is a binary image to be eroded, B is the binary structural element, z is the vector, or the initial size of the window and E is the area in I_E which comes under z.

– Dilation process: the basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels. Now there arises a situation that of existence small tumor like microcalcification which must be enlarged to its original size, and this is done by the dilation mathematical operator.

$$I_D = I_E \oplus B = \cup_{b \in B} I_E \tag{8}$$

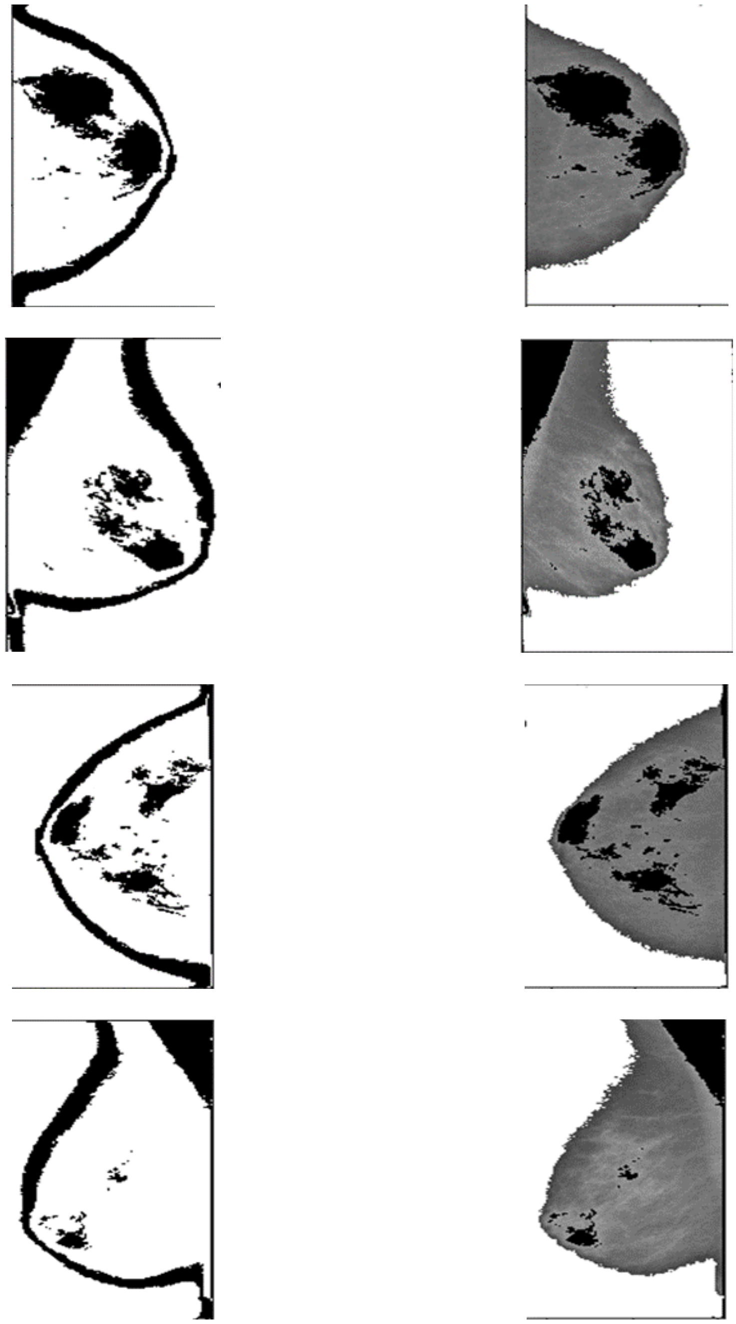
Where, I_E is the image to be dilated, B is the binary structural element, b is the vector or the initial size of the window.

– Removing unconnected regions: this is done to fill holes, to remove some small parts in segmented image which cannot be added as tumor and sometimes pectoral muscles too.

$$I_T = RUC(I_D) \quad (9)$$

Where, I_T is the segmented tumor portion in the image.

