ORIGINAL PAPER

A New method for promote the performance of deep learning paradigm in diagnosing breast cancer: improving role of fusing multiple views of thermography images

Mahsa Ensafi[1] · Mohammad Reza Keyvanpour[1] · Seyed Vahab Shojaedini[2]

Abstract

Purpose Breast cancer is one of the deadliest cancers among women worldwide which its early detection may significantly reduce its mortality rate. Thermgraphy is a new, non-invasive, non-painful, and low-cost modality that detects abnormalities by detecting heat from the breast surface.

Method Recent research has applied deep learning to early breast cancer diagnosis via thermography, using only the frontal view of thermograms. We combine several views of thermal images to improve the performance of pre-trained deep learning architectures in this article. This goal is achieved by combining frontal-45 data with lateral-45 and lateral45 thermograms to construct a detection model that utilizes transfer learning.

Result Research in this area uses the Database for Mastology Research (DMR) with infrared images. In this study, transfer based deep learning methods are demonstrated to be effective in fusing several views of thermograms to diagnose breast cancer in a manner that can result in a sensitivity increase of 2-15 percent and a specificity increase of 2-30 percent compared to other deep learning-based or handcrafted schemes.

Conclusion Using multiple views of thermograms and transfer learning, this paper proposes a method for improving breast cancer diagnosis. Using methods based on deep learning and methods based on hand-crafted features, we evaluated the performance of the proposed model. Using the obtained results as a basis for future research, the proposed design can be improved and developed as a valid approach in interpreting breast thermography images.

Keywords Breast cancer · Thermography · Deep learning · Transfer learning

1 Introduction

Cancer refers to abnormal cell growth that may be benign or malignant. The benign lesion does not affect the surrounding tissues, while the malignant lesion spreads throughout the body and consumes on the nutrients of other tissues [1].

Breast cancer is one of the fatal cancers which is common among women but may affect men. According to the World Health Organization(WHO) official reports, from 2016 to 2020, approximately 7.8 million cases of breast cancer have been reported which 685,000 of them have

died. Early detection of breast cancer leads to easier treatment in parallel with reduction of the mortality rate caused by this disease [2]. Based on this fact, breast screening which is performed by using different imaging techniques plays an important role in the early diagnosis and treatment of breast lesions. consequently, periodical breast screening is recommended for women aged 40 to 75 years old. Among screening modalities, the histopathology examination is the gold-standard screening method [3] in which an expert examines biopsy (i.e., removed sample of tissue) in laboratory by capturing microscopic images. These images make it possible to understand the microanatomy of cells, tissues, and organs by examining the relationship between structure and function [4]. The invasive nature of this method in parallel with its highly dependence to physician specialty may restrict its application. In addition, this method is so expensive and time consuming [4]. Based on the above limitations some non-invasive methods have been developed for breast cancer screening which the most common among

✉ Mohammad Reza Keyvanpour
keyvanpour@alzahra.ac.ir

[1] Faculty of Engineering and Technology, Alzahra University, Tehran, Iran

[2] Department of Biomedical Engineering, Iranian Research Organization for Science and Technology, Tehran, Iran

them is mammography. Unfortunately, this technique has several major disadvantages, for instance, this modality requires expensive equipment which is not available in rural and remote areas [2]. Furthermore, the sensitivity of this technique is ineffective for woman younger than 45 years old or women with dense breasts. In addition, mammography uses X-rays for imaging, and exposure to radiation increases the risk of breast cancer. Finally, compression of the breast area during imaging causes anxiety and pain in some patients [5]. Another method of breast cancer screening is ultrasound imaging which uses sound waves that are safer than X-rays to penetrate the breast [4]. This method has drawbacks including low image contrast and low sensitivity [4, 6], therefore is considered as a complementary method for mammography.

In recent years thermography has been introduced as a non-invasive, painless, without physical contact, and low-cost technique for screening the breast cancer [7]. This modality detects the presence of breast lesions by measuring the temperature distribution of the skin [2, 8]. Healthy tissues have a certain temperature distribution due to metabolic activity, but cancer cells increase the rate of metabolic activity, which produces heat [2]. The heat emitted from the breast surface may be converted into visual information (i.e., thermograms) through so called thermal cameras [8]. Women of any age may use thermal imaging, even with dense breasts that are not recommended for mammography. This technique does not use external radiation and also, it has the potential of predicting cancer 8-10 years earlier than mammography [9]. In 1982, thermography was approved by the Food and Drug Administration (FDA) as an adjunct to diagnosis of breast cancer [9]. There are two protocols for such imaging including Dynamic Infrared Thermography (DIT) and Static Infrared Thermography (SIT). In DIT, imaging is performed after applying of cold stress. In spite of the cold, the cancerous areas remain almost unchanged, while the healthy areas show a good decrease in temperature. In contrast, SIT measures the temperature distribution without considering its changes over time. This method requires more difficult environmental conditions and more time for the patient to adapt to the examination room [10].

Although thermography is a promising technique, but unfortunately analyzing thermographic images is difficult [11]. Manual interpretation of thermograms is done by human experts which its performance is highly dependent to ability and conditions of the specialist who analyses the image. Therefore, the automated analyzing of thermograms plays an important role in diagnosis of breast cancer. In recent years, parameters such as advancement of infrared cameras, the increase in the processing capacity of computers, the development and explosion of digital image processing and machine learning algorithms and finally the cheaper and easier access to Graphical Processing Unit(GPU)-based

cloud computing resources, have made thermography an effective method for automatic diagnosis of breast lesions [5]. Considering the above reasons machine learning techniques have been widely used for breast thermography over the last decade. There are two strategies for diagnosing breast cancer using machine learning techniques. The first on is based on traditional methods and the second one includes deep learning schemes. Traditional methods consist of two steps in which firstly, texture or statistical features are extracted manually from the input thermograms and then the extracted features are classified. The main disadvantage of this group of methods was the need for a specialist to extract the features. It is also challenging to distinguish benign and malignant lesions because of similarity between lesions and variety of characteristics [12].

