Alexandria Engineering Journal (2021) 60, 4701–4709

HOSTED BY

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

Deep learning in mammography images segmentation and classification: Automated CNN approach

Wessam M. Salama [a,*], Moustafa H. Aly [b]

[a] Department of Basic Science, Faculty of Engineering, Pharos University, Alexandria, Egypt
[b] Department of Electronics and Communications Engineering, College of Engineering and Technology, Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt

Received 24 February 2021; revised 14 March 2021; accepted 23 March 2021

KEYWORDS

Mammography;
Breast cancer;
Segmentation;
Deep learning;
U-Net;
Transfer learning

Abstract In this work, a new framework for breast cancer image segmentation and classification is proposed. Different models including InceptionV3, DenseNet121, ResNet50, VGG16 and MobileNetV2 models, are applied to classify Mammographic Image Analysis Society (MIAS), Digital Database for Screening Mammography (DDSM) and the Curated Breast Imaging Subset of DDSM (CBIS-DDSM) into benign and malignant. Moreover, the trained modified U-Net model is utilized to segment breast area from the mammogram images. This method will aid as a radiologist's assistant in early detection and improve the efficiency of our system. The Cranio Caudal (CC) vision and Mediolateral Oblique (MLO) view are widely used for the identification and diagnosis of breast cancer. The accuracy of breast cancer diagnosis will be improved as the number of views is increased. Our proposed frame work is based on MLO view and CC view to enhance the system performance. In addition, the lack of tagged data is a big challenge. Transfer learning and data augmentation are applied to overcome this problem. Three mammographic datasets; MIAS, DDSM and CBIS-DDSM, are utilized in our evaluation. End-to-end fully convolutional neural networks (CNNs) are introduced in this paper. The proposed technique of applying data augmentation with modified U-Net model and InceptionV3 achieves the best result, specifically with the DDSM dataset. This achieves 98.87% accuracy, 98.88% area under the curve (AUC), 98.98% sensitivity, 98.79% precision, 97.99% F1 score, and a computational time of 1.2134 s on DDSM datasets.

1. Introduction

The common type of cancer among woman is the breast cancer accounting for about one third of cancer [1,2], and has a high mortality rates of about 17% [3]. Breast cancer plays an important role in increasing morality rate. Mammography screening

* Corresponding author at: Department of Electronics and Communications, College of Engineering and Technology, Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt.
E-mail addresses: wessam.salama@pua.edu.eg (W.M. Salama), mosaly@aast.edu (M.H. Aly).

Peer review under responsibility of Faculty of Engineering, Alexandria University.

https://doi.org/10.1016/j.aej.2021.03.048
1110-0168

technique [4] plays an important role in breast cancer diagnosis. Machine learning is outperforming the conventional handcrafted technique. It helps in selecting the most important features. Deep learning [5] plays an important role in enhancing the results in the field of biomedical engineering, in particular the Deep Convolutional Neural Networks (CNNs), which can be easily applied with state-of-the-art efficiency. A. Shrestha et al. developed a new algorithm to describe deep learning [6]. The models included InceptionV3, DenseNet121, ResNet50, VGG16 and MobileNetV2 models, for the classification process [7–9].

Scanning Mammograms of 28, 27, 342 women were studied retrospectively to assess the effect of optical density and the number of views on cancer detection [10]. 76% of the predicted invasive cancers were observed with an MLO view and an optical density of less than 1.4. It increased to 95% when the optical density was greater than 1.4 with MLO and CC views. A retrospective analysis of 83 histologically proven breast cancers using a paired *t*-test for cancer diagnosis was performed in the three mammographic views (CC, MLO) and Medio Lateral (ML) and in combinations of views. It was found that the exposure to mass description was substantially improved when a double-view method was used.

The CNN is a deep architecture that is used in image processing which compromises two main layers; the convolutionary layer and the pooling layer. The output of the neurons that are connected to the local area network at the input by sharing weights and biases is calculated by the convolutional layer. The pooling layer subsamples the output of the convolutional layer and reduces the data size. Learning millions of parameters of a deep CNN includes a vast number of training images as well as the availability of its ground truth, which actually prohibits many superior deep CNNs from being applied to medical applications.

Different CNNs for the mass detection task are explained by J. Arevalo et al. [11]. Their experimentation was performed on the Breast Cancer Digital Repository Film Mammography (BCDR-FM) dataset. D. Abdelhafizet al. [12] introduced a framework for a pre-trained CNN on DDSM database. Breast cancer classification algorithm from scratch is implemented by L. Tsochatzidis et al. [13] and improves the ability to segment normal and irregular breast tissue based on the application of deep-learning medical imaging technology.

The U-Net model is the main model in images segmentation. O. Ronneberger et al. suggested a U-Net model for biomedical images segmentation [14]. N. Alamet al. [15] developed an automated segmentation approach for biomedical images. In their technique, the region of interest (ROI) was manually extracted, and the wavelet-based process was performed to improve the frequency of spatial picture. The contour of the calcify area was extracted based on segmentation technique which was explained in the work of S. Duraisamy et al. [16]. Finally, the modified segmentation of the U-Net model was applied to extract the ROI (breast region) and remove the unwanted regions, in some papers in the classification of breast cancer [17–21].

