ORIGINAL PAPER

Automated early breast cancer detection and classification system

Asmaa A. Hekal[1] · Ahmed Elnakib[1] · Hossam El-Din Moustafa[1]

Abstract
Early detection of breast cancer is clinically important to reduce the mortality rate. In this study, a new computer-aided detection (CAD) and classification system is introduced to classify two types of mammogram tumors (i.e., mass and calcification) as either benign or malignant. In this CAD system, the tumor-like regions (TLRs) are identified using the automated optimal Otsu thresholding method. Afterward, deep convolutional neural networks (CNNs) process the extracted TLRs to extract relevant mammogram features, investigating AlexNet and ResNet-50 architectures. The normalized extracted CNN features are further input to a support vector machine classifier to decode the classes of mammogram structures (i.e., Benign Calcification, Benign Mass, Malignant Calcification, and Malignant Mass nodules). The experimental results are tested on 2800 mammogram images from the Curated Breast Imaging Subset of Digital Database of Screening Mammography, a publicly available dataset. The accuracy of the proposed CAD system, to classify the ROI into one of the four classes, achieves 0.91 using AlexNet and 0.84 using ResNet-50 models, using fivefold cross-validation. Comparison results with the related methods confirm the advantages of the proposed CAD system.

Keywords Deep learning · Breast cancer · Convolutional neural networks (CNN) · Computer-aided detection system · Support vector machine

1 Introduction

According to the World Health Organization (WHO), breast cancer is the most globally common cancer in women [1]. Each year, around 2.1 million women are diagnosed with breast cancer, resulting in the highest number of women cancer-related deaths. A recent report in 2018 estimates around 627,000 women cancer-related deaths, accounting for around 15% of all women cancer deaths [1,2]. Low-energy X-rays mammography is considered the standard screening tool to identify the breast cancer abnormalities. Mammogram screening has shown to reduce breast cancer mortality by approximately 20% in high-resource settings [3]. Early detection is clinically important to reduce the cancer mortal-

ity rate [4,5]. However, the early detection of breast cancer on screening mammography is challenging due to the small sizes of potential nodules with respect to the entire breast [4]). Therefore, computer-aided detection (CAD) systems for breast cancer are clinically important to reduce the workload and improve the detection accuracy of radiologists and experts [5].

Recently, convolutional neural networks (CNNs) show impressive performance in the field of pattern recognition, classification, object detection [6,7], and more specifically in the field of breast cancer detection [8–16]. For example, Lévy and Jain [8] used three different network architectures (a CNN model, AlexNet, and GoogLeNet) to classify benign and malignant breast masses, using areas around the breast masses, cropped from the DDSM dataset, achieving accuracies of 60%, 89%, and 92.9%, on the three tested models, respectively. Castro et al. [10] trained a CNN architecture for mass classification into benign or malignant and validated it on three available full mammogram datasets (INbreast, CBIS-DDSM, and a Breast Cancer Digital Repository (BCRP)). The reported sensitivities on the tested datasets are 80%, 80%, and 60%, respectively. To classify breast masses as benign or malignant, Tsochatzidis et

✉ Ahmed Elnakib
   nakib@mans.edu.eg

   Asmaa A. Hekal
   asmaahekal3@gmail.com

   Hossam El-Din Moustafa
   hossammoustafa@ieee.org

[1] Electronics and Communication Engineering Department, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt
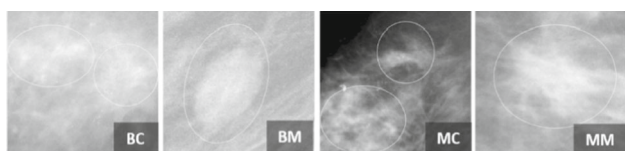
Fig. 1 Typical samples for the four classes of the ROI CBIS-DDSM database [17]. From left to right: Benign Calcification (BC), Benign Mass (BM), Malignant Calcification (MC), and Malignant Mass (MM) images. Nodules are outlined using white circles

al. [12] investigated different deep CNNs (AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-101, ResNet-152, GoogLeNet, Inception-BN-v2), showing that the fine-tuning of the pre-trained ResNet-101 has achieved the best accuracies on DDSM and Region of Interest (ROI) CBIS-DDSM databases of 78.5% and 75.3%, respectively. Chun-minga et al. [13] classified mammography images of a dataset, collected from DDSM and CBIS-DDSM databases, into five classes, i.e., benign calcification (BC), benign mass (BM), malignant calcification (MC), and malignant mass (MM) and normal. They fused two deep CNN networks, achieving an accuracy of 91%. Ragab et al. [14] used AlexNet and SVM classification to classify breast masses in as benign or malignant with an accuracy of 87.2% on preprocessed ROI CBIS-DDSM. Falconí et al. [15] applied a preprocessing step followed by feature extraction using ResNet50 and MobileNet models to classify benign and malignant ROI CBIS-DDSM breast masses, achieving an accuracy of 78.4% and 74.3%, for the two models, respectively. Ansar et al. [16] used the MobileNet model to classify breast masses into benign or malignant, achieving an accuracy of 74.5% on preprocessed ROI CBIS-DDSM data. Although different methods have achieved considerable success to detect and classify breast cancer, there is still a need to investigate how to further improve the classification accuracy. This paper focuses to develop an approach for the early detection of the breast cancer, using the ROI CBIS-DDSM [17] (see Fig. 1), to overcome the following limitations of the existing work:

