1212: DEEP LEARNING TECHNIQUES FOR INFRARED IMAGE/VIDEO UNDERSTANDING

Screening of breast cancer from thermogram images by edge detection aided deep transfer learning model

Subhrajit Dey[1] · Rajarshi Roychoudhury[2] · Samir Malakar[3] 🔟 · Ram Sarkar[2]

Abstract

Breast cancer, the most common invasive cancer, causes deaths of thousands of women in the world every year. Early detection of the same is a remedy to lessen the death rate. Hence, screening of breast cancer in its early stage is utmost required. However, in the developing nations not many can afford the screening and detection procedures owing to its cost. Hence, an effective and less expensive way of detecting breast cancer is performed using thermography which, unlike other methods, can be used on women of various ages. To this end, we propose a computer aided breast cancer detection system that accepts thermal breast images to detect the same. Here, we use the pre-trained DenseNet121 model as a feature extractor to build a classifier for the said purpose. Before extracting features, we work on the original thermal breast images to get outputs using two edge detectors - Prewitt and Roberts. These two edge-maps along with the original image make the input to the DenseNet121 model as a 3-channel image. The thermal breast image dataset namely, Database for Mastology Research (DMR-IR) is used to evaluate performance of our model. We achieve the highest classification accuracy of 98.80% on the said database, which outperforms many state-of-the-art methods, thereby confirming the superiority of the proposed model. Source code of this work is available here: https://github.com/subro608/thermogram

✉ Samir Malakar
malakarsamir@gmail.com

Subhrajit Dey
subhrajitdey.agt@gmail.com

Rajarshi Roychoudhury
rroychoudhury2@gmail.com

Ram Sarkar
raamsarkar@gmail.com

[1] Department of Electrical Engineering, Jadavpur University, Kolkata, India

[2] Department of Computer Science and Engineering, Jadavpur University, Kolkata, India

[3] Department of Computer Science, Asutosh College, Kolkata, India

1 Introduction

The main reason of cancer is the uncontrolled cell division which takes place due to mutation of the genes in the cells. Breast cancer is a disease of the cancer family, caused by the genetic abnormalities that take place in the breast cells and poses high mortality risk.[1] Early detection of breast cancer can be mitigated its fatality with appropriate medical treatment. Hence, early screening plays a vital role in saving the lives of thousand of women affected by breast cancer.

Various types of methods are used for detection of breast cancer, some of those have been discussed by Borchartt et al. [4]. Diagnosis by using mammography is one of the most popularly used methods [16]. In mammography, breast images are captured using low energy X-rays and then diagnosed to detect the cancer. Rouhi et al. [29] and Gao et al. [15] have proposed mammograhy based breast cancer detection techniques, where convolutional neural network (CNN) models are employed. However, mammography is not a very accurate process and also the use of X-rays may damage the breast cells due to exposure to the radiation.[2] Therefore, researchers are trying the alternatives which include the use of other technologies such as ultrasound, thermography, magnetic resonance imaging (MRI), tomography etc. [2, 6, 26, 28]. Among these alternatives, thermography is found to be more effective in detecting breast cancer [2, 26, 28]. Screening using thermography follows the hypothesis: the temperature profile of the cancerous tissues is higher than the normal tissues. Besides, thermography does not use any external radiation like in the case of mammography. Therefore, it can be considered as a safer process in terms of tissue damage over the others. Moreover, it is an effective process on women of all ages even with breast implants.

To capture the thermogram images, digital infrared thermal imaging (DITI) is used [17, 20]. However, the problem is that in the diagnosis part, there is a need of a specialist doctor who can diagnose the images to predict the possibility of breast cancer. The major problem in this method is that in the developing countries like India, Bangladesh, Sri Lanka and Nepal there are not many people who can afford a specialized doctor for diagnosis. Besides, there is always a chance of human error, and therefore a specialist doctor sometimes suggests to take second opinion from another specialist. This scenario increases the diagnosis cost further. Hence, demand for an automatic system for diagnosis and screening of subjects having breast cancer is raised. In many of the research works, the researchers have used feature engineering for this task [2, 3, 31]. Considering the recent advancement of deep learning based models, researchers have started using such models for breast cancer detection [25] especially CNN based achitectures.