Data augmentation technique creates new samples of the training data set by performing random transitions process to available datasets [22]. This has a lot effects, including speeding up the process of convergence and avoiding overfitting. The easiest approach to small datasets is to perform fine transitions (translation, zooming, flipping, mirroring, rotation, etc.).

Utilizing a pre-trained model instead of using model from scratch is called transfer learning. Training the neural network from scratch needs substantial data and computational power [23].

The main contribution in this paper is divided into two phases; classification phase based on different deep models, and segmentation phase before classification. First, different models including InceptionV3, DenseNet121, ResNet50, VGG16 and MobileNetV2 are employed to classify our MIAS, DDSM and CBIS-DDSM images into benign or malignant. Second, the segmentation phase based on applying modified segmentation of the U-Net model is utilized to extract ROI (breast region) and remove unwanted regions. This step plays a vital role in enhancing the images to be suitable as input to the classification phase and improving the system performance. After segmentation phase, our different deep learning models are applied to the segmented images to classify them into benign or malignant. The data augmentation technique is employed on DDSM and MIAS datasets to resolve the scarcity of datasets. Also, transfer leaning is used to minimize consuming time and computing resources.

The novelty in our work is based on exploring the most powerful models that are the most successful end-to-end deep models in computer vision, including InceptionV3, DenseNet121, ResNet50, VGG16 and MobileNetV2. These models are explored on three different mammography datasets which explain the distinct cases of breast thickness, breast size, and age of the patients taken from the MIAS, DDSM and CBIS-DDSM databases.

The rest of this paper is organized as follows. Details of the methodology used in this paper for the segmentation and classification of mammograms are explained in Section 2. Section 3 introduces and discusses the obtained results. Section 4 represents the main conclusions.

2. Methodology

This paper introduces novel strategies to segment and classify mammography. Pre-trained modified U-Net model and different deep learning models are utilized. This includes versions of InceptionV3, DenseNet121, ResNet50, VGG16 and MobileNetV2, as the beginning of scratch training leads to overfitting, time consuming and a need for high computing power. Data augmentation and fine-tuning are often used to resolve the scarcity of mammography images. Figs. 1 and 2 illustrate our proposed frame work (Data augmentation + modified U-Net model + Classifier Networks).

2.1. Transfer learning

Transfer learning is the golden key for using small datasets, e.g. medical images, which are impossible to collect in vast quantities than most datasets. A great deal of data, power and time is required to train deep learning models from scratch. So, pre-trained models and only fine tuning are used to solve these problems. For transfer learning, we apply the pre-trained InceptionV3, DenseNet121, ResNet50, VGG16 and MobileNetV2 models [7–10].

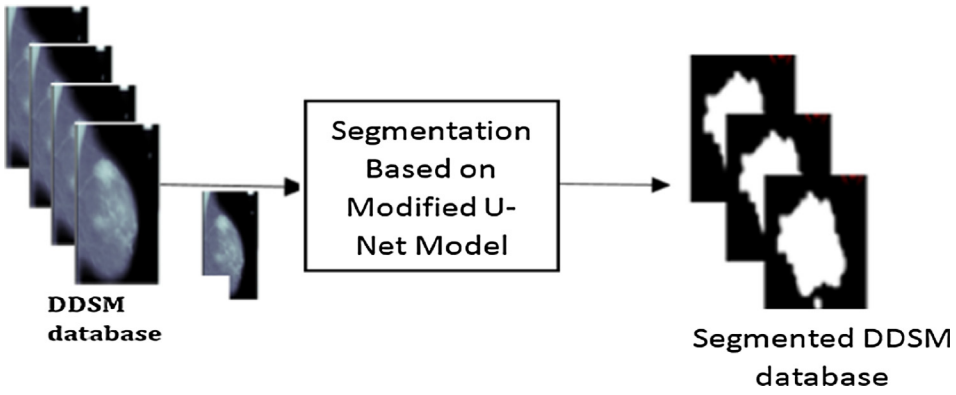

Fig.1 Our proposed segmentation framework.