– Building new CNN models require huge amount of data for training, to avoid overfitting. In addition, these methods are computationally expensive, especially in the training phase.
– Training with the whole details of the ROI image represents an overhead and may lead to reduce the detection accuracy.
– There is a need to design advanced methods and develop new ideas to improve the diagnosis accuracy, especially for early cases, where the nodule size is small with respect to the image size.

The main features/contributions of this work are as follows:

– Unlike the traditional CNN, which work on the original mammogram or cropped regions, the proposed CAD system applies an automated optimal dynamic thresholding method to extract the tumor-like regions (TLRs), eliminating the need for processing the whole ROI image for detecting breast nodules and enabling the CNN model to focus on the fine TLRs details, improving the chance of the early detection of the small size nodules.
– Performance is evaluated on the challenging standard CBIS-DDSM dataset [17], achieving superior performance over the competing methods.

2 Materials and methods

The proposed deep learning system, see Fig. 2, is designed based on three processing stages. In the first sage, the TLRs are identified based on an automated thresholding method. In the second stage, a transfer deep learning is performed by investigating two architectures, i.e., ResNet-50 and AlexNet. Finally, a classification stage is used to classify TLR images based on the normalized deep extracted features.

2.1 Collected database (CBIS-DDSM)

To test the proposed CAD system, the ROI CBIS-DDSM [17] images are utilized, composed of 3549 ROI mammogram images (1852 calcification (1132 BC and 720 MC) and 1697 mass (913 BM and 784 MM) images), see Fig. 1. CBIS-DDSM is an updated and standardized version of the Digital Database for Screening Mammography (DDSM) [18]. Image type has been converted from DICOM format to PNG format. Figure 1 shows typical samples of CBIS-DDSM database of the four classes (i.e., BC, BM, MC, and MM).

2.2 Extracting tumor-like regions (TLRs)

The first stage in the proposed CAD system aims at extracting TLRs to enable the CNN model to focus on the fine TLRs details, improving the chance of the early detection of the small size nodules and eliminating the need for processing the unnecessary details of the original CBIS-DDSM image. TLRs is a grey-level image, where the regions with similar grey levels as the tumors are extracted. To achieve this goal, a Gaussian smoothing is firstly applied to the original ROI image, IM_{ORG}, in order to enable the proposed Otsu method to focus on the global image features. Let the Gaussian kernel be G. The image is convolved with G [19,20]:

$$IM_{Smooth} = IM_{ORG} * G \qquad (1)$$

Fig. 2 Diagram of the proposed CAD system for breast cancer detection. First, tumor-like regions (TLRs) are identified from the input raw image by Otsu thresholding. Second, TLRs are processed using a CNN model to form a feature descriptor. Finally, classification is performed to produce the breast classes (BC, BM, MC, and MM)
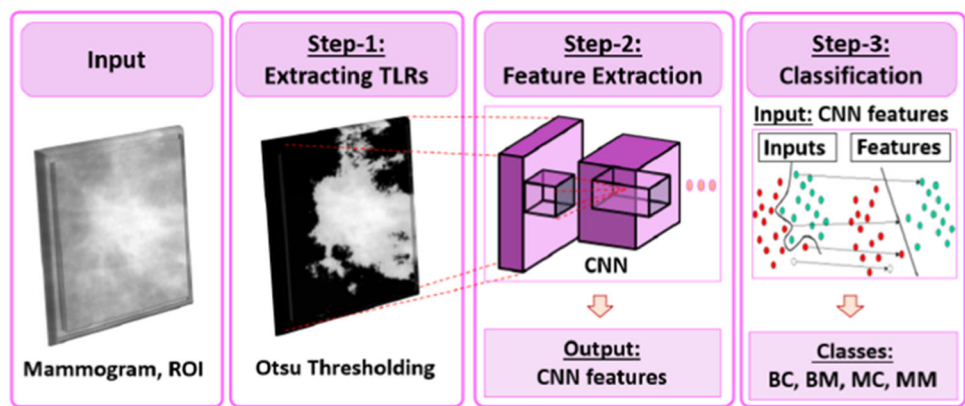