LeCun et al. [23] first introduced the CNN architecture which has proved to be effective in various image classification and object recognition problems over the years. However, the problem with CNN is that to achieve better classification accuracy, sometimes, we need heavy and complex architectures which require large amount of data and computational power to train the model. To resolve this, researchers have come up with the concept of

---

[1]https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html
[2]https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/limitations-of-mammograms.html

transfer learning. In this concept, pre-trained CNN models are used as feature extractors or sometimes, or these models fine-tuned using the target dataset. To have a competent pre-trained model, researchers use large and complex CNN models trained on big datasets like ImageNet which has approximately 20,000 categories of images. This concept has been used successfully in the past for various medical image analysis tasks like skin cancer detection [12], lung cancer detection [24], breast cancer detection from thermogram images [13], and COVID-19 detection from chest X-ray images [7, 9], as well as for various pattern classification problems like high resolution satellite image classification [5], fraction detection [22], and infrared pedestrian detection [18]

Keeping the above facts in mind, in this work, we first extract features using the pre-trained DenseNet121 model from thermal breast images and then employ a classifier on the extracted features to detect whether a subject has breast cancer or not. To extract the minute details like the blood vessels and deformation in the breast images, we use the edge information of the gray-scale thermal breast images and converted the 1-channel thermal images to 3-channel images. We look for the minute details in breast images because they might help in accurate detection of the breast cancer from thermal images.[3] To be specific, we incorporate texture analysis methods with these edge detectors from the breast thermogram images inspired by the works [10, 37] where edge detectors are used to improve the image based crack detection in concrete [10] and face recognition [37]. In this work, we use Roberts and Prewitt edge detection techniques as they preserve most of the minute edges. To perform the experiments, we use the DMR-IR dataset introduced in [33] that contains infrared breast images. In a nutshell, the main contributions of our work are as follows:

– Use of edge detectors for generating edge-prominent 3-channel breast images to be fed to a CNN model: two channels having images obtained after applying Roberts and Prewitt edge detection methods, and one being the original gray-scale image.
– Developed a classification system with the help of features extracted using the pre-trained DenseNet121 model.
– Compared the classification performance of our model with other pre-trained models like VGG19, VGG16, DenseNet169 and Xception.
– Performed Grad-CAM analysis on the classifier to better understand its working principle.
– Obtained state-of-the-art results by the proposed model despite having less number of trainable samples.

The rest of the paper is organized as follows: first in Section 2, we discuss some of the past methods on thermogram image based breast cancer detection developed by the other researchers. We explain working procedure of overall architecture of present method and some of its key components in Section 3 while in Section 4, we first illustrate performance of our model and then compare it with other models. Finally, in Section 5, we conclude our paper mentioning some possible future extension of the work.

2 Related work

Several research attempts have been made to develop breast cancer detection methods using thermography. Most of these attempts have used feature engineering techniques [1, 2, 14,

---

[3]https://thebreastthermographycenter.com/what-is-breast-thermography/

27, 28, 30–32, 35] while a few have used deep learning models [13, 36, 39] for breast cancer detection. In the work [14], the authors used Curvelet transform based feature extraction for breast cancer detection. In another work [32], K-means clustering method has been used by the authors for breast cancer detection using color features extracted from thermogram images. Discrete wavelet transform is used on segmented breast thermograms in the work [27] to compute initial feature point image from which features can be extracted. We also find the use of bio-inspired optimization techniques like Grey wolf optimizer, Particle Swarm Optimizer, Moth Flame optimizer and Firefly Algorithm optimizer for segmentation of thermograms images in the work [30].