In order to eliminate or reduce the above limitations, recently, researchers have focused on deep learning schemes in diagnosing breast cancer in thermograms. Deep learning is a sub-field of artificial intelligence and machine learning which may abstract the high-dimensional expression of the image through convolution and down-sampling in order to effectively classify the images [13]. This tool takes several signals as inputs, combines them linearly by using several filter banks, and then passes the combined signals through nonlinear processes to generate the output features [13].

In general, cancer detection methods using thermography can be divided into two approaches, traditional and based on deep learning. In the following, these two approaches will be examined.

2 Related work

In some researches the statistical features were applied in order to analyze breast thermograms. For instance, the authors of [14] proposed a system that classifies breast thermographic images into normal and abnormal categories by using mean, variance, standard deviation, skewness, elongation, entropy, amplitude, and median. The disadvantage of this method is the ineffectiveness of the extracted features and low performance. Another example of this group of methods is research reported in [15] in which a local tissue feature extraction technique called Block Variance (BV) has been developed for analysis of the breast thermal images.

In some other members of this group of methods some researches made use of statistical and texture features. Authors in [16] analyzed the 7 statistical plus 17 texture features extracted from both breasts of a thermogram that quantifies the symmetric and asymmetric thermal patterns of the breasts. Then, three feature sets: FStat (First order statistical features), STex (Second-order Texture Feature) and SSigFS (Statistically significant first order and second order features) are fed into six most widely used

classifiers for comparing the efficiency of each feature set. The obtained results showed that SSigFS provides more accuracy in comparison to these features. In this article, out of 24 features extracted, the selected features are less effective, which leads to reduced performance.

Some of methods utilized asymmetry analysis in order to interpret breast thermograms. For instance [17] proposed a new approach for breast thermogram image analysis by developing a fully automated right and left breast segmentation in order to asymmetry analysis, by using breast shape features and polynomial curves. Based on the gray surface matrix and histogram texture features were extracted from the segmented images and were classified by using traditional classifiers such as Support Vector Machine(SVM).

In some studies, methods have been introduced to extract Region of Interest (ROI) from breast thermograms. The article [7] discuses on several segmentation methods, that have been published between 2010 and 2021.

In [18], authors proposed an automatic method to segment the ROI based on color intensities, thresholding operators, local contrast enhancement and statistical operators. In this paper, Artificial Neural Network(ANN) was used to classify handmade features as normal and abnormal. However, some features which have been extracted in that article are not optimally effective.

To overcome the limitations of traditional methods which are based on handcrafted features, in recent years, deep learning has widely been used to achieve high-accuracy diagnostic models in order to analyze breast thermograms. The common deep learning algorithm in the field of image analysis is Convolution Neural Network (CNN) [19].

In some studies, transfer learning has been used to extract distinctive features of thermograms. The authors of [20] considers application of Infrared Thermography(IRT) in the diagnosis of various diseases such as skin and breast cancer detection in preliminary stages, dry eye syndromes, and ocular issues, liver disease, diabetes diagnosis and last but not least the novel COVID-19 virus. In the second stage they have proposed an autonomous breast tumor classification system by employing inception-v3 architecture.

In some studies, transfer learning has been used to extract distinctive features of thermograms. The authors of [21] utilized transfer learning g followed by the fine-tuning are performed to reduce consuming computational power and data augmentation to solve the problem of data lack. For high efficiency, they have combined the deep neural network model with some classifiers which one of them is SVM.

Another example of applying deep networks to interpret thermographic breast image is research which has been carried in [22]. In this article different pre-trained deep architectures such as resnet18, resnet34, resnet50, resnet152, vgg16 and vgg19 were applied by using Pytorch libraries.

Article [23] addresses the drawbacks of this method by not proposing a method for enriching the image feature.

The goal of several researches is to investigate the behavior of different recent deep learning methods [24]. For this, the authors developed classifiers based on CNN and modified inception v4 model (MV4) for identifying breast disorders in both grayscale and color images. Their purpose in introducing MV4 was to decrease the number of layers used to speed up the run time while having a minimal effect on accuracy level. The limitation of this method is ineffectiveness for grayscale images.

In this paper, a new deep learning scheme is presented in order to improve diagnosing breast cancer in thermograms which makes use of fusing multiple views of this type of images. The obtained results indicate that the performance of deep networks is promoted by our proposed integration of several views of thermographic images, compared to using one view of image at the input. To achieve this goal, first we build a diagnostic model for process of only-front images by utilizing transfer learning concept and subsequently, the extended model is developed by emerging both lateral-45˚and lateral45˚view images. State-of-the-art pre-trained CNN transfer learning models such VGG, ResNet, DenseNet and EfficientNet are used for fitting purpose in simple and extended approaches. Finally, the effect of incorporating lateral images on the diagnostic capacity of the models, are evaluated.

The rest of the paper is structured as follows. Section 3, explains the proposed scheme introduced in this article. In Sects. 4 and 5 the experimental analysis and comparison with other techniques are discussed and finally, conclusion remarks are presented in Sect. 6.