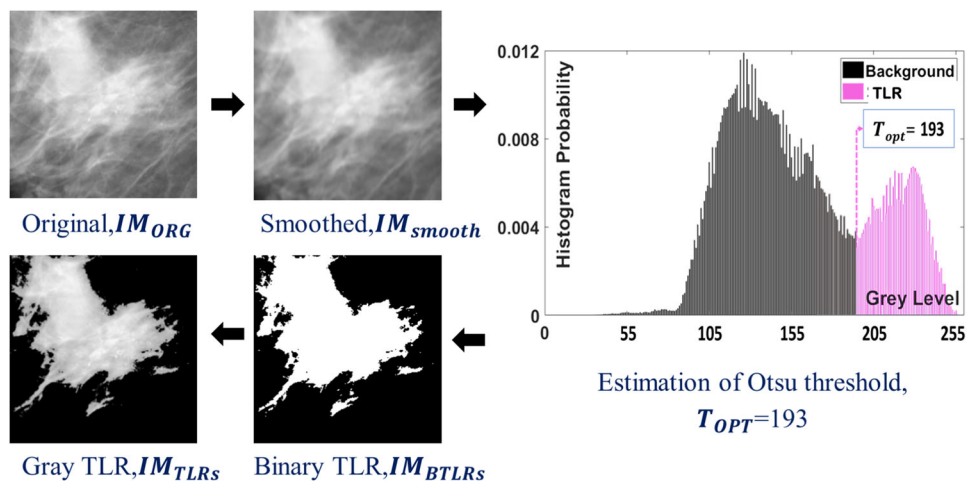**Fig. 3** Extracting tumor-like regions (TLRs) by processing the smoothed image using Otsu thresholding

Table 1 Summary of the pre-trained ResNet50 and Alexnet CNN models

CNN Model description				Pre-trained Network Parameters		
CNN model	I/P	#CL	#FC	Database	#Classes	Size
AlexNet	227 × 227	5	3	ILSVRC 2012	1000+	62.4 MB
ResNet-50	224 × 224	49	1	ILSVRC 2014	1000	102 MB

The symbol "#" in the table indicates the number, I/P indicates the input image size, FC denotes fully connected, CL denotes convolutional layers, and size indicates the CNN model size

where * denotes the convolution operator and IM_{Smooth} is the smoothed image. The Gaussian kernel G is defined as follows:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}} \qquad (2)$$

where (x, y) is the pixel's Cartesian coordinates and σ denotes the standard deviation of the Gaussian kernel. Afterward, the Otsu method, an optimum global thresholding method, proposed by Otsu [21], is used to extract TLRs, utilizing the zeros and the first-order cumulative moments of the gray-level histogram. Otsu thresholding is optimum in the sense that it maximizes the between class variance, a well-known measure used in statistical discriminant analysis [22,23]. Assume that the grey level histogram of the image,

IM_{Smooth}, is composed of L pins. Let $p(i)$ be the probability of an intensity grey level $i, i \in \{0, L-1\}$. Let T be an arbitrary random threshold. Let the weights ω_0 and ω_1 define the probabilities of the two classes, separated by the threshold T, as follows: $\omega_0(T) = \sum_{i=0}^{T-1} p(i)$ and $\omega_1(T) = \sum_{L-1}^{i=T} p(i)$. The means of the two classes are:

$$\mu_0(T) = \frac{\sum_{i=0}^{T-1} p(i)}{\omega_0} \quad \text{and} \quad \mu_1(T) = \frac{\sum_{L-1}^{i=T} p(i)}{\omega_1}$$