Besides, in the work [21], authors have used the combination of thermography along with mammography and have shown that this combination enhances the sensitivity and specificity of classification as compared to when only mammography is used. There are authors who have used block variance which is a texture feature extraction method, on breast thermograms, [28]. Another use of thermogram images found in the work by Okuniewski et al. [26]. In this work the authors have used contour classification of breast thermograms to facilitate breast cancer detection. Acharya et al. [2] have first co-occurrence and run lengths based features from breast thermograms and then these features are fed to support vector machine (SVM) for classification. Use of representational learning and texture analysis methods on the breast thermograms is proposed in the [1]. In another work, Silva et al. [35] have first extracted region of interest (ROIs) from sequential Dynamic Infrared Thermography (DIT) images and then extract features from these ROIs to perform breast cancer detection using SVM.

In another work, the authors use a computer aided diagnosis (CAD) based breast cancer detection technique along with CNN, which uses thermal images as input data [39]. The authors of the work [36] have proposed a segmentation technique for thermal images of breasts using curvature function and gradient vector flow and then uses CNNs for the classification. Utilization of optimization algorithms for tuning CNNs for breast cancer detection can be found in the research works like [11], where the authors extract features from breast thermograms and then classify them using CNN optimized by Bayes algorithm. The concept of transfer learning for the detection of breast cancer from thermograms has been applied in works reported in [13].

3 Proposed method

In this work, we have proposed a method which uses the concept of transfer learning for the detection of breast cancer from thermogram images. Our method uses the DenseNet121 pre-trained model as a feature extractor and using it we have built a classifier for detection of breast cancer. The reason for using the DenseNet architecture is based on the fact that this architecture diminishes the vanishing gradient problem. Along with using the pre-trained DenseNet121 model, we have also used two edge detectors, Prewitt and Roberts, to extract edge information from the thermal breast images. We have then concatenated the outputs from these two edge detectors with the original gray-scale breast image to obtain an edge prominent 3-channel image. This is important because the pre-trained model can extract features from 3-channel images only. The overall architecture of the proposed work is shown in Fig. 1 and the associated modules are described in the following subsections.
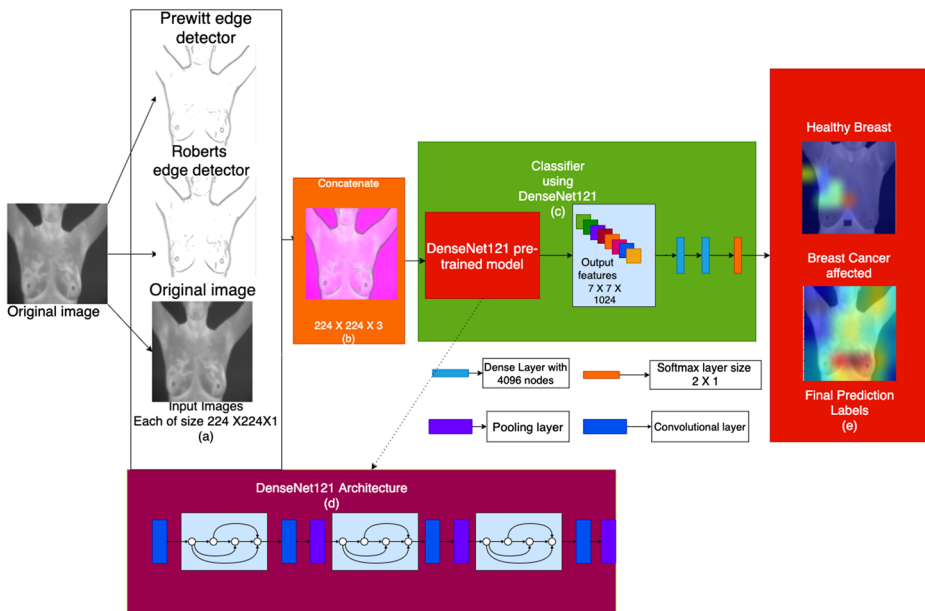
Fig. 1 The overall architecture of the proposed model. Here, block **a** shows an instance of original gray-scale thermal image along with two edge extracted images by Roberts and Prewitt methods while the output image after concatenating the original image with the outputs of the edge detectors is shown in block (**b**). The block **c** represents working of the classifier using pre-trained DenseNet121 model, **d** represents the DenseNet121 architecture with three Dense-Blocks and the transition layers between them, and **e** represents the final prediction labels. We have also shown the Grad-CAM images of the two classes for better understanding of the working of the classifier

3.1 Edge detectors

For breast cancer classification, it is important to look at the minute details like blood vessels and deformations on the breasts, in order to determine whether the patient has breast cancer or not and for that we have extracted edge information from the thermal images and combined these edge information with the original thermal images to make them more information rich. At this end, we have used two well-known edge detection techniques namely, Roberts and Prewitt to generate edges from original gray-scale thermogram images as these detectors help in preserving most of the minute edges. These two edge marked images are used with the original image to form a 3-channel image which is input to the pre-trained DenseNet121 model.