Superimposing this segmented region on breast mass: it is important to superimpose the separated tumor over the image so that we can find the exact position of the tumor which can aid in finding the severity of tumor. The resultant images using segmentation procedure are shown in Figure 5. The result of adaptive morphological segmentation from Figures 5(a) segmented tumor in binaryscale and 5(b) segmented tumor in grayscale.



(a) (b)

Figure 5. Result of adaptive morphological segmentation (a) segmented tumor in binaryscale and (b) segmented tumor in grayscale

2.2. Adaptive and versatile feature extraction form the extracted breast tumor

For the efficient classification process, there is a need for first order and higher order features to be collected. So, in this framework first order attributes were computed, and it includes entropy, modified entropy, standard deviation (SD), modified standard deviation (MSD), energy, modified energy, asymmetry, modified skewness, and range value of the histogram. Along with these other features like mean, SD, smoothness, third moment, entropy, skewness, kurtosis, variance, mode, interquartile range, and percentiles or quintiles are also extracted constituting to 28 first order features.

To make the process furthermore efficient spatial inter-relationships of the pixels is carried out and it is done by computing grayscale co-occurrence matrix (GLCM). The 2D histogram of grayscale intensity for a pair of pixels is called the GLCM. The extracted second order features includes energy, contrast, correlation, homogeneity, entropy, maximum probability, inverse different moment (IDM), variance, sum average, sum entropy, sum variance, difference entropy, difference variance, autocorrelation, dissimilarity, cluster shade, cluster prominence, correlation information 1, and correlation information 2. Sometime there are situation when physical features matter. So, this work has concentrated on collecting the physical features as well which includes size, shape, and density of the tumor. So, this works collects large number of features which acts as a strong platform for these unlabeled data to perform unsupervised learning.

2.3. Congregate unsupervised deep neural network

Since there are only unlabeled data, supervised learning is quite impossible, and it may also give more false positives. So, to make it into an unsupervised classifier labeling must be done within the classifier and this will do the clustering based on the similarities among features. This labelling strategy creates a dataset with the features to be trained along with their labels. The primary stage of this network is training where data along with the labels plays the important part. The input and its features now step into first part of the training phase in which labelling takes place and this step is the man aid of this network. Since here labelling happens in the network itself the data of any type and size can be used for the processing. This approach accepts inputs I and its corresponding features fe(I) for training. Now the input and its features are subjected for computing the distance matrix using Euclidean distance and with the ward linkage.

$$D_{euclidean}(x) = \left\| fe(I_i) - fe(I_j) \right\|_2 \tag{10}$$

Based on this the input will selects its closest clusters. Now it is time to select the approximate cluster such that to create the proximal matrix. This cluster selection is carried out by ward linkage which is depicted below:

$$c_i, l_{ward} = \sum_{x \in c_i} D_{euclidean}(x) \tag{11}$$

where c_i is clusters, I is number of clusters and x is input.

Now once the data select its exact cluster the whole process completes and the dendogram is created. If the data fails to find the cluster, then the whole process of calculating Euclidean distance and ward linkage resumes and the process goes on till it find its cluster. This process is an iterative process which yields lward as the label for the data. Now the training data has features of inputs *f(x)* and its labels lward which is fed into the first layer of dense network with window size 12×12. The hidden layer h is described as:

$$h_i = f(w * X + b) \tag{12}$$

Now the hidden layer output is compiled using rmsprop optimizer which would eliminate the space for redundant data thereby improving the accuracy.

$$T_{output} = compile(h_i) \tag{13}$$

3. RESULTS AND DISCUSSION

The performance of the proposed algorithm is verified by using standard mammogram image dataset and some of the performance measures such as accuracy, sensitivity, and specificity. The curated breast imaging subset-digital databased for screening mammography (CBIS–DDSM) dataset [25] are used for training and testing of proposed algorithm. This dataset is updated version of DDSM and contains 2,620 scanned film mammography images. Out of this dataset, in this paper, 280 images are taken as training dataset and 80 images are taken as testing dataset.