The between class variance, σ_b^2, can be expressed as follows [22,23]: $\sigma_b^2(T) = \omega_0\omega_1(\mu_0 - \mu_1)^2$. To estimate the TLRs, an algorithm of six steps is followed. First, set initial random values for the class probabilities and means, i.e., $\omega_0, \omega_1, \mu_0$, and μ_1. Second, for all possible thresholds $T = 1, 2, ..., L-1$: Update ω_0, ω_1, μ_0, and μ_1 accordingly.
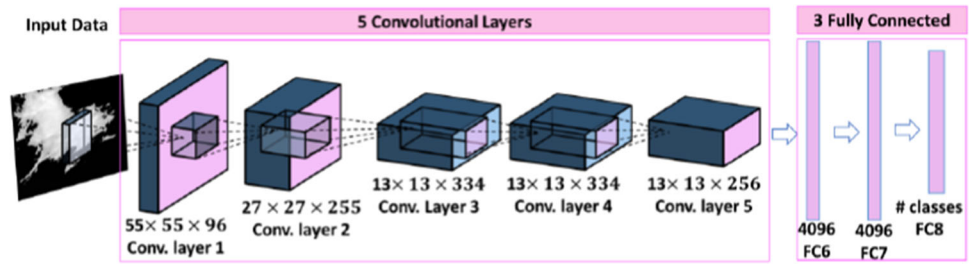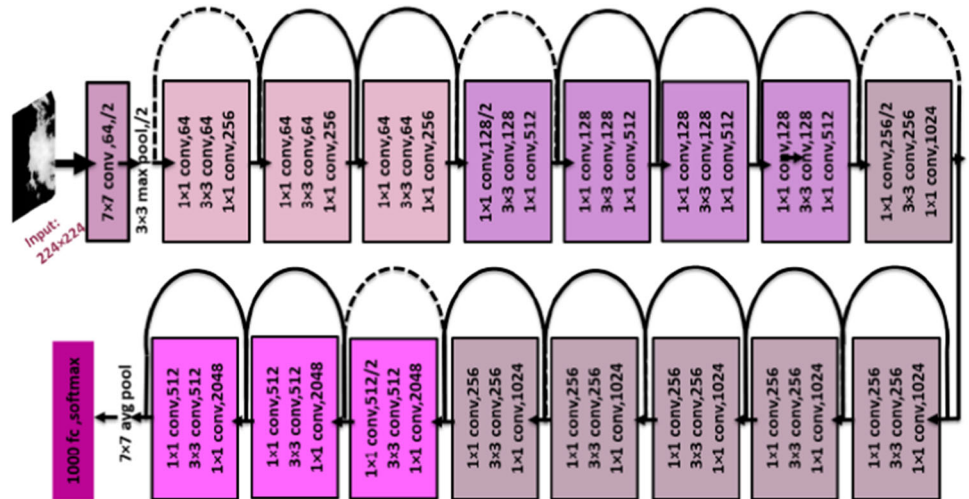
Fig. 4 AlexNet [27] architecture

Fig. 5 ResNet-50 [28] architecture

Third, compute $\sigma_b^2(T)$ for each threshold, T. Forth, determine the desired optimal global threshold, T_{OPT}, that globally maximizes the between class variance to correspond to the maximum $\sigma_b^2(T)$ for all T values. Fifth, compute IM_{BTLRs}, where all pixels less than T_{OPT} are classified as background (binary zero), and the remaining pixels are TLRs (binary one), (see Fig. 3). Finally, Output the grey-level TLRs image IM_{TLRs}, by the pixel-by-pixel multiplication of IM_{BTLRs} with the original image IM_{ORG} as follows: $IM_{TLRs}=IM_{ORG} * IM_{BTLRs}$. Figure 3 illustrates the step-by-step algorithm of the proposed method for extracting TLRs. Note that the TLRs are not the tumor class, but it contains the regions of the grey levels that are similar to the tumor grey levels (bright pixels).

2.3 Utilized CNN models

A convolutional neural network (CNN) consists of two types of layers; convolutional layers to extract lower- and higher-order image features, and fully connected layers to perform the classification [24–26]. AlexNet [27] and ResNet [28] CNN models are from the most commonly used popular architectures, that achieve a remarkable success in medical applications (i.e., lung cancer detection [29] and face recognition [30]), and in particular breast cancer detection [8,11,12,14]. Therefore, they have been adopted in the proposed system.

AlexNet has five convolution layers, three pooling layers, and three fully connected layers [27,31] (see Table 1). The AlexNet CNN architecture is shown in Fig. 4 The layers of Conv. Layer 1 to Conv. Layer 5, in Fig. 4, are the convolution layers. Each neuron in the convolution layers computes a dot product between their weights and the local region that is connected to the input volume [32]. Each convolutional layer is followed by a pooling layer in order to perform a down sampling operation along the spatial dimensions to reduce the amount of the computations and improve the robustness [32]. Additionally, the fully connected layers are FC6, FC7 and FC8, as shown in Fig. 4. The role of the fully connected layers are to process the extracted convolutional layers' features and output the most relevant feature vector for classification [32].

ResNet is a short name for Residual Network. The basic idea of a ResNet model is to skip blocks of convolutional layers by using shortcut connections (see Fig. 5) [12]. ResNet-50 is based on a residual learning framework, where layers within a network are reformulated to learn a residual mapping rather than the desired unknown mapping between the inputs and outputs. Such a network is easier to optimize and consequently enables training of deeper networks, which correspondingly leads to an overall improvement in the network capacity and performance [12].

In order to reduce the training overhead, the weights of the convolutional layers of the pretrained models are transferred without training, and only the fully connected layers are

trained using the CBIS-DDSM data. To apply transfer learning of the CNN models, the last fully connected layer of the pre-trained models (FC8 layer in AlexNet or FC1000 layer in ResNet50) is replaced with a shallow classifier (SVM). The vectors of activities of the FC7 layer in AlexNet or the flatten layer (just before FC1000) in Res-Net50, represent the feature descriptor of the input CBIS-DDSM image. Features are further normalized between 0 and 1 to the input of the SVM classifier.