Roberts and Prewitt edge detection techniques are known as a gradient based edge detection method. In gradient based edge detection technique, we convolve the image with horizontal and vertical derivative masks. These masks are also known as horizontal and vertical operators. These operators are used to perform quantitative analysis of the change in pixel intensities that leads to the identification of edges. They determine the presence of edges by calculating the difference between the corresponding pixels of an image which is analogous to the derivative in the signal domain.

Let, Δ_x and Δ_y are the horizontal and vertical edge operators respectively which are convolved with a gray-scale image (say, I_g) and generate two gradient images (say, G_x and G_y) i.e.,

$$G_x = I_g * \Delta_x$$

and

$$G_y = I_g * \Delta_y$$

Here '*' is the convolutional operator. Please note that Δ_y is formed by 90° rotation of Δ_x and vice-versa. The magnitude of the gradient (say, G) is

$$G = \sqrt{G_x^2 + G_y^2}$$

The pixel coordinate notion helps us approximate the gradient calculation, where we use appropriate masks to calculate the gradient at a particular pixel. We obtain the final edges using a threshold value (say, th) which is set here as the mean of values appear in G i.e.,

$$th = \frac{1}{(MN)} \sum_{x=1}^{M} \sum_{y=1}^{N} G(x, y)$$

Finally an edge image (say, I_e) is obtained using following equation.

$$I_e(x, y) = \begin{cases} edge\ pixel & : G(x, y) < th \\ background\ pixel & : otherwise \end{cases}$$

Based on different edge operators used, there are different edge detection techniques. Here, we have used Roberts and Prewitt edge operators, described below, to extract the edges required in our model.

3.1.1 Roberts edge detector

The Roberts operators highlight the regions of high spatial frequency which has a high probability to denote the presence of edges. Roberts operators use masks of size 2×2, and the gradient calculations are lightweight. To determine the possibility of a pixel for being edge pixel only four neighbouring pixels of it are examined. Here, we have used Roberts cross operator as it is better choice than its horizontal or vertical version.

Here,

$$\Delta_x = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}$$

and

$$\Delta_y = \left(\begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix}\right)$$

There are no parameters to set. However, its main disadvantages are: i) it is very susceptible to noise as it uses a small sized masks, and ii) it does not perform well unless the edge is prominent. That is why we also use Prewitt edge detection technique. We show an example of edge detected image on applying Roberts cross operator based edge detection on a thermal breast image (see Fig. 2a) in Fig. 2c.

3.1.2 Prewitt edge detector

Prewitt edge detector uses masks of size 3×3. It is used for detecting two kinds of edges: horizontal and vertical. The advantage of using Prewitt edge detector is its simplicity. By
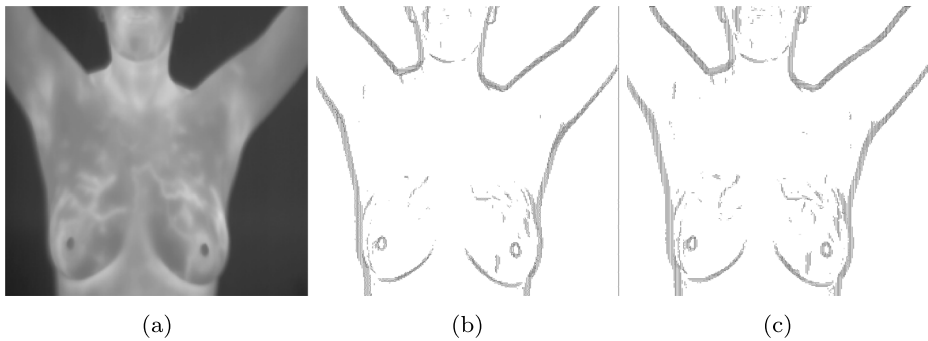
(a) (b) (c)