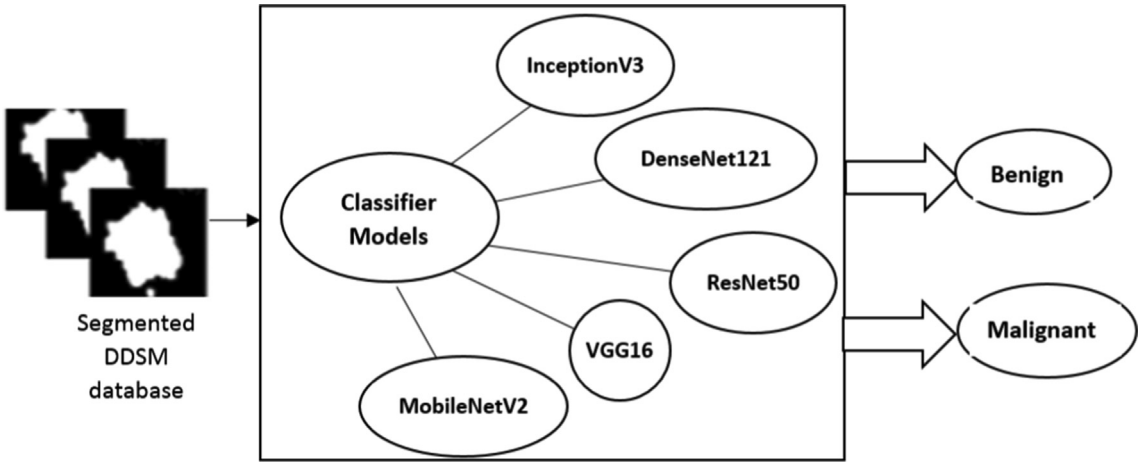

Fig.2 Our proposed classification framework.

2.2. Data augmentation effect

To achieve the best performance, we need a vast number of training samples. Therefore, data augmentation technique is performed to increase the number of original input data by creating new samples of training data. In this paper, the rotation is applied where each original image rotates 0, 90, 180 and 270 degrees. Therefore, each image is augmented into four images.

2.3. Segmentation network

The segmentation of the ROI is a crucial step in the automated analysis of mammography images for early detection of breast cancer. The segmentation could be considered as a classification task to classify each pixel in the dataset images, either ROI (breast region) nor background (BG) [24]. The input of this phase is the mammography images and the output is the ROI (breast regions) images. Then, the ROI images are masked with the original mammography images to prepare them as an input to the classifier phase. Our mammography images are segmented based on the modified U-Net model [25]. The modified segmentation of U-Net is compromised of the encoder and decoder network. The traditional CNN, which contains semantic information and less spatial information is called encoder part. However, spatial information is also important for segmentation of semantic information. The particular information from the decoder part is fed into U-Net, where semantic information is extracted from the lower most layer of the U-Net network. The decoder part contains the high-resolution features, these features are extracted from the encoder portion, skipping the link and providing the fine segment structures. The leaky ReLU and normalization of the instance are utilized instead of rectified linear unit (ReLU) and standardization of batch is, respectively, 32×32 pixels for the size of the patch and 42 feature maps that are used for the highest layer.

2.4. Deep learning models

InceptionV3, DenseNet121, ResNet50, VGG16 and MobileNetV2 models are considered the most successful deep CNN for classification in the computer vision field, to classify mammography. In order to begin the fine-tuning process on the mammography dataset, certain parameters must then be changed. For example, 1000 classes are replaced with the two classification classes, benign and malignant, layer.

2.4.1. InceptionV3

The ImageNet is utilized to train the InceptionV3 model. The advantage of this design is to increase the computing ability of the network. The number of iterations and the rate are adjusted to 10^6 and 10^{-4}, respectively, and the number of the epoch is adjusted to be 60 in our proposed work.

2.4.2. DenseNet121

Dense CNN (DenseNet121) is a fully connected neural feed forward network. DenseNet121 compromises a function map for every layer. The next layer is fed from the character map of previous layer. Among the benefits of DenseNet121 is that it needs fewer parameters. In addition, the number of iterations and the rate are adjusted to 10^5 and 10^{-2}. The number of the epoch is adjusted to 80.

2.4.3. ResNet50

The ResNet50, also known as the deep residual network, is one of the pre-trained models that use the ImageNet. The ResNet50 skips one or more layers and manages the gradient vanishing problem. One of the main benefits of ResNet50 is its ease of optimization. In addition, the accuracy of the sample can be improved by increasing the depth of the model. The two or three layers of the ResNet50 model are directly attached to either layer, not even the neighboring layer, using the ReLu nonlinear activation function. The ResNet50 uses forward and backward propagation methods. The number of iterations and the rate are set to 10^4 and 10^{-3}, respectively, and the number of the epoch is 130.

2.4.4. Vgg16

The AlexNet is updated to VGG16 with a larger number of layers. The generalization of the model is directly proportional to the number of layers. The benefit of VGG16 is the use of only 3×3 convolutional filters. The number of iterations and the rate should be 10^5 and 10^{-4}, respectively, to re-train the VGG16 model, and the number of the epoch is 80.

2.4.5. MobileNetV2

The MobileNetV2 consists of two types of blocks. One is residual block with stride of 1, and the other one is a block with stride of 2 for downsizing. There are three layers for both blocks. The first layer is 1×1 convolution with ReLU, and the second layer is the depth-wise convolution. The third layer is another 1×1 convolution, but, without any nonlinearity. The number of iterations and the rate are set to 10^7 and 10^{-5}, and the number of the epoch is 160.