2.4 Classification

To classify the mammogram images, a SVM shallow classifier is used. SVM is a supervised machine learning algorithm that sorts data in categories. The idea of SVM is to formulate a computationally efficient way of learning by separating hyper planes in a high dimensional feature space [33,34]. The input of the classifier is the vectors of activities (FC7 in AlexNet or flatten layer in ResNet-50) and the output is the ROI class. The SVM classifier is chosen using binary non-linear kernel to reflect the variability of the classes. For the case of four classes, We used one-versus-all SVM, in which we fuse four binary trained classifiers, each processes a class, where the data from the class is treated as positive, and the data from all the other classes is treated as negative. The performance is evaluated for either the ResNet-50 or AlexNet, each along with the SVM classifier based on a standard performance evaluation metric, i.e., the classification accuracy, in order to select the best model/classifier.

2.5 Performance evaluation

To test the proposed system, we used six standard metrics to evaluate the system performance, i.e., Accuracy (ACC), Sensitivity (SEN), Specificity (SPE), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F1 score (FSC) defined as follows [35]:

$$ACC = \frac{\text{No. of correct assessments}}{\text{No. of assessments}} \tag{1}$$

$$SEN = \frac{\text{No. of true positive assessments}}{\text{No. of positive assessments}} \tag{2}$$

$$SPE = \frac{\text{No. of true negative assessments}}{\text{No. of negative assessments}} \tag{3}$$

$$PPV = \frac{\text{No. of true positive assessments}}{\text{No. of positive assessments}} \tag{4}$$

$$PNV = \frac{\text{No. of true negative assessments}}{\text{No. of negative assessments}} \tag{5}$$

Fig. 6 Comparison results with the ground truth (GT) and other threshold (T)-based methods

$$FSC = 2 * \frac{\text{PPV} * \text{SEN}}{\text{PPV} + \text{SEN}} \tag{6}$$

3 Experimental results and discussions

This section illustrates, in details, the experimental setup, comparison results, and related discussions.

3.1 Experimentation setting

To evaluate the proposed system, CBIS-DDSM ROI dataset is used. To train the classifier, data is randomly divided into training set (70%) and testing set (30%). To train the deep learning model (AlexNet and ResNet-50), the Bayesian optimizer is used to minimize the binary cross entropy function, using a learning rate of $10^{(-4)}$. During training, the data is shuffled using a mini-patch size of 128. The maximum number of epochs is set to 20.

3.1.1 Threshold cases

To verify the advantages of using the automated adaptive Otsu thresholding to identify TLRs, different values of fixed thresholds are investigated and the performance results of the proposed system have been carried and compared with values of fixed threshold cases. Assuming that the ROI grey levels are normalized between 0 and 1, Case I uses original ROI dataset. Case II to case VII use ROI with thresholds T of 0.60 to 0.80 with a step of 0.05.

Table 2 Accuracy (ACC), Sensitivity (SEN), Specificity (SPE), Positive Predictive Value (PPV), Negative Predictive Value (NPV), and F1 score (FSC) between the different reported cases, for classification breast nodules into benign mass and malignant mass (BM–MM), or into benign calcification and malignant calcification (BC–MC), or into benign calcification, benign mass, malignant calcification, and malignant mass (BC–BM–MC–MM) using the two investigated modals (i.e., ResNet-50, and AlexNet) and the SVM classifier