3 Proposed method

In this article, we present a method for diagnosing breast cancer by analyzing thermograms based on deep learning approaches. Fig. 1 shows the structure of the proposed method. The thermograms used in this article may be divided into three types based on the direction in which the images were captured. These types include frontal, lateral 45-right, lateral 45-left which firstly were preprocessed as follows. In the first step, we removed the blurred and bandaged images. The precise details of the original image may not be obtained from blurred thermograms. Fig. 2a shows an example of this type. Furthermore, some thermograms were bandaged for reasons such as surgery. This covering may damage the thermal pattern of the breast. Fig. 2b shows an example of such images. In addition to remove the mentioned images, we cropped unwanted parts like neck and armpit in images. Then right and left breast thermograms are converted
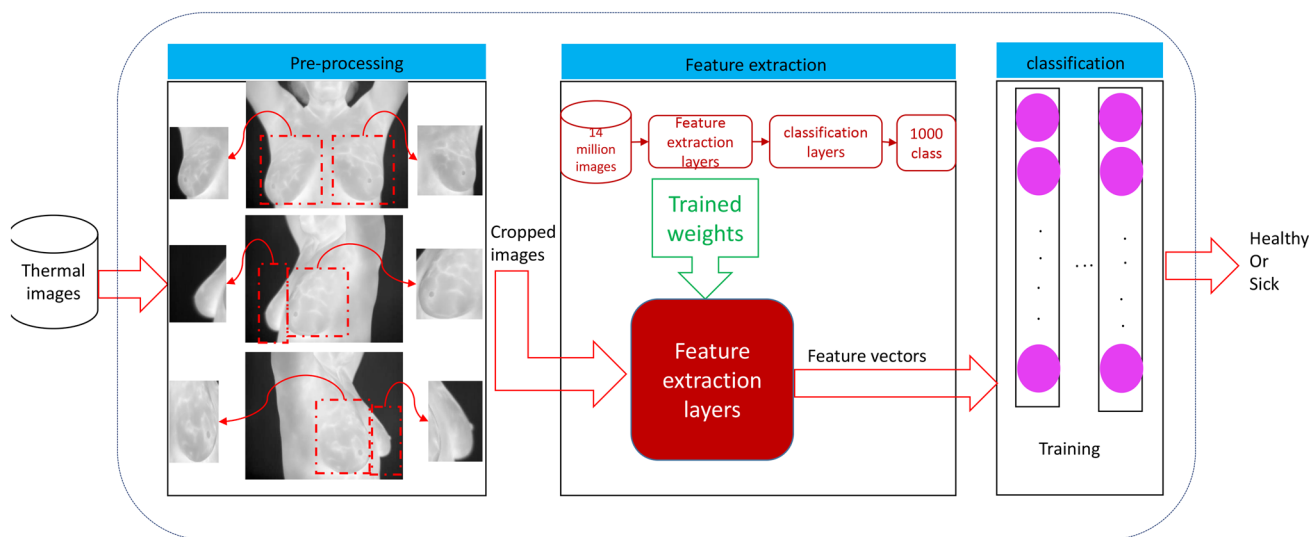
Fig. 1 structure of the proposed method

from RGB to gray scale. The process of image processing in CNN includes feature extraction and classification [5, 25]. As shown in Fig. 3, In the feature extraction section, the convolution layer is responsible for extracting the feature through the learnable filters. In other words, the convolution layer creates convolved image containing features (also known feature maps) of the previous layer image. The convolution of a filter is performed on a two-dimensional image via Eq. (1).

$$g(x,y) = I(x,y)^*f(x,y) = \Sigma_{-\infty}^{\infty}\Sigma_{-\infty}^{\infty}I(x-u,y-v)f(u,v) \quad (1)$$

where I(x,y) is the input image and f(x,y) is the filter and g(x,y) denotes convolved images. while that the value of jth node in L^{th} convolution layer is calculated through Eq. (2) [26].

$$X_j^L = f(\sum_{i\epsilon M_j}X_j^{L-1*}W_{ij}^L + b_J^L) \quad (2)$$

where f(.) is non-linear function and w^L and b^L indicates weight and bias of Lth layer respectively.

The pooling layer reduces the dimensions of the feature maps. Pooling is performed by maximizing or averaging on the convolved image, depending on the pool-size. Therefore, the size of output of convolutional layer(o_{conv}) is obtained through Eq. (3) [27].

Fig. 2 a. blurred thermogram **b.** bandaged breast **c.**normal thermogram

$$o_{conv} = \begin{cases} o_w = \frac{W-F_w+2p}{S} + 1 \\ o_h = \frac{H-F_h+2p}{S} + 1 \end{cases} \quad (3)$$

where P and S indicate size of padding and the stride of the convolution operation respectively [26]. Finally, in the fully connected section, dense or fully connected layers in the classification part may learn.

features by incorporating a set of parameters.

Despite of its great potential in extracting features and classification but, CNN requires a large amount of data to adapt. In several cases which the available database is small, transfer learning is an effective and helpful way to training networks. The concept of transfer learning is shown in Fig. 4. The weights of a pre-trained model can be transferred to a model with a similar application and the model can start with those weights. The pre-trained model is trained with large amounts of data and learns weights and biases during training [28]. Pre-trained CNN models such as VGG [29], ResNet [30], DenseNet [31], Efficient-Net [32], are available which were trained with the ImageNet database include more than 14 million images and 1000 classes.