3. Results and discussion

As described above, the classification and segmentation accuracy of screening mammograms can be interpreted on the basis of the end-to-end training models implemented in this article. 322 breast images are taken from the MIAS, 564 breast images are taken from the DDSM and 330 from the CBIS-DDSM datasets are examined in order to determine the effectiveness of the suggested techniques [26–28]. For each patient, our new used databases include CC and MLO images of two breasts. The cases are clustered into one of the two categories: Benign or Malignant. DDSM, MIAS and CBIS-DDSM contain 534, 302 and 300 cases which have both MLO and CC view of each breast. Therefore, the DDSM and MIAS augmented images are 2136 and 1208 images, respectively. The classification criteria are based on the availability of mammogram-containing tumors, as the MIAS, DDSM and CBIS-DDSM databases contain certain non-tumor mammograms. The Intersect over Union (IoU),Dice Coefficients (DC), Accuracy, Sensitivity or recall, Precision, $F1$- score, Area Under the ROCcurve (AUC) and computational time are our evaluation metrics. Accuracy is a measure of true predictions as in Eq. (1). Precision is the positive predictive value or fraction of detected malignant cases that match the ground truth as in Eq. (2). Sensitivity is the true positive rate or the fraction of true malignant cases that are detected malignant as in Eq. (3). The $F1$-score is the harmonic mean of Precision and Sensitivity and represents a more generalized form balancing both. It gives the sample sets similarity and diversity as illustrated in Eq. (4). The AUC tests the whole two-dimensional region below the entire ROC curve and gives an overall output calculation over all possible classification thresholds. All these metrics are defined as follows [29]

$$ACC = (TP + TN)/(TP + FP + TN + FN) \tag{1}$$

$$Pr = TP/(TP + FP) \tag{2}$$

$$Se = TP/(TP + FN) \tag{3}$$

$$F1 = 2(Pr \times Se)/(Pr + Se) \tag{4}$$

$$IoU = \frac{|y \cap y'|}{|y \cup y'|} \tag{5}$$

Table 2 Segmentation results based on modified U-Net model for the MIAS, DDSM and CBIS-DDSM databases based on MLO view.

Datasets	IoU	DC
DDSM	92.99%	91.89%
MIAS	90.78%	89.99%
CBIS-DDSM	87.96%	88.65%

Table 1 Number of samples of training and testing for all datasets used, techniques for data augmentation and number of images produced.

Databases	Samples of training	Samples of testing	Total samples	Number of images after the process of augmentation
DDSM	451	113	564	1804
MIAS	257	65	322	1028
CBIS-DDSM	NA	330	330	NA

Table 3 Classification results with and without data augmentation for different models for the DDSM, MIAS and CBIS-DDSM based on the MLO view.

Classification results with and without data augmentation										
Model	**Accuracy %**		**AUC %**		**Sensitivity %**		**Precision %**		**F1-score %**	
	Without augmentation	**With**	**Without Augmentation**	**With**	**Without Augmentation**	**With**	**Without augmentation**	**With**	**Without Augmentation**	**With**
DDSM Database										
InceptionV3	88.87	93.85	87.99	92.99	88.74	93.54	88.59	93.19	87.99	93.23
DenseNet121	85.97	91.57	84.99	90.89	85.89	91.49	85.84	91.52	85.85	91.35
ResNet50	84.54	89.58	83.99	88.99	83.94	89.84	84.12	90.23	84.21	90.32
VGG16	82.81	85.98	81.99	84.88	82.74	85.94	82.43	84.99	82.34	85.32
MobileNetV2	80.97	82.77	80.89	82.65	81.74	82.98	80.88	82.76	80.99	82.65
MIAS Database										
InceptionV3	86.77	91.65	85.89	90.99	86.84	91.34	85.48	91.19	85.88	90.89
DenseNet121	83.87	89.99	82.88	90.32	83.77	89.59	83.74	90.22	83.75	89.88
ResNet50	82.65	88.58	81.79	87.99	81.96	88.54	82.46	88.43	82.65	87.89
VGG16	80.98	86.78	80.87	86.98	80.84	86.94	80.83	86.78	80.84	86.79
MobileNetV2	79.97	84.77	79.54	84.85	79.84	84.98	80.21	84.86	80.11	84.95
CBIS-DDSM Database										
InceptionV3	84.21	90.55	84.87	90.49	84.87	90.23	83.99	90.59	83.96	90.69
DenseNet121	82.47	87.69	82.65	87.35	82.57	87.69	81.99	87.42	82.34	87.58
ResNet50	81.65	86.78	80.89	86.89	80.99	86.94	81.32	86.53	81.22	86.23
VGG16	80.98	84.68	80.77	84.68	81.89	84.94	81.99	84.88	81.84	84.89
MobileNetV2	79.82	83.47	78.99	83.95	79.87	83.99	79.99	83.98	79.89	83.99

Table 4 Segmentation with classification results with and without data augmentation for different models based on the MLO view.