Fig. 2 The outputs of different edge detection techniques when applied on original thermal breast image shown in (**a**): **b** Prewitt operator and **c** Roberts cross operator

using this method, we get an approximation of the magnitude and also Prewitt edge detector can detect edges and their orientations. For Prewitts,

$$\Delta_x = \begin{pmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{pmatrix}$$

and

$$\Delta_y = \begin{pmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix}$$

We show an example of edge detected image on applying Prewitt edge detection technique on a thermal breast image (see Fig. 2a) in Fig. 2b.

3.2 DenseNet121 model

Using pre-trained models for achieving state-of-the-art image classification accuracy is gaining a lot of popularity nowadays. It becomes very useful when data in use and availability of computational power are limited. In our case, both constraints are applicable. That is why we use pre-trained DenseNet121 model as feature extractor to perform the present classification task.

In order to improve the information flow between layers, different connectivity patterns [19] are proposed in DenseNet. Here, each layer gets its input from the output of all the previous layers and so each layer gets a collective knowledge from all the previous layers. If the number of layers is n then in a traditional convolutional architecture, the number of connections will also be n but, in the case of DenseNet, this number is $n(n+1)/2$. Concatenation is used to combine all the outputs from previous layers to be fed to the present layer. Because of this dense connectivity, DenseNet architecture can reuse features which helps it in reducing the number of parameters required for training. As feature reutilization takes place in DenseNet, so it does need to train the redundant features as can be seen in other CNN architectures. Also, DenseNet facilitates information and gradients flow throughout the architecture. Because of this reason DenseNet is capable of handling the vanishing gradients problem. The problem that arises in concatenating outputs from different layers is their sizes and so to solve this DenseBlocks are used. In each of these DenseBlocks the feature maps are of same size but the number of filters used are different. There are layers

between these DenseBlocks which take care of the downsampling by using convolution of kernel size 1×1 and 2×2 average pooling, known as transition layers. These DenseBlocks together with the transition layers constitute the DenseNet architecture.

For training the thermogram breast images, we first resize the images to 224×224 and then applied the edge detectors- Roberts and Prewitt to extract edge information. For passing the images through the DenseNet121 pre-trained model for feature extraction, we use the edge images generated by the edge detectors and concatenate them with the corresponding original gray-scale images to generate the 3-channel image. We set the number of epochs to 50 and learning rate to 2e-4 as at this value the convergence is achieved faster, whereas for values greater than or less than this, the optimization may stuck in local optima.. We set these values after thorough experimentation.

For the optimizer, we use root mean square propagation (RMSProp) because in the case of other optimizers the convergence is rarely achieved and they normally get stuck in the local optima. We use ReLU as the activation function as suggested by the authors of DenseNet models [19]. After the features are extracted, we pass them through two fully connected (FC) layers with 4096 nodes in each. Any CNN model has two parts, one is the feature extraction part (here, we have used pre-trained weights to perform this) and the other is the classification part constituting a number of FC layers. In our case, after getting the features extracted by the pre-trained part we train those features on our own classification part. In VGG19 the number of nodes in FC layers is 4096 and in DenseNet121 it is 1024. We know that with the increase in number of nodes the trainable parameter increases and with the increased number of trainable parameters, a classification model can learn in a better way. That is why we chose the number of nodes as 4096 in the FC layers.. At the last, we add a softmax layer of size 2×1 to classify the input thermogram images. Here softmax layer is used for the binary classification because in the case of softmax increasing the output value of one class makes the others go down (sigma=1), which is exactly what we deal with in this breast cancer classification model. That is why instead of using sigmoid activation function, we use softmax activation function for this binary classification model. Also ReLU activation function is used in the previous layers because of which the output values of those layers are huge and when these values are passed though sigmoid then they skew its output. Based on the fact that for sigmoid outputs categorical-cross entropy is used as the loss function [38], we also use this loss function during the training phase of our model.