Case	Description	BM - MM						BC - MC						BC-BM-MC-MM	
		ACC	SEN	SPE	PPV	NPV	FSC	ACC	SEN	SPE	PPV	NPV	FSC	ACC	FSC
*Using imbalanced data (the whole 3549 images), 70% train-30% test															
I	ResNet-50	0.66	0.65	0.68	0.66	0.66	0.66	0.70	0.78	0.65	0.76	0.64	0.68	0.63	
(Original Data)	AlexNet	0.73	0.69	0.75	0.75	0.74	0.75	0.74	0.73	0.74	0.74	0.74	0.73	0.70	
II	ResNet-50	0.85	0.84	0.94	0.85	0.95	0.85	0.84	0.86	0.87	0.83	0.87	0.84	0.75	
(Threshold = 0.60)	AlexNet	0.80	0.76	0.73	0.87	0.73	0.81	0.78	0.77	0.76	0.80	0.76	0.79	0.75	
III	ResNet-50	0.87	0.86	0.85	0.89	0.86	0.87	0.82	0.84	0.84	0.80	0.84	0.82	0.74	
(Threshold = 0.65)	AlexNet	0.82	0.82	0.82	0.82	0.82	0.82	0.79	0.82	0.84	0.73	0.84	0.78	0.79	
IV	ResNet-50	0.90	0.91	0.91	0.90	0.88	0.90	0.86	0.84	0.83	0.90	0.83	0.87	0.80	
(Threshold = 0.70)	AlexNet	0.92	0.92	0.92	0.91	0.92	0.92	0.84	0.84	0.94	0.84	0.94	0.84	0.88	
V	ResNet-50	0.90	0.91	0.91	0.90	0.91	0.90	0.86	0.85	0.86	0.86	0.86	0.85	0.78	
(Threshold = 0.75)	AlexNet	0.92	0.89	0.88	0.97	0.88	0.93	0.85	0.86	0.86	0.84	0.86	0.85	0.85	
VI	ResNet-50	0.89	0.89	0.96	0.89	0.96	0.89	0.86	0.86	0.86	0.86	0.86	0.86	0.79	
(Threshold = 0.80)	AlexNet	0.89	0.88	0.92	0.88	0.92	0.88	0.81	0.81	0.81	0.80	0.81	0.81	0.86	
Proposed Method*	ResNet-50	0.92	0.92	0.92	0.91	0.92	0.92	0.88	0.90	0.90	0.90	0.90	0.89	0.81	
	AlexNet	0.95	0.92	0.96	0.91	0.96	0.91	0.86	0.80	0.90	0.80	0.90	0.80	0.90	
**Using balanced data (2800 image, 700 image per each class), 70% train-30% test															
Proposed Method**	ResNet-50	0.97	0.97	0.97	0.97	0.97	0.97	0.91	0.88	0.86	0.95	0.86	0.91	0.89	
	AlexNet	0.97	0.97	0.97	0.98	0.97	0.97	0.90	0.93	0.88	0.86	0.86	0.90	0.93	

Fig. 7 Confusion matrix classification using the proposed Otsu's method and the AlexNet using the balanced data

3.1.2 Data balancing

The experiments are run through two setting (*a*) imbalance data that contains the whole data (3549 images of imbalanced number of images per each class), and (*b*) balanced data which contains 2800 images (700 image per each class) The proposed system adopted balancing the data to avoid the bias toward the classes of large numbers.

3.2 Comparison results

3.2.1 Experiments using the whole imbalanced data

Two investigated deep models (i.e., AlexNet and ResNet-50) have been compared, using the different values of thresholds, for breast nodule diagnosis, i.e., classification of breast masses into benign and malignant (BM or MM), and classification of breast calcification into benign and malignant (BC or MC). Table 2 overviews the achieved comparison results. Otsu's thresholding, besides being automated and adaptive to each input TLR, achieves the best performance for both the investigated ResNet-50, and AlexNet models, as shown in Table 2. In addition, Table 2 demonstrates that when classification breast mass the AlexNet model achieves slightly better results than the ReseNet-50 model. This may be due to the large number of layers of the Resenet50 (50 layers) compared to the AlexNet model (five convolutional and two fully connected layers), that lead to data overfitting in the case of ResNet-50 model, taking into account limited size of the standard ROI CBIS-DDSM data. Figure 6 shows visual comparison results of the proposed method using the whole data. As shown in Fig. 6, the proposed method shows superior results over the competing fixed thresholding methods. In addition to the ability of classifying breast calcification into benign or malignant (BC or MC), see Table 2, the proposed system is tested to detect the malignancy in case of calcification nodules, which is more challenging, since the nodule shape in case of the calcification nodules is complex and distributed. In this experiments, ResNet-50 achieves slightly better results, due to the complex task of the nodule diagnosis, in case of the calcification nodules.

Furthermore, the proposed system is tested to classify the ROI data into one of the four classes, i.e., BC, MC, BM, or MM. As shown in Table 2, the proposed system achieves accuracies of 0.81 and 0.90 for using ResNet-50 and AlexNet models, respectively, in case of using the whole imbalanced data. As expected, the accuracy of the 4-class classifier is significantly below that of a 2-class classifier. This is due to the added challenge of differentiating between benign and malignant, in additional to distinguishing between mass and calcification.

3.2.2 Experiments using the balanced data

To remove the bias toward the classes of large number, we adopted balancing the data, such that each class contains 700 images (a total of 2800 images selected randomly). As expected, using balanced data further improve the performance (see Table 2). Figure 7 shows the confusion matrix for classifying the four classes using Alexnet and balanced data using our proposed method. As shown in the figure, the total accuracy of all classes reaches 93.2%, which outperforms the case of using imbalanced data, as expected. In order to investigate the sensitivity of the proposed method to the selected training data, fivefold cross-validation is applied, achieving relatively low standard deviation, which indicates that our proposed method is insensitive to the training data selection (see Table 3).