Segmentation with classification results without data augmentation										
Model	**Accuracy%**		**AUC%**		**Sensitivity%**		**Precision%**		**F1-score%**	
	Without augmentation	**With**	**Without Augmentation**	**With**	**Without Augmentation**	**With**	**Without augmentation**	**With**	**Without Augmentation**	**With**
DDSM Database										
Inception V3 + Modified segmentation of the U-Net	95.78	98.87	95.69	98.88	95.88	98.98	95.54	98.79	94.78	97.99
DenseNet121 + Modified segmentation of the U-Net	93.99	96.99	94.54	97.14	94.77	97.47	94.89	97.49	94.34	97.24
ResNet50 + Modified segmentation of the U-Net	93.57	96.87	93.97	96.47	94.21	97.01	94.61	97.11	93.89	96.99
VGG16 + Modified segmentation of the U-Net	92.67	95.97	92.34	96.11	92.81	95.88	92.81	96.21	92.69	95.99
MobileNetV2 + Modified segmentation of the U-Net	90.87	93.88	91.47	93.97	91.21	93.78	90.97	93.87	90.99	92.99
MIAS Database										
Inception V3 + Modified segmentation of the U-Net	93.88	96.87	93.79	97.21	93.98	96.88	93.97	96.99	93.79	97.12
DenseNet121 + Modified segmentation of the U-Net	91.89	95.78	91.94	94.99	90.99	95.77	91.59	95.69	90.88	95.84
ResNet50 + Modified segmentation of the U-Net	90.67	94.27	90.91	94.87	90.98	94.78	90.81	94.61	89.89	94.59
VGG16 + Modified segmentation of the U-Net	89.47	93.87	89.54	93.61	89.78	93.79	89.81	93.78	89.49	93.69
MobileNetV2 + Modified segmentation of the U-Net	88.45	91.42	87.99	91.27	88.81	90.78	88.87	91.43	88.99	90.99
CBIS-DDSM Database										
Inception V3 + Modified segmentation of the U-Net	91.78	94.18	91.29	94.65	91.65	94.76	91.97	94.65	91.89	94.12
DenseNet121 + Modified segmentation of the U-Net	90.32	92.34	90.44	92.56	89.99	92.77	90.59	92.69	89.78	92.84
ResNet50 + Modified segmentation of the U-Net	89.47	91.88	89.78	91.87	88.58	91.78	89.81	91.61	88.89	91.59
VGG16 + Modified segmentation of the U-Net	87.32	90.79	87.21	91.01	87.78	90.65	87.32	90.98	87.45	90.69
MobileNetV2 + Modified segmentation of the U-Net	86.45	89.78	85.99	90.27	85.89	89.98	86.87	90.23	85.99	89.99

Table 5 Classification results with and without data augmentation for different models for the DDSM, MIAS and CBIS-DDSM based on MLO and CC views.

Segmentation with classification results without data augmentation										
Model	**Accuracy%**		**AUC%**		**Sensitivity%**		**Precision%**		**F1-score%**	
	Without augmentation	**With**	**Without augmentation**	**With**	**Without Augmentation**	**With**	**Without augmentation**	**With**	**Without Augmentation**	**With**
DDSM Database										
Inception V3 + Modified segmentation of the U-Net	96.45	99.43	96.87	99.22	96.96	99.12	96.86	98.99	95.88	98.98
DenseNet121 + Modified segmentation of the U-Net	95.89	98.89	95.44	98.21	95.67	98.98	95.36	98.69	95.63	98.33
ResNet50 + Modified segmentation of the U-Net	94.87	97.97	94.24	97.87	95.98	98.11	95.69	97.98	94.49	97.79
VGG16 + Modified segmentation of the U-Net	93.87	96.21	93.65	96.89	93.88	96.78	93.88	96.89	93.79	96.87
MobileNetV2 + Modified segmentation of the U-Net	92.52	94.98	92.87	94.87	92.98	94.98	91.99	94.87	92.87	94.54
MIAS Database										
Inception V3 + Modified segmentation of the U-Net	94.32	97.87	94.59	98.01	94.88	97.78	94.32	97.32	94.52	97.99
DenseNet121 + Modified segmentation of the U-Net	92.89	96.48	92.88	95.99	92.99	96.57	92.98	96.59	92.88	96.21
ResNet50 + Modified segmentation of the U-Net	91.57	95.63	91.89	95.01	91.99	95.48	91.51	95.22	91.08	95.36
VGG16 + Modified segmentation of the U-Net	90.54	94.36	90.95	94.77	90.88	94.89	90.88	94.35	90.99	94.32
MobileNetV2 + Modified segmentation of the U-Net	89.99	93.99	88.89	93.98	89.98	93.28	89.87	93.85	89.87	92.99
CBIS-DDSM Database										
Inception V3 + Modified segmentation of the U-Net	93.21	96.01	93.59	96.11	93.32	96.12	93.63	96.32	93.29	96.12
DenseNet121 + Modified segmentation of the U-Net	92.96	94.96	92.64	94.36	91.89	94.37	92.21	94.89	91.88	94.21
ResNet50 + Modified segmentation of the U-Net	91.27	93.58	91.81	93.81	91.22	93.96	91.32	93.32	90.99	93.29
VGG16 + Modified segmentation of the U-Net	89.96	92.89	89.89	92.98	89.98	92.85	89.98	92.38	89.89	92.23
MobileNetV2 + Modified segmentation of the U-Net	88.95	90.32	88.59	91.99	87.79	91.98	88.98	92.01	87.99	92.21

Table 6 Computational time of the segmentation with classification and data augmentation system.