4 Result and analysis

In this work, we have designed a breast cancer detection technique from thermogram images of breasts utilizing the concept of transfer learning. The experiments are performed on DMR-IR dataset. The experiments are conducted to show the effectiveness of the edge detectors pair (i.e., Roberts and Prewitt) and the chosen DenseNet121 pre-trained model. In the first set of experiments, we have tried with different pairs of edge detectors along with the DenseNet121 pre-trained model whereas, in the second set of experiments, we have used the Robert and Prewitt edge detectors to compare performance of DenseNet121 model with the performances of other pre-trained models. Before performing these sets of primary experiments, we have done three additional experiments to check the effectiveness of pre-set learning rate and number of nodes in FC layers in the classifier designed here, and to decide the optimum number of epochs to be used during training. In this section, we have first described the dataset in use and then analyzed the experimental outcomes of the said

Table 1 The dataset distribution for training and evaluation of the proposed model

Dataset	# train samples		# test samples	
	Healthy	Affected	Healthy	Affected
DMR-IR dataset	300	580	80	160

experiments. Finally, we have provided Grad-CAM image analysis and comparative study with the state-of-the-art techniques.

4.1 Dataset

For helping in early breast cancer detection, a thermogram image dataset was prepared and made available to the research community by Silva et al. [33]. This dataset was formed by taking 20 sequential images at 15 seconds interval in between. The images were taken when the breast temperature was same as environment. Therefore, cooling the breast by using air stream before capturing images was performed. The dataset, downloaded from the link provided in [8], is pre-divided into train and test sets. The train set contains thermograms of 29 patients having breast cancer and 15 healthy cases. In the test set, there are thermogram records of 4 healthy subjects and 8 breast cancer patients. For each case, 20 breast thermograms along with their ROIs images are there in the dataset. The distribution of the dataset used in this work for training and evaluation of proposed model is shown in the Table 1.

4.2 Ablation study

We have already mentioned in Section 3.2 that we have set learning rate as 2e-4 and number of nodes in FC layers as 4096. In this section we have made ablation study to ensure about the effectiveness of these selections as well as to search presence of any alternative(s). For experimentation, we have used 3-channel breast thermogram images, constructed using edge images generated using Roberts and Prewitt, and original gray-scale image. The results of ablation study with varying learning rate and number of nodes in the FC layers are shown in Figs. 4 and 3 respectively. In both the cases, we have used pre-trained DenseNet121 model for extracting features from the edge prominent 3-channel breast images and set the number of epochs to 50. These results indicate the superiority of the present choices of learning rate and number of nodes in the FC layers and hence, we have continued with these values for rest of the experiments (Fig. 4).

Additionally, to decide an optimum number of epochs that should be used to train the classifier for rest of the experiments. For this experiment, we have used six pre-trained models- VGG16, VGG19, DenseNet121, DenseNet169, Inception and Xception to extract features from 3-channel edge-prominent breast images. Figure 5 shows the experimental outcomes. From these results, we can see that by increasing the number of epochs beyond 50 leads to overfitting in the models. Hence, we set number of epochs to 50 for rest of the experiments.