The training images are subjected under various algorithms for its prior processing to make the classification more accurate. The algorithm produces various range of accuracy with different number of features and training dataset. The algorithm has been updated at each step by adding or reducing features or training dataset. Table 1 gives the summarized performance of proposed algorithm over different features and different number of images in training dataset. As the Table 1 clearly reveals that low number of features gives good learning to proposed algorithm. Since deep learning requires more space for its training it suffered from overfitting problem, and this leads to low accuracy. Then various experiments were carried out with different number images in the dataset and different number of features. From the Table 1, it is obvious that the algorithm performs better when it has lesser data and lesser features. Thus, finalization was made to train the network with 12 features of 280 images which yields a good training accuracy of 96.1. This is lower than the accuracy of training with 12 features of 250 images, but the variation is negligible. Hence used the last case which can accept good amount of dataset in training a good number of features. The proposed algorithm suffers in its performance measures with higher number of hidden layers. This changes in performance happens due to more hidden layers along with large number of datasets which creates over fitting problem resulting in huge variation of performance measures. The analysis of performance measures with respect to different hidden layers for proposed algorithm for testing dataset are summarized in Table 2.

Table 1. Accuracy of proposed algorithm over different number of features and different number of images in dataset

Number of features	Dataset	Accuracy
29	280	50
29	250	50.5
20	280	55.8
20	250	56.2
19	280	63.5
19	250	79
16	280	73.2
16	250	75.8
14	280	81.9
14	250	82.5
12	280	96.1
12	250	96.50

Table 2. Performance measures with different number of hidden layers for testing dataset

Number of hidden layers	Performance metrics				
	Accuracy (%)	Precision (%)	Recall (%)	Sensitivity (%)	Specificity (%)
4	62	65	56	72	80
3	65	71	62	79	84
2	82	84	74	87	89
1	96	89	84	92	95

The results in Table 2 shows that proposed algorithm gives good accuracy for a smaller number of hidden layers. The performance of proposed algorithm is also compared with some existed algorithms [21]–[24] which are used for feature extraction and detecting of breast cancer tumor. These algorithms were designed using conventional machine learning algorithms such as support vector machine (SVM), decision tree (DT), Naïve Bayes (NB), and k nearest neighbor (KNN). The comparison of algorithms is given in Table 3.

Table 3. Comparison of performance for various learning-based algorithms

Method	Used algorithm	Achieved maximum accuracy
Kim *et al.* (2012) [21]	SVM	0.8458
Park *et al.* (2014) [22]	Semi supervised learning, SVM, NB, and random forest	0.725, 0.528, 0.592, and 0.664
Sountharrajan *et al.* (2017) [23]	SVM, NB, and DT	0.7925, 0.7725, and 0.7725
Abien *et al.* (2018) [24]	SVM and KNN	0.9375 and 0.9357
Proposed	DNN	0.96

4. CONCLUSION

In this paper, an automatic diagnosis algorithm for detecting breast cancer based on clustering based unsupervised learning is presented. The proposed algorithm was designed using thresholding and DNN. The tumor in mammogram image was extracted using Otsu thresholding-based segmentation in this proposed algorithm. The various tumor features which were extracted from the tumor are used about prediction of image like that image has tumor or not. The experimental results show that the proposed algorithm provides accuracy up to 96% for detection of breast cancer. The results also show that the performance of proposed algorithm was better than performance of existed algorithms in the literature.

REFERENCES

[1] "GLOBOCAN 2008: cancer incidence and mortality worldwide–IARC." https://www.iarc.who.int/news-events/globocan-2008-cancer-incidence-and-mortality-worldwide/

[2] "Breast cancer statistics | world cancer research fund international," *WCRF International*. https://www.wcrf.org/cancer-trends/breast-cancer-statistics/ (accessed Sep. 02, 2022).

[3] F. Ilyas, "'Over 40,000 die of breast cancer every year in Pakistan," *DAWN.COM*, Mar. 11, 2017. https://www.dawn.com/news/1319675 (accessed Sep. 02, 2022).