Firstly, the results of only-front view classification are obtained by using front-view dataset including 132 and 84 healthy and sick images respectively. The networks used in this test are DenseNet121 [31], EfficientNetB0 [32] and VGG19 [29]. Then, the frontal and lateral views are merged and in order to obtained more accurate classification results. In this case, the number of healthy images is 237, including 100 frontal view, 68 left 45 view and 71 right 45 views. The number of sick images also is 222, including 83 frontal view, 68 left 45 view and 71 right 45 views. VGG16 [29], ResNet50 [30] and DenseNet201 [31] are the networks are used in this experiment. The pseudo-code of Fusing Multiple views of Thermography Images(FMVT) shows the process of the proposed method in this paper and the Table 1 illustrates detailed of hyper parameters of the used models.
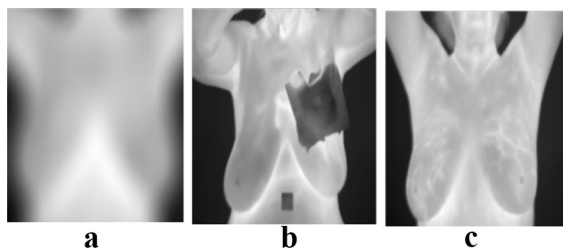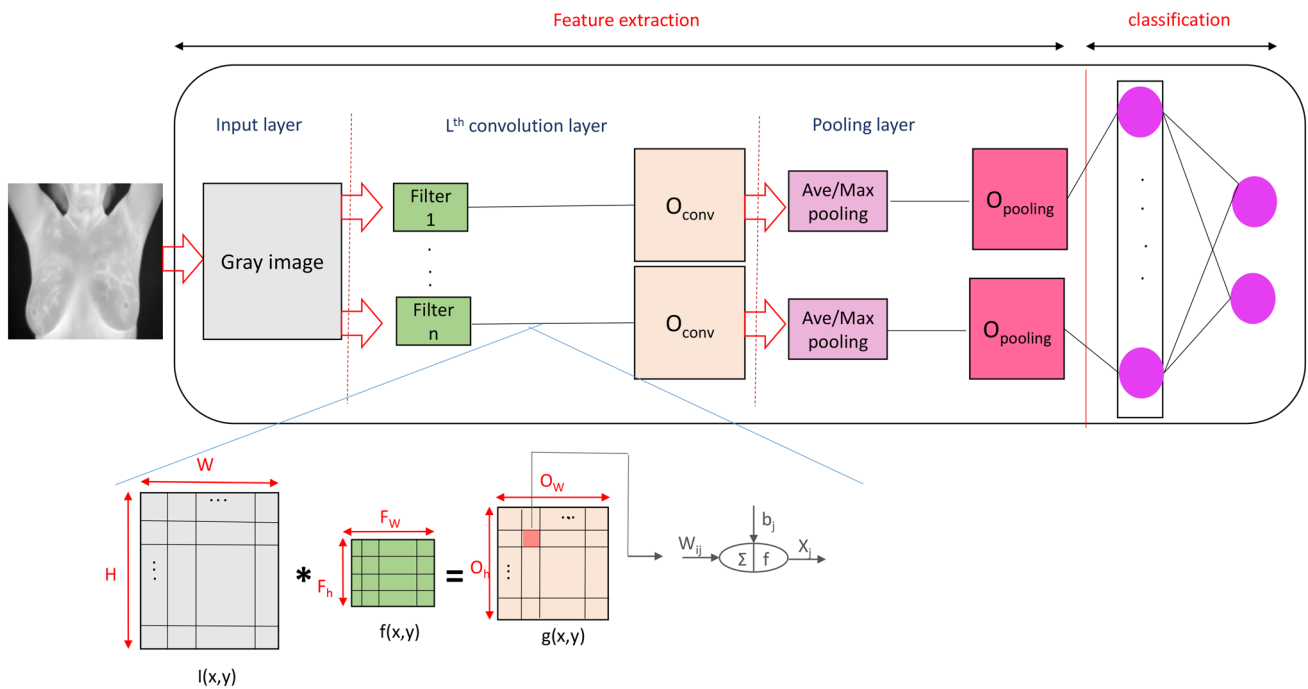
Fig. 3 Description of CNN layers

Pseudo- Code of FMVT

1. Input: breast thermograms in 2 views (frontal and 45°)
2. Output: label(healthy/sick)
3. prepare thermograms
4. while (criterion do not stop)
5. merge only frontal and 45°view
6. load feature extraction layer of selected pre-trained network
7. freeze weights
8. use the weights in step 6 to feature extraction of inputs
9. add and use GAP layer to transition from feature maps to an output prediction
10. add and use some Dense layer to map output of the last layer to 2 types (healthy or sick)
11. return prediction

Fig. 4 concept of transfer learning [35]

4 Experiments and results

The proposed algorithm was applied to a set of breast thermograms known as DMR [33] to evaluate its performance. In this database, IR imaging has been undertaken with two static and dynamic protocols. During the static protocol, five thermograms were recorded: a frontal, two laterals of the left at 45 and 90, and two laterals of the right at 45 and 90. In dynamic protocol, 20 sequential images are recorded in front position and 2 side images. The average number of images belonging each of patient is 27 and the size of the obtained thermograms is 640×480 pixels [34]. In addition to Infrared (IR) images, this database also includes digital mammograms and clinical data. These data were obtained from healthy and patients with any breast lesions, including cancer in the screening ward and gynecology department in the Antônio Pedro University Hospital in Brazil.

The data set was splited into four subsets with equal percentage (25%) to make possible utilizing fourfold cross-validation strategy. In order to implement learning models, we have used Tensorflow framework which are open-source machine learning libraries in Python. To run the program, we used Google Colab, a GPU framework provided by Google.

After performing a test, any of following situations may occur in the interpretation of the results. The first is about cases were correctly identified as sick by the model which are named true positive (TP). There were also misdiagnosed cases as sick by the classifier, such results are called false
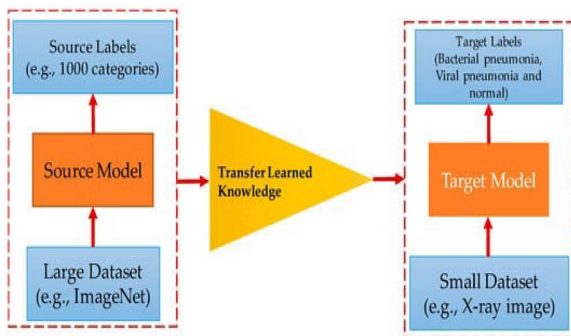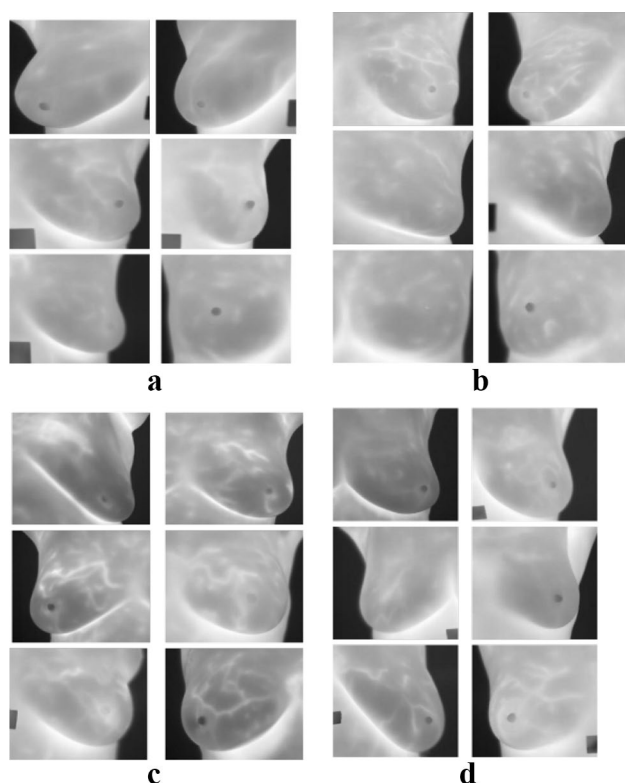
Fig. 5 examples of classification of only-front view **a**. false positive **b**. false negative **c**. true positive **d**. false negative