Model	Time, s
Inception V3 + Modified segmentation of the U-Net	1.2134
DenseNet121 + Modified segmentation of the U-Net	2.2365
ResNet50 + Modified segmentation of the U-Net	2.3254
VGG16 + Modified segmentation of the U-Net	1.9897
MobileNetV2 + Modified segmentation of the U-Net	1.6587

Table 7 Quantitative distinction between our model and the state of the art for the task of classification of the DDSM database.

Reference	Number of mammograms Databases	Name of Database	Accuracy %	AUC %	Sensitivity %	Precision %	F1-score %	DC %	IoU %
Our proposed work	**1804**	**DDSM**	**98.87**	**98.88**	**98.98**	**98.79**	**97.99**	**91.89**	**92.99**
[30]	DDSM	NA	NA	NA	NA	NA	NA	88	NA
[31]	200	MIAS and DDSM	97.73	NA	92.50	NA	NA	NA	NA
[32]	402	MIAS	NA	NA	NA	NA	NA	NA	NA
[33]	300	DDSM	98	NA	97.40	NA	NA	NA	NA
[34]	251	MIAS	96	NA	83	NA	NA	NA	NA
[35]	728	IN-breast	95.64	NA	97.14	NA	NA	NA	NA
[36]	1804	DDSM	97.98	98.46	97.63	96.51	95.97	NA	NA

Table 8 Performance comparison of current CADs using combination MLO and CC view mammographic features.

Reference	Accuracy %	AUC %	Sensitivity %	Precision %	F1-score %	DC %	IoU %
Our proposed work	**99.43**	**99.22**	**99.12**	**98.99**	**98.98**	**94.79**	**94.89**
[37]	NA	NA	85	NA	NA	NA	NA
[38]	NA	NA	91	NA	NA	NA	NA
[39]	NA	NA	93	NA	NA	NA	NA
[40]	NA	NA	80	NA	NA	NA	NA
[41]	83.13	NA	77.08	NA	NA	NA	NA
[42]	93.98	NA	97.37	NA	NA	NA	NA
[43]	96.6	NA	95	NA	NA	NA	NA

$$DC = \frac{2.|y \cap y'|}{|y| + |y'|} \tag{6}$$

where FP is the false positive of the non-lesion pixel segmented as a lesion pixel that means the sample is malignant which is wrong diagnosis, and FN is the false negative of lesion pixel segmented as a non-lesion pixel that means the sample is benign which is wrong diagnosis. TP is the true positive that means that the database sample is malignant which is correct diagnosis, while TN is the true negative that means that the database sample is benign which is correct diagnosis. The IoU is utilized to quantify the percentage overlap between the target mask and our prediction malignant or benign dignosis. When IoU increases, the system performance is enhanced. Here, y represents the ground truth mask, and y' represents the probability map generated by neural network. Moreover, DC is called loss function.

In our work, mammogram datasets are chosen using the Python simulator to test the proposed method. The number of analysis and research samples for all datasets is illustrated in Table 1. Data augmentation is conducted on samples that are rotated at four angles of 0, 90, 180 and 270 degrees to increase accuracy, as shown in Table 1.

Table 2 explains the segmentation results of IoU and DC for our databases based on the modified U-Net model. The obtained classification results with and without data augmentation are represented in Table 3 for the MIAS, DDSM and CBIS-DDSM databases utilizing the MLO view.

Moreover, Table 4 explains the results of classification with and without data augmentation after utilizing the modified U-Net model to segment the databases based on the MLO view.

Table 5 introduces our proposed CAD system based in the combination between MLO and CC views.

Table 6 explains the computing time of all the system. The proposed segmentation and classification systems are compared with other recent CAD systems and the results are shown in Table 6. Our proposed system is compared with other systems illustrated in Refs. [30–37] and the results are shown in Tables 3-6. The comparison declares the superiority of our proposed system in accuracy, AUC, precision and F1 score as shown in Tables 7 and 8.