[4] S. Charan, M. J. Khan, and K. Khurshid, "Breast cancer detection in mammograms using convolutional neural network," in *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, Mar. 2018, pp. 1–5, doi: 10.1109/icomet.2018.8346384.

[5] M. Shareef, M. A. Ashraf, and M. Sarfraz, "Natural cures for breast cancer treatment," *Saudi Pharmaceutical Journal*, vol. 24, no. 3, pp. 233–240, May 2016, doi: 10.1016/j.jsps.2016.04.018.

[6] S. S. Gorin, "Delays in breast cancer diagnosis and treatment by racial/ethnic group," *Archives of Internal Medicine*, vol. 166, no. 20, p. 2244, Nov. 2006, doi: 10.1001/archinte.166.20.2244.

[7] I. Khalkhali, I. Mena, and L. Diggles, "Review of imaging techniques for the diagnosis of breast cancer: a new role of prone scintimammography using technetium-99m sestamibi," *European Journal of Nuclear Medicine*, vol. 21, no. 4, pp. 357–362, Apr. 1994, doi: 10.1007/bf00947973.

[8] J. T. Dobbins and D. J. Godfrey, "Digital x-ray tomosynthesis: current state of the art and clinical potential," *Physics in Medicine and Biology*, vol. 48, no. 19, pp. R65–R106, Sep. 2003, doi: 10.1088/0031-9155/48/19/r01.

[9] S. Beura, B. Majhi, and R. Dash, "Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer," *Neurocomputing*, vol. 154, pp. 1–14, Apr. 2015, doi: 10.1016/j.neucom.2014.12.032.

[10] A. J. Bekker, M. Shalhon, H. Greenspan, and J. Goldberger, "Multi-view probabilistic classification of breast microcalcifications," *IEEE Transactions on Medical Imaging*, vol. 35, no. 2, pp. 645–653, Feb. 2016, doi: 10.1109/tmi.2015.2488019.

[11] M. Saha, M. K. Naskar, and B. N. Chatterji, "Mammogram denoising by curvelet transform based on the information of neighbouring coefficients," in *Proceedings of the 2015 Third International Conference on Computer, Communication, Control and Information Technology (C3IT)*, Feb. 2015, pp. 1–6, doi: 10.1109/c3it.2015.7060180.

[12] A. K. Bhandari, A. Kumar, S. Chaudhary, and G. K. Singh, "A novel color image multilevel thresholding based segmentation using nature inspired optimization algorithms," *Expert Systems with Applications*, vol. 63, pp. 112–133, Nov. 2016, doi: 10.1016/j.eswa.2016.06.044.

[13] I. Zyout, J. Czajkowska, and M. Grzegorzek, "Multi-scale textural feature extraction and particle swarm optimization based model selection for false positive reduction in mammography," *Computerized Medical Imaging and Graphics*, vol. 46, pp. 95–107, Dec. 2015, doi: 10.1016/j.compmedimag.2015.02.005.

[14] A. S. Becker, M. Marcon, S. Ghafoor, M. C. Wurnig, T. Frauenfelder, and A. Boss, "Deep learning in mammography," *Investigative Radiology*, vol. 52, no. 7, pp. 434–440, Jul. 2017, doi: 10.1097/rli.0000000000000358.

[15] J. Wang, X. Yang, H. Cai, W. Tan, C. Jin, and L. Li, "Discrimination of breast cancer with microcalcifications on mammography by deep learning," *Scientific Reports*, vol. 6, no. 1, Jun. 2016, doi: 10.1038/srep27327.

[16] N. Dhungel, G. Carneiro, and A. P. Bradley, "The automated learning of deep features for breast mass classification from mammograms," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016*, Springer International Publishing, 2016, pp. 106–114, doi: 10.1007/978-3-319-46723-8_13.

[17] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. G. Lopez, "Representation learning for mammography mass lesion classification with convolutional neural networks," *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 248–257, Apr. 2016, doi: 10.1016/j.cmpb.2015.12.014.