Fig. 6 examples of classification of emerged frontal and lateral view **a**. true negative **b**. false positive **c**. true positive **d**. false negative

positives (FP). On the other hand, there were cases that were correctly diagnosed as healthy by the classifier, which are called true negatives (TN), and finally, there were cases that were misdiagnosed as healthy by the classifiers, which are called false negatives (FN).

Table 2 shows the confusion matrix obtained for applying DenseNet121, EfficientNetB0 and VGG19 to diagnosis the nature of thermograms, among of them the VGG19 network has had a greater number of true negatives and true positives. In contrast Fig. 5 shows the true negative and true positive outputs of only-front image classification using same deep networks.

As shown in Table 3, in this case, network DenseNet201 has obtained better true negative items than the emerged frontal and lateral view. Some examples of results true negative and true positive output of emerged frontal and lateral view classification using DenseNet201 shown in Fig. 6.

After evaluating the six models (only-front and emerged frontal and lateral view), we found that the specificity and sensitivity improve due adding lateral view. This indicates that some of the images that were misdiagnosed in the first test were correctly identified in the second test of their class. For instance, Fig. 7 shows examples that were identified as false negatives and false positives in the only-front view
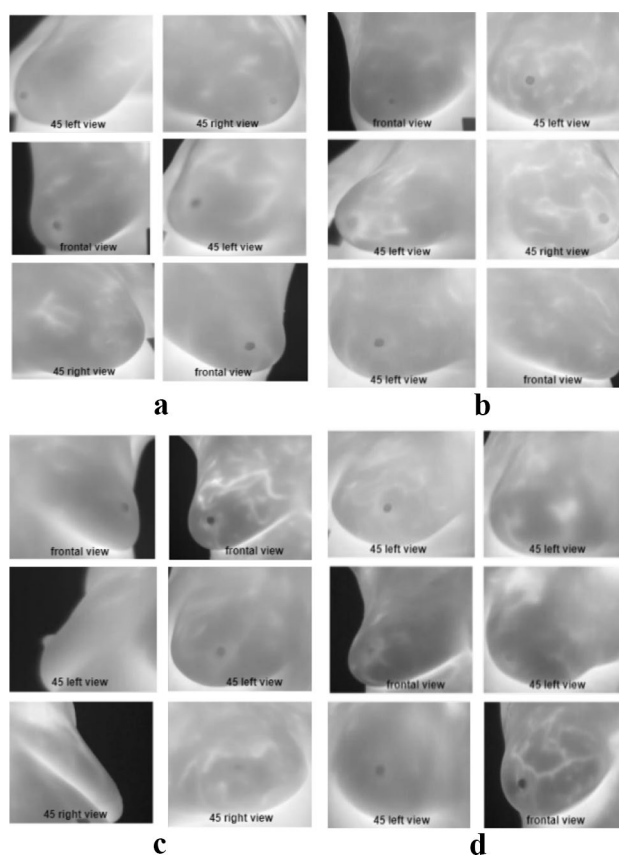
experiment, but most of them were correctly identified in the emerged frontal and lateral view experiment.

5 Discussion

The proposed method may successfully detect breast cancer by merging three views of thermograms. Furthermore, we have applied the transfer learning approach in order to reduce undesirable effects of the insufficient data volume.

5.1 Evaluation criteria

In order to evaluate the proposed model, common criteria have been used, including:

- Accuracy
  The accuracy parameter indicates how many samples identified right (according to Eq. (4)) [6].

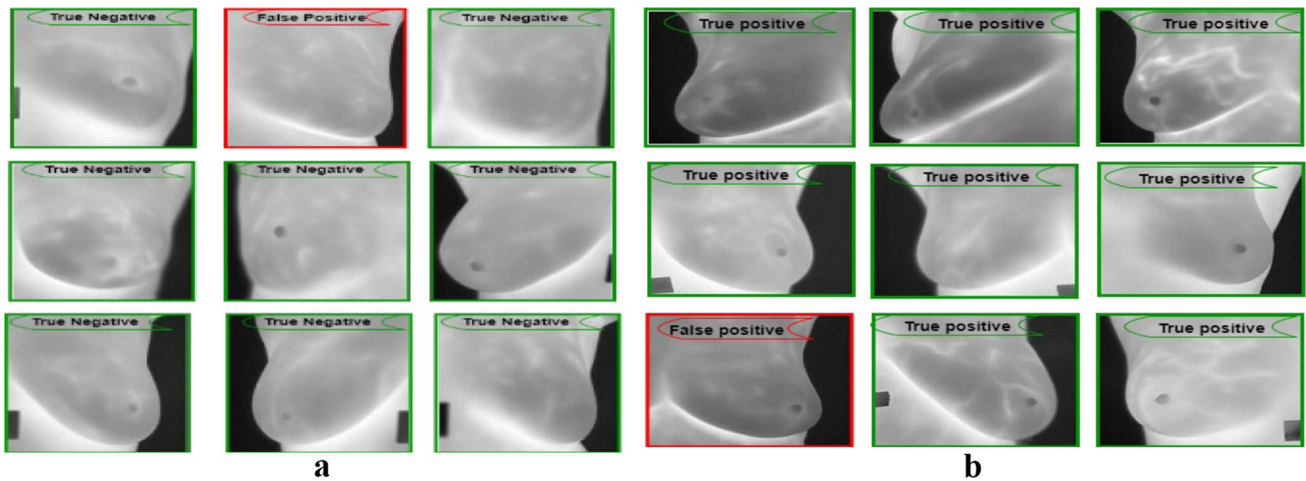$$ACC = \frac{TP + TN}{(TP + FP + TN + FN)} \quad (4)$$