4.3 Performance analysis of edge detector pairs

The results, after using Roberts and Prewitt with the original image to generate a 3-channel image and feeding it into DenseNet121 are shown in the Table 2. Along with it, we have also used the Sobel and Canny edge detector techniques in combination with Roberts and
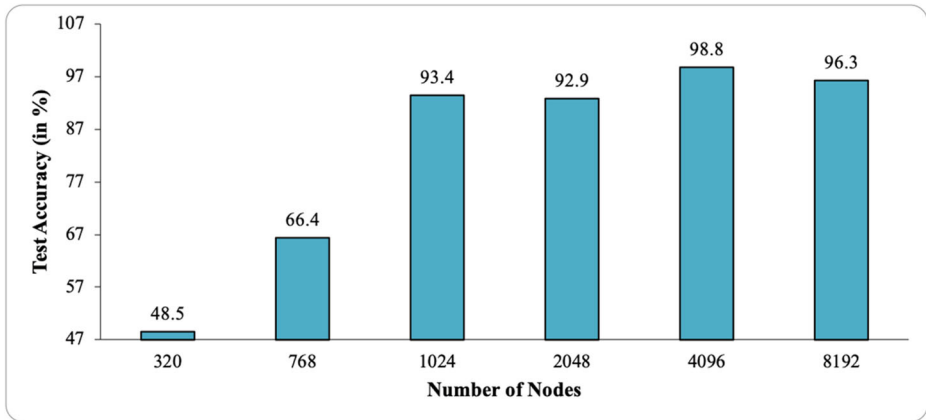
Fig. 3 Results of the ablation study with varying number of nodes in the FC layers

Prewitt. These results are shown in Table 2 for the combinations of Sobel and Roberts (SR), Sobel and Prewitt (SP), Sobel and Canny (SC), Prewitt and Canny (PC), and Roberts and Canny (RC) edge detectors. It is to be noted that in each combination we have concatenated original image with the edge images generated using the corresponding edge detectors to form 3-channel edge-prominent breast images. In the table, we have shown the precision, recall, F1-score and accuracy of the classifiers using pre-trained model of DenseNet121 as feature extractor. We can observe from the tables that use of PR combination yields better results with an accuracy of 98.80% which is better than that of other five combinations (i.e., SP, SR, SC, RC and PC). The SR and SP combination provide 97.50% and 96.70% recognition accuracy respectively which means use of edge image generated using Sobel edge detector to form 3-channel edge-prominent image generates lower accuracy than while using Robert and Prewitt. The reason for the poor performance of combinations with Sobel edge detector can be attributed to the signal to noise ratio. We know with increase in noise levels, the gradient magnitude of the edges starts giving inaccurate values which in turn
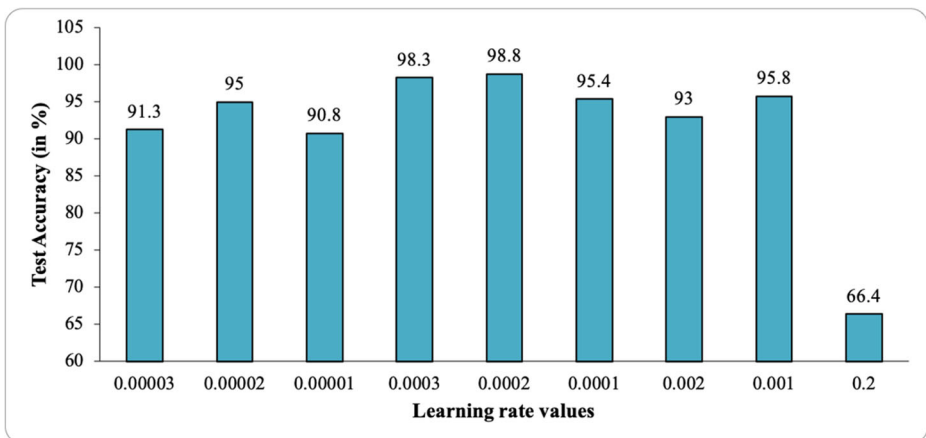

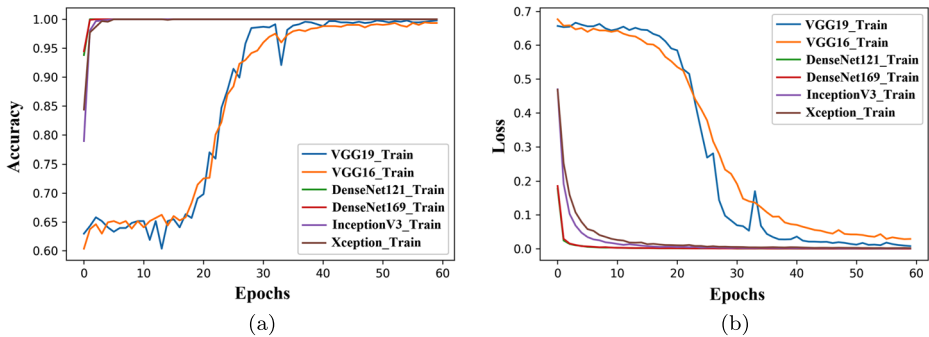**Fig. 4** Ablation study results with varying learning rate