[18] G. Carneiro, J. Nascimento, and A. P. Bradley, "Automated analysis of unregistered multi-view mammograms with deep learning," *IEEE Transactions on Medical Imaging*, vol. 36, no. 11, pp. 2355–2365, Nov. 2017, doi: 10.1109/tmi.2017.2751523.

[19] A. Dubrovina, P. Kisilev, B. Ginsburg, S. Hashoul, and R. Kimmel, "Computational mammography using deep neural networks," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging &amp$\mathsemicolon$ Visualization*, vol. 6, no. 3, pp. 243–247, Mar. 2016, doi: 10.1080/21681163.2015.1131197.

[20] J. Hai *et al.*, "Multi-level features combined end-to-end learning for automated pathological grading of breast cancer on digital mammograms," *Computerized Medical Imaging and Graphics*, vol. 71, pp. 58–66, Jan. 2019, doi: 10.1016/j.compmedimag.2018.10.008.

[21] W. Kim *et al.*, "Development of novel breast cancer recurrence prediction model using support vector machine," *Journal of Breast Cancer*, vol. 15, no. 2, p. 230, 2012, doi: 10.4048/jbc.2012.15.2.230.

[22] C. Park, J. Ahn, H. Kim, and S. Park, "Integrative gene network construction to analyze cancer recurrence using semi-supervised learning," *PLoS ONE*, vol. 9, no. 1, p. e86309, Jan. 2014, doi: 10.1371/journal.pone.0086309.

[23] S. S, K. M, S. E, and R. C, "Automatic classification on bio medical prognosisof invasive breast cancer," *Asian Pac J Cancer Prev*, vol. 18, no. 9, pp. 2541–2544, 2017, doi: 10.22034/APJCP.2017.18.9.2541.

[24] A. F. M. Agarap, "On breast cancer detection," in *Proceedings of the 2nd International Conference on Machine Learning and Soft Computing - ICMLSC '18*, 2018, pp. 5–9, doi: 10.1145/3184066.3184080.

[25] R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin, "A curated mammography data set for use in computer-aided detection and diagnosis research," *Scientific Data*, vol. 4, no. 1, Dec. 2017, doi: 10.1038/sdata.2017.177.

BIOGRAPHIES OF AUTHORS

Khalid Shaikh is a serial entrepreneur, a technocrat, and a business strategist. He is the founder and CEO of Prognica Labs, an Artificial Intelligence powered health-tech company, IntelliDent AI–a Dental AI startup and Affaan Technologies, an AI Consulting and Development Company. He is a computer engineer and has a solid record of more than 20 years of high level performance in the healthcare and finance industries. In addition to his professional commitments, he also gives back to the aspiring entrepreneur community by serving as an advisor and mentor. He can be contacted at email: khalid@prognica.com.

Sabitha Krishnan received Masters in Engineering in Biomedical engineering and BE in Computer Science. She has 6+ years of R&D experience in AI/ML in healthcare. Her primary areas of research are artificial intelligence, computer vision, medical image analysis, biometrics, and digital forensic. She can be contacted at email: sabitha@prognica.com.

Dr. Rohit Thanki (Ph.D.) received his Ph.D. in electronics and communication engineering from C.U. Shah University, an ME in communication engineering from G.H. Patel College of Engineering and Technology, and a BE in electronics and communication engineering from Atmiya Institute of Technology and Science, India. His primary areas of research are artificial intelligence, computer vision, digital watermarking, medical image analysis, multimedia security, biometrics, and compressive sensing. He has published 11 books with Springer and one with CRC Press, and has contributed 13 book chapters in edited books published by Elsevier, Springer, CRC Press, and IGI Global. He has published over 20 research articles in refereed and indexed international journals, and is a reviewer for journals such as ACM Transactions on Multimedia Computing, Communications and Applications, IEEE Access, IEEE Consumer Electronics Magazine, IET Image Processing, IET Biometrics, Soft Computing, Imaging Science Journal, Signal Processing: Image Communication, and Computers & Electrical Engineering. He can be contacted at email: rohit@prognica.com.