4. Conclusion

In this paper, a new framework is proposed based on different deep learning models, including InceptionV3, DenseNet121, ResNet50, VGG16 and MobileNetV2, for breast cancer diagnosis using digitized mammograms with high accuracy and low computational time. The proposed framework utilizes the modified segmentation of the U-Net model for the segmentation process. The diagnosis performance is evaluated in terms of the IoU, DC, accuracy, sensitivity, precision, AUC, F1- score and computational time, where the training is initialized by weights of a network that has already been trained using another dataset. The data augmentation technique is used to resolve data shortcomings and introduce variety to the dataset, which enhances the generalization capability of the pre-trained network and thus alleviates over-fitting, as a large amount of training data is needed. This study is very beneficial and shows that there is no need for a human interface with pre- or post-processing or hand-crafted features. The data augmentation with modified segmentation of the U-Net model and InceptionV3model achieves the best performance: 98.87% accuracy, 98.88% area under the curve (AUC), 98.98% sensitivity, 98.79% precision, 97.99% F1 score and computational time 1.2134 s on DDSM datasets. The proposed frame work when utilizing the combination between MLO and CC view achieves better performance than utilizing MLO view only, where the metrics are enhanced to: 99.43% accuracy, 99.22% AUC, 99.12% sensitivity, 98.99% precision, 98.98% F1 score. The obtained results reveal that our proposed models achieve better performance than that in the literature with more than 10%.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, CA Cancer J. Clin. 65 (1) (2015) 5–29.

[2] D. Saslow, C. Boetes, W. Burke, S. Harms, M.O. Leach, C.D. Lehman, E. Morris, E. Pisano, M. Schnall, S. Sener, R.A. Smith, American cancer society guidelines for breast screening with MRI as an adjunct to mammography, CA Cancer J. Clin. 57 (2) (2007) 75–89.

[3] I. Maniecka-Bryła, M. Bryła, P. Bryła, M. Pikala, "The burden of premature mortality in Poland analysed with the use of standard expected years of life lost, BMC Public Health 15 (1) (2015) 101.

[4] R.A. Hubbard, K. Kerlikowske, C.I. Flowers, B.C. Yankaskas, W. Zhu, D.L. Miglioretti, Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: A cohort study, Ann. Intern. Med. 155 (8) (2011) 481–492.

[5] L. Tsochatzidis, L. Costaridou, I. Pratikakis, Deep learning for breast cancer diagnosis from mammograms - a comparative study, J. Imaging 5 (37) (2019) 1–11.

[6] A. Shrestha, A. Mahmood, Review of deep learning algorithms and architectures, IEEE Access 7 (2019) 53040–53065.

[7] S. Targ, D. Almeida, K. Lyman, Resnet in Resnet: Generalizing residual architectures, In International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, 2-4 May, pp. 1–4, 2016.

[8] X. Zhang, J. Zou, K. He, J. Sun, Accelerating very deep convolutional networks for classification and detection, IEEE Trans. Pattern Anal. Mach. Intell. 38 (10) (2015) 1943–1955.

[9] H. Shin, H.R. Chang, G. Roth, L. Lu, X. Mingchen, I. Ziyue, Y. Nogues, D. Jianhua, Mollura, R.M. Summers, Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, IEEE Trans. Med. Imaging, 35 (5) (2016) 1285–1298.

[10] S. Sasikala, M. Bharathi, M. Ezhilarasi, M. Ramasubba Reddy, S. Arunkumar, Fusion of MLO and CC view binary patterns to improve the performance of breast cancer diagnosis, Curr. Med. Imaging 14 (4) (2018) 651–658.

[11] J. Arevalo, F.A. González, R. Ramos-Pollán, J.L. Oliveira, M. A.G. Lopez, Representation learning for mammography mass lesion classification with convolutional neural networks, Comput. Methods Programs Biomed. 127 (15) (2016) 248–257.

[12] D. Abdelhafiz, C. Yang, R. Ammar, S. Nabavi, Deep convolutional neural networks for mammography: Advances, challenges and applications, BMC Bioinf. 20 (11) (2019) 1–20.

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

[14] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention, 9351, Springer, Cham, 2015, pp. 234–241.

[15] N. Alam, A. Oliver, E.R. Denton, R. Zwiggelaar, Automatic segmentation of microcalcification clusters, Annual Conference on Medical Image Understanding and Analysis, 894, Springer, Cham, 2018, pp. 251–261.

[16] S. Duraisamy, S. Emperumal, Computer-aided mammogram diagnosis system using deep learning convolutional fully complex-valued relaxation neural network classifier, IET Comput. Vis., 11 (8) (2017) 656–662.

[17] E. Deniz, A. Şengür, Z. Kadiroğlu, Y. Guo, V. Bajaj, Ü. Budak, Transfer learning based histopathologic image classification for breast cancer detection, Health Inform. Sci. Syst. (1) (2018) 1–7.

[18] S. Kwok, Multiclass classification of breast cancer in whole-slide images, International conference image analysis and recognition, 10882, Springer, Cham, 2018, pp. 931–940.

[19] Li, Chen, Dan Xue, Hu. Zhijie, Hao Chen, Yao. Yudong, Yong Zhang, Mo Li, Qian Wang, Xu. Ning, A Survey for breast histopathology image analysis using classical and deep neural networks, in: International Conference on Information Technologies in Biomedicine. Springer, Cham, vol. 1011, 2019, pp. 222–233.

[20] D.A. Ragab, M. Sharkas, S. Marshall, J. Ren, Breast cancer detection using deep convolutional neural networks and support vector machines, PeerJ 7 (4) (2019) e6201.