Fig. 7 some result of emerged frontal and lateral view classification **a**. prediction of false positive samples in only-front view experiment, **b**. prediction of false negative samples in only-front view experiment

- Precision

    The precision is the number of positive cases that are correctly predicted and calculated by Eq. (5) [6].

$$Pre = \frac{TP}{(TP + FP)} \qquad (5)$$

- Sensitivity

    As shown in Eq. (6), this parameter indicates the number of positive samples that have been identified correctly [6].

$$Sn = \frac{TP}{(TP + FN)} \qquad (6)$$

- Specificity

    As shown Eq. (7), this parameter demonstrates the number of negative samples that have been identified right [6].

$$Sp = \frac{TN}{(TN + FP)} \qquad (7)$$

Tests and comparisons were performed in two different scenarios.

5.2 First scenario

Since the proposed method of this article is based on deep learning, so in the first scenario, the performance of the proposed method was compared with other methods, which are also in the field of deep learning and are considered the family of the method of this article. In this scenario of comparisons, we first examined the selected deep neural networks with *frontal* breast thermograms. Training with such data was not the main aim of this research and has been performed in various other articles. However, the above procedure was performed in order to verify the correctness

Table 1 details of the utilized deep models

Network	Structure	Configuration	Number of non-trainable parameters	Number of trainable parameters
VGG16	13 convolution 5 max pooling	GAP->FC(2048)->FC(128)->sig	7,635,264	8,392,449
VGG19	16 convolution 5 max pooling	GAP->FC(512)->FC(256)->sig	12,944,960	7,473,665
EfficientNetB0	16 MBConv block 1 Conv block	GAP->FC(1024)->sig	4,049,571	1,312,769
ResNet50	50 layers deep	GAP->FC(2048)->FC(1024)->FC(512)->FC(128)->sig	660,608	29,910,785
DenseNet121	121 layers deep	GAPFC(1024)FC(512)FC(256)FC(128)sig	1,149,440	7,626,817
DeseNet201	201 layers deep	GAPFC(1024)FC(512)FC(256) FC(128) sig	2,404,928	17,327,617
Optimizer	Batch size	Activation function at intermediate layers	Loss function	---
Adam	32	ReLU	Binary_Crossentropy	---

Table 2 confusion matrix in only-front view **a** DenseNet121 **b** EfficientNetB0 **c** VGG19

a		
-	healthy	sick
healthy	108	18
sick	24	66
b		
-	healthy	sick
healthy	122	11
sick	11	73
c		
-	healthy	sick
healthy	124	6
sick	8	78

of our implementation by matching its result with the results published in other articles.

Since the tests of this research are performed in the form of fourfold cross validation, in Table 4, the evaluation parameters were presented separately for each fold of the networks tested. These results are obtained when networks were trained solely on frontal data. Two important points emerge from these results that determine how to continue our tests. First, among the tested networks, the results of different folds of VGG19 are quite similar, and therefore this architecture may be considered as a model with desired fitting compared to others. Another is that the average of fourfold results in VGG19 (i.e., final evaluation parameter) are significantly consistent with the published results of other studies.

After ensuring the accuracy of the implementations as described above, an attempt was made to improve the results by merging the images with lateral angles. At this stage, according to Table 5, the results obtained from DensNet201 show a significant advantage over other structures both in correctness and concentration.

As shown in the results, by adding lateral views, we slightly improved the results so that the specificity on the side was 1% and the sensitivity was 2% versus only-front view. This shows the effect of adding different views of thermograms in improving the results.

Table 3 confusion matrix in emerged frontal and lateral view **a.** VGG16 **b.** ResNet50 **c.** DenseNet201.

a		
-	healthy	sick
healthy	210	18
sick	27	204
b		
-	healthy	sick
healthy	223	13
sick	14	209
c		
-	healthy	sick
healthy	225	16
sick	12	206

5.3 Second scenario

In this scenario, the performance of the proposed method compared with either traditional feature based methods or deep learning based methods.

Consequently, based on which was demonstrated in both above tables, in the two main comparison tables (6 and 7) the results of VGG19 for front view scenario and DensNet201 for combinational scenario (frontal plus lateral views) have been compared and summarized with other deep and non-deep alternatives.

Table 6 shows comparison of our methods with some other deep learning-based studies. The results of some of these studies have been adapted from the articlebecause the combination of data in the training scenario and their testing were exactly the same as the research in this article. However, because the combination of training and test data in some of these studies was not as comprehensive as our research, for the fairness of the comparisons, we first implemented the method of that article and then experiments have been performed with the data structure which has been used in this article.

As shown in the Table 6, the evaluation parameters obtained by our method is higher than that those obtained by other methods. This superiority is at least 2 percent for both sensitivity and specificity of the proposed method against its closest alternative. In [21] which led to the best results among alternatives of deep learning family, 4 types of data augmentation have been used including horizontal and vertical flip, rotation between 0–45 degrees, 20% zoom, and normalized noise. However, the performance parameters obtained in this study are somewhat lower than those were obtained by proposed method.