Table 3 shows that the proposed system achieves a better results to the most similar related work [13]. Note that our experiments do not consider the normal (N) class, since the utilized standard ROI CBIS-DDSM does not contain any normal record. To demonstrate the advantage of the proposed system for breast cancer identification, Table 3 compares the results of the proposed system with the related state-of-the-art methods. Table 3 shows the superior of the proposed system for classifying breast masses into benign and malignant. This is due to that the proposed automated thresholding improves the detection rate, since the deep model focuses to detect tumors within only the TLR regions.

Table 3 Comparison results of Accuracy between related methods and our proposed system Accuracy for case V

Method	Database	Description	Classes	ACC
Lvy et al. [8]	Cropped DDSM	GoogLeNet	BM,MM	0.93
Castro et al. [10]	Full mammogram CBIS-DDSM	CNN model	BM,MM	0.85
Tsochatzidis et al. [12]	ROI CBIS-DDSM	ResNet-101	BM,MM	0.75
Chun-minga et al. [13]	Cropped DDSM+ cropped CBIS-DDSM	Two fused CNN networks	BC,BM,MC,MM,N	0.91
Ragab et al. [14]	ROI CBIS DDSM	Preprocessing+ AlexNet + SVM	BM,MM	0.87
Falconí et al. [15]	ROI CBIS-DDSM	Preprocessing+ ResNet-50	BM,MM	0.78
Ansar et al. [16]	ROI CBIS-DDSM	Preprocessing+ MobileNet	BM,MM	0.75
Proposed Method (5-fold cross-validation, balanced data)	ROI CBIS-DDSM	TLRs+ResNet50+SVM	BM,MM	0.94 ± 1.8
			BC,MC	0.89 ± 2.9
			BM,MM,BC,MC	0.84 ± 2.1
		TLRs+AlexNet+SVM	BM,MM	0.95 ± 2.5
			BC,MC	0.87 ± 1.7
			BM,MM,BC,MC	0.91 ± 1.8

In CBIS-DDSM ROI dataset

3.3 Limitations

Although the proposed method has achieved promising results comparing with the competing methods, it fails to correctly classify all the images. Last column in Fig. 6 shows an example of incorrect classification. In the future, we will try adding more features and/or different CNN models to improve further the classification accuracy.

4 Conclusion

In this work, a CAD system for early detect breast cancer based on deep learning is proposed. Unlike related work, the utilized CNN models extract features from the TLRs based on automated adaptive Otsu thresholding, in order to improve the training capabilities of the deep learning model. A SVM classifier is used to classify mammogram images into the nodule classes, i.e., BC, MC, BM, and MM. Experiments results on the ROI CBIS-DDSM data confirm the superior of the proposed method over other related works. In the future, other databases will be investigated to test the robustness of the proposed system. In addition, other feature models will be tested in order to improve the performance.

References

1. WHO. World health organization. https://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/ (2020). Accessed 7 Nov 2020
2. DeSantis, C.E., Miller, K.D., GodingSauer, A., Jemal, A., Siegel, R.L.: Cancer statistics for African Americans. CA Cancer J. Clin. **69**(3), 211 (2019)
3. Chen, H.L., Yang, B., Wang, G., Wang, S.J., Liu, J., Liu, D.Y.: Support vector machine based diagnostic system for breast cancer using swarm intelligence. J. Med. Syst. **36**(4), 2505 (2012)
4. Shen, L., Margolies, L.R., Rothstein, J.H., Fluder, E., McBride, R., Sieh, W.: Deep learning to improve breast cancer detection on screening mammography. Sci. Rep. **9**(1), 1 (2019)
5. Hamidinekoo, A., Denton, E., Rampun, A., Honnor, K., Zwiggelaar, R.: Deep learning to improve breast cancer detection on screening mammography. Med. Image Anal. **47**, 45 (2018)
6. Chan, H.P., Samala, R.K., Hadjiiski, L.M., Zhou, C.: Deep Learning in Mammography and Breast Histology, An Overview and Future Trends, Medical Image Analysis, pp. 3–21. Springer, Berlin (2020)
7. Altaf, F., Islam, S.M., Akhtar, N., Janjua, N.K.: Going deep in medical image analysis: concepts, methods, challenges, and future directions. IEEE Access **7**, 99540 (2019)
8. Lévy, D., Jain, A.: Breast mass classification from mammograms using deep convolutional neural networks, arXiv preprint arXiv:1612.00542 (2016)
9. Yi, D., Sawyer, R.L., CohnIII, D., Dunnmon, J., Lam, C., Xiao, X., Rubin, D.: Optimizing and visualizing deep learning for benign/malignant classification in breast tumors, arXiv preprint arXiv:1705.06362 (2017)
10. Castro, E., Cardoso, J.S., Pereira, J.C.: Elastic deformations for data augmentation in breast cancer mass detection. In: 2018 IEEE