[21] C. Liang, M. Li, Z. Bian, W. Lv, D. Zeng, J. Ma, Establishment of a deep feature-based classification model for distinguishing benign and malignant breast tumors on full-filed digital mammography, J. Southern Med. Univ. 39 (1) (2019) 88–92.

[22] S.C. Wong, A. Gatt, V. Stamatescu, and M.D. McDonnell, Understanding data augmentation for classification: when to warp, in: 2016 International Conference on Digital Image Computing: IEEE Techniques and Applications (DICTA), Gold Coast, QLD, Australia, pp. 1–6, 2016.

[23] Y. Chen, T. Zheming, Z. Yang, S. Holly, L. Norford, Transfer learning with deep neural networks for model predictive control of HVAC and natural ventilation in smart buildings, J. Cleaner Prod. 254 (119866) (2020) 1–10.

[24] M.B. Tayel, A.M. Elbagoury, Breast infrared thermography segmentation based on adaptive tuning of a fully convolutional network, Curr. Med. Imaging 16 (5) (2020) 611–621.

[25] M.S. Hossain, Micro-calcification segmentation using modified u-net segmentation network from mammogram images, J. King Saud University-Comput. Inform. Sci., in Press, online 4 Nov. 2019. http://doi.org/10.1016/j.jksuci.2019.10.014.
[26] .
[27] . Accessed 7 June 2019.
[28] https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM. Accessed 1 June 2019.
[29] M. Long, Y. Cao, Z. Cao, J. Wang, H. Zhu, M.I. Jordan, Transferable representation learning with deep adaptation networks, IEEE Trans. Pattern Anal. Mach. Intell. 41 (12) (2018) 3071–3085.
[30] A. Amyar, R. Modzelewski, H. Li, S. Ruan, Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation, Comput. Biol. Med. 126 (2020) 104037.
[31] M. Dong, X. Lu, Y. Ma, Y. Guo, Y. Ma, K. Wang, An efficient approach for automated mass segmentation and classification in mammograms, J. Digit. Imaging 28 (2) (2015) 613–625.
[32] A.K. Mohanty, M. Senapati, S. Beberta, S.K. Lenka, Retraction Note to: Mass classification method in mammograms using correlated association rule mining, Neural Comput. Appl. 23 (2) (2013) 273–281.
[33] W. Xie, Y. Li, Y. Ma, Breast mass classification in digital mammography based on extreme learning machine, Neuro Comput. 173 (3) (2016) 930–941.
[34] K.U. Sheba, G.S. Raj, An approach for automatic lesion detection in mammograms, Cogent Eng., 5 (1) (2018) Article 1444320, pp. 1–16, http://doi.org/10.1080/23311916.2018.1444320.
[35] M.A. Al-antari, M.A. Al-masni, M.T. Choi, S.M. Han, T. Seong Kim, A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification, Int. J. Med. Inf. 117 (2018) 44–54.
[36] W.M. Salama, A.M. Elbagoury, M.H. Aly, Novel breast cancer classification framework based on deep learning, IET Image Proc. 14 (13) (2020) 3254–3259.
[37] R.G. Blanks, M.G. Wallis, R.M. Given-Wilson, Observer variability in cancer detection during routine repeat (incident) mammographic screening in a study of two versus one view Mammography, J. Med. Screening 6 (3) (1999) 152–158.
[38] S. Paquerault, N. Petrick, H.P. Chan, B. Sahiner, M.A. Helvie, Improvement of computerized mass detection on mammograms: Fusion of two-view information, Med. Phys. 29 (2) (2002) 238–247.
[39] S.J. Kim, W.K. Moon, N. Cho, J.H. Cha, S.M. Kim, J.G. Im, Computer-aided detection in digital Mammography: Comparison of craniocaudal, mediolateral oblique, and mediolateral views, Radiology 241 (3) (2006) 695–701.
[40] B. Sahiner, H.P. Chan, L.M. Hadjiiski, M.A. Helvie, C. Paramagul, J. Ge, Joint two-view information for computerized detection of micro calcifications on Mammograms, Med. Phys., 33 (1) (2006) 2574–2585.
[41] R.D. Dantas, M.Z. do Nascimento, R. de Souza Jacomini, D.C. Pereira and R.P. Ramos, Fusion of two-view information: SVD based modeling for computerized classification of breast lesions on Mammograms, in: Mammography-recent advances. InTech., pp. 261–278, 2012.
[42] L. Sun, L. Li, W. Xu, W. Liu, J. Zhang, G. Shao, A novel classification scheme for breast masses based on multi-view information fusion, in: Proceedings of 4th IEEE International Conference on Bioinformatics and Biomedical Engineering (iCBBE), 2010, pp. 1–4.
[43] S.M. Sasikala, Ezhilarasi, S. Arun Kumar, Detection of breast cancer using fusion of MLO and CC view features through a hybrid technique based on binary firefly algorithm and optimum-path forest classifier, in: Applied Nature-Inspired Computing: Algorithms and Case Studies, pp. 23–40. Springer, Singapore, 2020.