As mentioned at the beginning of the article, an important group of methods for diagnosis of cancer in breast thermographic images are methods based on feature extraction and traditional classifiers. Thus, in the second scenario of comparisons, the results of the method of this paper were compared with the results of traditional methods based on manual extraction of features (i.e., non-family schemes). Most of these methods have analyzed left and right breast temperature by applying a segmentation method and extracting features such as texture, statistical, GLCM, Haralick, wavelet, HOG and etc.

Table 7 clearly demonstrates the superiority of the proposed method of this paper over all its alternatives in scenario of handcrafted features. It is observed that in comparison with the previous table, the precision parameter has been removed because in the references in this table, the value of this parameter has not been reported.

In [15], although a sensitivity of 2% more than our method is reported, but its specificity is less than the proposed method by extent of 10%. Of course, here, the explanations that we provided earlier about the importance of specificity over sensitivity in medical applications are true. Apart from this one case in other cases of Table 7, both the sensitivity and specificity of proposed method show superiority over its alternatives.

Table 4 result of three pre-trained network in only-front view.

fold	DenseNet121				EfficieNetB0			
	ACC	Prr	Sn	Sp	ACC	Pr	Sn	Sp
1	0.62	0.50	0.65	0.61	0.83	0.82	0.70	0.91
2	0.81	0.88	0.66	0.93	0.90	0.90	0.86	0.93
3	0.90	0.88	0.83	0.94	0.90	0.85	0.89	0.91
4	0.87	0.75	1	0.78	0.96	0.91	1	0.90
ave	0.80	0.75	0.78	0.81	0.90	0.87	0.86	0.91

VGG19			
ACC	Pr	Sn	Sp
0.90	0.94	0.80	0.97
0.94	0.91	0.91	0.93
0.92	0.81	1	0.88
0.96	0.95	0.95	0.96
0.93	**0.90**	**0.91**	**0.94**

Table 5 result of three pre-trained network in emerged frontal and lateral view.

fold	VGG16				ResNet50			
	ACC	Prr	Sn	Sp	ACC	Pr	Sn	Sp
1	0.85	0.86	0.82	0.88	0.94	0.93	0.96	0.92
2	0.89	0.85	0.92	0.87	0.92	0.91	0.89	0.93
3	0.93	0.89	1	0.87	0.96	0.96	0.96	0.96
4	0.92	0.91	0.92	0.91	0.92	0.92	0.92	0.93
ave	0.90	0.88	0.91	0.88	0.94	0.93	0.93	0.94

DenseNet201			
ACC	Pr	sn	Sp
0.92	1	0.84	1
0.95	0.94	0.96	0.95
0.93	0.94	0.93	0.94
0.93	0.89	0.98	0.90
0.93	**0.94**	**0.93**	**0.95**

Table 6 Comparison of our model with articles with deep learning-based feature extraction

Approach	ACC	Pr	Sn	Sp
Ours-only frontal	**0.93**	**0.90**	**0.91**	**0.94**
Ours-frontal and lateral	**0.93**	**0.94**	**0.93**	**0.95**
[20]	0.80	0.71	0.83	0.77
[21]	0.92	0.94	0.91	0.93
[36]	0.71	-	0.78	0.65

Table 7 Comparison of our model with articles with manually feature extraction

article	ACC	Sn	Sp
Ours-only frontal	**0.93**	**0.91**	**0.94**
Ours-frontal and lateral	**0.93**	**0.93**	**0.95**
[14]	0.85	0.87	0.83
[15]	0.90	0.95	0.85
[16]	0.88	0.80	0.93
[17]	0.90	0.87	0.93
[37]	0.91	0.93	0.90
[38]	0.89	0.86	0.90
[39]	0.88	0.87	0.89
[40]	0.91	0.87	0.94

6 Conclusion

This paper presents a new method for improving breast cancer diagnosis by using deep neural networks. The proposed scheme attempts to detect cancerous breasts by transfer learning-based models which are fitted by fusion of multiple views of thermographic images. The well known DMR database was utilized to train and test the fitted models. To evaluate the performance of the proposed algorithm, its results and those belong to the its several alternatives were compared in two scenarios in terms of sensitivity, specificity, accuracy and precision.

In the first scenario, the comparisons are performed in deep learning paradigm in such way that the proposed scheme is compared with some techniques that may be considered in its family. In contrast the second scenario compares the performance of the proposed scheme with methods which are based on handcrafted features and may be said that are not a member of its family.

In the scenario of comparing with the methods of deep learning field, it may be observed that the proposed scheme capable of distinguish healthy and cancerous tissue with better performance parameters than its alternatives. For instance, although the superiority of the sensitivity parameter of the proposed method over some of these alternatives reaches to 15%, even compared to the closest alternative, we see at least 2% superiority for both sensitivity and specificity.

In the second scenario, the superiority of the proposed method over alternatives which are based on handcrafted features is more significant, so that in all cases except for an alternative, both sensitivity and specificity of the proposed method is better than the existing methods. In the only exception, where the sensitivity of the proposed method is 2 percent less than one of the alternatives in this family, in contrast, its specificity is 10 percent higher, which practically makes the proposed method more reliable in diagnosing cancerous tissue.

Based on the better sensitivity and specificity of the proposed method compared to both family and non-family methods, it may be concluded that the proposed scheme can be developed and improved in future researches as a choice with acceptable practical potential in interpreting the breast thermographic images.

Declarations

Ethics approval Not applicable.

Consent for publication Yes.

Conflict of interest The co-authors declare no conflict of interests.