EMBS International Conference on Biomedical and Health Informatics (BHI), pp. 230–234. IEEE (2018)

11. Xi, P., Shu, C., Goubran, R.: Abnormality detection in mammography using deep convolutional neural networks. In: 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA), pp. 1–6. IEEE (2018)

12. Tsochatzidis, L., Costaridou, L., Pratikakis, I.: Deep learning for breast cancer diagnosis from mammograms—a comparative study. J. Imaging **5**(3), 37 (2019)

13. Chun-ming, T., Xiao-mei, C., Xiang, Y., Fan, Y., et al.: Five classification of mammography images based on deep cooperation convolutional neural network. Am. Sci. Res. J. Eng. Technol. Sci. ASRJETS **57**(1), 10 (2019)

14. Ragab, D.A., Sharkas, M., Marshall, S., Ren, J.: Breast cancer detection using deep convolutional neural networks and support vector machines. PeerJ **7**, e6201 (2019)

15. Falconí, L.G., Pérez, M., Aguilar, W.G.: Transfer learning in breast mammogram abnormalities classification with mobilenet and nasnet. In: 2019 International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 109–114. IEEE (2019)

16. Ansar, W., Shahid, A.R., Raza, B., Dar, A.H.: Breast cancer detection and localization using MobileNet based transfer learning for mammograms. In: International Symposium on Intelligent Computing Systems, pp. 11–21. Springer (2020)

17. Lee, R.S., Gimenez, F., Hoogi, A., Miyake, K.K., Gorovoy, M., Rubin, D.L.: A curated mammography data set for use in computer-aided detection and diagnosis research. Sci. Data **4**(1), 1 (2017)

18. Bowyer, K., Kopans, D., Kegelmeyer, W., Moore, R., Sallam, M., Chang, K., Woods, K.: The digital database for screening mammography. In: Third International Workshop on Digital Mammography, vol. 58, p. 27 (1996)

19. Al-Najdawi, N., Biltawi, M., Tedmori, S.: Mammogram image visual enhancement, mass segmentation and classification. Appl. Soft Comput. **35**, 175 (2015)

20. Lukac, R.: Single-Sensor Imaging: Single-Sensor Imaging: Methods and Applications for Digital Cameras. CRC Press, Boca Raton (2018)

21. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62 (1979)

22. Deepa, S., Subbiah, B.V.: Efficient ROI segmentation of digital mammogram images using Otsu's N thresholding method. Natl. J. Adv. Comput. Manag. **4**, 1 (2013)

23. Swetha, T., Bindu, C.H.: Detection of breast cancer with hybrid image segmentation and Otsu's thresholding. In: 2015 International Conference on Computing and Network Communications (CoCoNet), pp. 565–570. IEEE (2015)

24. LeCun, Y., Kavukcuoglu, K., Farabet, C.: Convolutional networks and applications in vision. In: Proceedings of 2010 IEEE International Symposium on Circuits and Systems, pp. 253–256. IEEE (2010)

25. Anagnostis, A., Asiminari, G., Papageorgiou, E., Bochtis, D.: A convolutional neural networks based method for anthracnose infected walnut tree leaves identification. Appl. Sci. **10**(2), 469 (2020)

26. Truong, T.D., Pham, H.T.T.: Breast cancer histopathological image classification utilizing convolutional neural network. In: International Conference on the Development of Biomedical Engineering in Vietnam, pp. 531–536. Springer (2018)

27. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. **25**, 1097 (2012)

28. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition (2015), arXiv preprint arXiv:1512.03385 (2016)

29. Elnakib, A., Amer, H.M., Abou-Chadi, F.E.: Early lung cancer detection using deep learning optimization. Int. J. Online Biomed. Eng. iJOE **16**(06), 82 (2020)

30. Moustafa, A.A., Elnakib, A., Areed, N.F.: Age-invariant face recognition based on deep features analysis. Signal Image Video Process. **14**, 1027 (2020)

31. Wu, Z., Shen, C., Van DenHengel, A.: Wider or deeper: revisiting the resnet model for visual recognition. Pattern Recogn. **90**, 119 (2019)

32. Suzuki, S., Zhang, X., Homma, N., Ichiji, K., Sugita, N., Kawasumi, Y., Ishibashi, T., Yoshizawa, M.: Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis. In: 2016 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), pp. 1382–1386 IEEE (2016)

33. Gunn, S.R., et al.: Support vector machines for classification and regression. ISIS Tech. Rep. **14**(1), 5 (1998)

34. Pisner, D.A., Schnyer, D.M.: Support Vector Machine, Machine Learning, pp. 101–121. Elsevier, Amsterdam (2020)

35. Zhu, W., Zeng, N., Wang, N., et al.: Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. NESUG Proc. Health Care Life Sci. Baltim. Md. **19**, 67 (2010)