References

1. Mahmood T, et al. A Brief Survey on Breast Cancer Diagnostic with Deep Learning Schemes Using Multi-Image Modalities. Journals & Magazines. 2020;8:165779–809.
2. Mashekova A, et al. Early detection of the breast cancer using infrared technology – A comprehensive review. Therm Sci Eng Prog. 2022;27.
3. El Chamieh C, et al. Statistical methods for evaluating the fine needle aspiration cytology procedure in breast cancer diagnosis. BMC Med Res Methodol. 2022;22.
4. Husaini MASA, et al. A Systematic Review of Breast Cancer Detection Using Thermography and Neural Networks. IEEE Access. 2020;8:208922–208937.
5. Sánchez-Cauce R, et al. Multi-input convolutional neural network for breast cancer detection using thermal images and clinical data. Comput Methods Programs Biomed. 2021;204.
6. Houssein EH, et al. Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review. Expert Syst Appl. 2020;167.
7. Dihmani H, et al. A review on Suspicious-Regions Segmentation Methods in Breast Thermogram Image. 2nd Innovative Research in Applied Science, Engineering and Technology (IRASET), Conference on. 2022.
8. Mačianskytė D, et al. Automatic Detection of Human Maxillofacial Tumors by Using Thermal Imaging: A Preliminary Study. Sensors. 2022;22.
9. Singh D, et al. Role of image thermography in early breast cancer detection- Past, present and future. Comput Methods Programs Biomed. 2020;183.
10. Resmini R, et al. A hybrid methodology for breast screening and cancer diagnosis using thermography. Comp Biol Med. 2021;135.
11. Barbosa VAF, et al. Deep-wavelet neural networks for breast cancer early diagnosis using mammary termographies. Deep Learning for Data Analytics. 2020;99–124.
12. Al-antari MA, et al. Evaluation of Deep Learning Detection and Classification towards Computer-aided Diagnosis of Breast Lesions in Digital X-ray Mammograms. Comput Methods Programs Biomed. 2020;196.
13. Baduge SK, et al. Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications. Autom Constr. 2022;141.
14. Lessa V, et al. Applying Artificial Neural Network for the Classification of Breast Cancer Using Infrared Thermographic Images. International Conference on Computer Vision and Graphics. 2016;429–438.

15. Sourav P, et al. Texture Analysis of Breast Thermogram for Differentiation of Malignant and Benign Breast. International Conference on Advances in Computing, Communications and Informatics (ICACCI). 2016;8–14.

16. Gogoi UR, et al. Discriminative Feature Selection for Breast Abnormality Detection and Accurate Classification of Thermograms. International Conference on Innovations in Electronics, Signal Processing and Communication (IESC). 2017;39–44.

17. Sathish D, et al. Asymmetry analysis of breast thermograms using automated segmentation and texture features. SIViP. 2017;11:745–52.

18. Sanchez-Ruiz D, et al. Automatic region of interest segmentation for breast thermogram image classification. Pattern Recognit Lett. 2020;135:72–81.

19. Angkoso CV, et al. Multiplane ConVolutional Neural Network (Mp-CNN) for Alzheimer's Disease Classification. International Journal of Intelligent Engineering and Systems. 2022;15(1):329–40.

20. Farooq MA, et al. Infrared Imaging for Human Thermography and Breast Tumor Classification using Thermal Images. 31st Irish Signals and Systems Conference (ISSC). 2020.

21. Zuluaga-Gomez J, et al. A CNN-based methodology for breast cancer diagnosis using thermal images. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization. 2020.

22. Fernández-ovies FJ, et al. Detection of breast cancer using infrared thermography and deep neural networks detection of breast cancer using infrared thermography and deep neural networks, in Bio informatics and Biomedical Engineering. Berlin, Germany: Springer-Verlag; 2019.

23. Roslidar R, et al. BreaCNet: A high-accuracy breast thermogram classifier based on mobile convolutional neural network. Mathematical Bioscience and Engineering. 2021;19:1304–31.

24. Husaini MASA, etal. Thermal-based early breast cancer detection using inception V3, inception V4 and modified inception MV4. Neural Comput Appl. 2022.

25. Yamashita R, et al. Convolutional neural networks: an overview and application in radiology. Insights Imaging. 2018;9:611–29.

26. Mardi M, et al. Temporal Convolutional Learning: A New Sequence-based Structure to Promote the Performance of Convolutional Neural Networks in Recognizing P300 Signals. Journal of Health Management and Informatics. 2021;8(1):68–77.

27. Esraa AM, et al. Deep learning model for fully automated breast cancer detection system from thermograms. PLoS One. 2022:17(1).

28. Krishna ST, et al. Deep Learning and Transfer Learning Approaches for Image Classification. International Journal of Recent Technology and Engineering (IJRTE). 2019;7:427–32.

29. Simonyan K. Very deep convolutional networks for large-scale image recognition. arXiv. 2015.

30. Kaiming H, et al. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016:770–778.

31. Huang G, et al. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017;4700–4708.

32. Tan M, et al. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning. PMLR. 2019;97:6105–6114. http://visual.ic.uff.br/.

33. [online]: Available from: http://visual.ic.uff.br/.

34. Silva LF, et al. A New Database for Breast Research with Infrared Image. J Med Imaging Health Inform. 2014;4:92–100.

35. Rahman T, et al. Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection using Chest X-ray. Appl Sci. 2020;10(9):3233.

36. Baffa MFO, et al. Convolutional Neural Networks for Static and Dynamic Breast Infrared Imaging Classification. 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). 2018:174–181.

37. Karim CN, et al. A new approach for breast abnormality detection based on thermography. Med Technol J. 2018;2:257–65.

38. Pramanik S, et al. A Computer-Aided Hybrid Framework for Early Diagnosis of Breast Cancer. Adv Comp Syst Sec. 2019:111–124.

39. Pramanik S, et al. Suspicious-Region Segmentation From Breast Thermogram Using DLPE-Based Level Set Method. IEEE Trans Med Imaging. 2019;38:572–84.

40. Sathish D, et al. Role of normalization of breast thermogram images and automatic classification of breast cancer. Vis Comput. 2019;35:57–70.