Fig. 5 Effect on training accuracy and loss with respect to epochs: **a** represent accuracy vs. epochs while **b** shows loss vs. epochs curve. Here, "accuracy" denotes training accuracy

degrades the edge detection performance. The classification accuracies for the combinations including Canny edge detector i.e., SC, RC and PC are 92.90%, 93.80% and 96.70% respectively. The relatively poor performance of combinations with canny can be attributed to the loss minute edges present in the thermogram images during the use of the Gaussian filter and non-maximum suppression or adding some extra pixels as edge pixels while performing edge linking by hysteresis. In Fig. 6 we have shown the comparative performance scores obtained using different ways of making 3-channel image. In this figure, seven terms are used: "Original3", "Original+SR", "Original+SP", "Original+PR", "Original+SC", "Original+RC" and "Original+PC". "Original3" means original thermal breast image concatenated with itself twice, the others mean original breast image is concatenated with outputs from corresponding edge detector pair. From these results, we can see that the PR combination gives the highest accuracy among the combinations while the SC combination gives the lowest. Therefore, we can safely comment that our idea of generating the edge-prominent image using PR combination is justified.

4.4 Comparison with other pre-trained models

We have shown the average precision, recall, F1-score and accuracy of different models formed by using pre-trained models - VGG19, VGG16, DenseNet121, DenseNet169 and Xception as feature extractor in Fig. 7. This has been done to compare the performance of DenseNet121 model with other state-of-the-art CNN models when all are used as a feature extractor. For comparison purpose in case of all the models, we have considered the edge detector combination of Prewitt and Roberts. After comparison of the results, it can be concluded that our proposed method gives the highest accuracy score of 98.80%. Also in terms of average precision, our model has the highest average precision of 99.00% compared to all the other models. We have also plotted the confusion matrix of this classifier as shown in Fig. 8, where it can be seen that the classifier predicts all the breast cancer cases correctly, while it incorrectly classifies 3 healthy breast images as the cancerous.

4.5 Grad-CAM analysis

Grad-CAM actually means gradient weighted class activation map which helps us to understand the working of a deep learning model (here in our case a classifier's internal working process). It does this by showing the regions of the input images on which the model

Table 2 Class-wise performance of the classifier using pre-trained DenseNet121 as a feature extractor and using combinations of edge detectors on the input breast images

Methods used	Thermal Image Class	Precision (in %)	Recall (in %)	F1-score (in %)	Accuracy (in %)
Prewitt and Roberts	Affected Breast	98.00	100.00	99.00	98.80
	Healthy Breast	100.00	96.00	98.00	
	Average	99.00	98.00	98.50	
Sobel and Roberts	Affected Breast	96.00	100.00	98.00	97.50
	Healthy Breast	100.00	93.00	96.00	
	Average	98.00	96.50	97.00	
Sobel and Prewitt	Affected Breast	95.00	100.00	98.00	96.70
	Healthy Breast	100.00	90.00	95.00	
	Average	97.50	95.00	96.50	
Sobel and Canny	Affected Breast	90.00	100.00	95.00	92.90
	Healthy Breast	100.00	79.00	88.00	
	Average	95.00	89.50	91.50	
Prewitt and Canny	Affected Breast	95.00	100.00	98.00	96.70
	Healthy Breast	100.00	90.00	95.00	
	Average	97.50	95.00	96.50	
Roberts and Canny	Affected Breast	91.00	100.00	96.00	93.80
	Healthy Breast	100.00	81.00	90.00	
	Average	95.50	90.50	93.00	

focuses on. In the same sense, we have also performed Grad-CAM analysis on the classifiers we have used which can be seen from the Figs. 9 and 10. In Fig. 9, we can see that the Grad-CAM analysis of breast images fed into the classifier using pre-trained model of
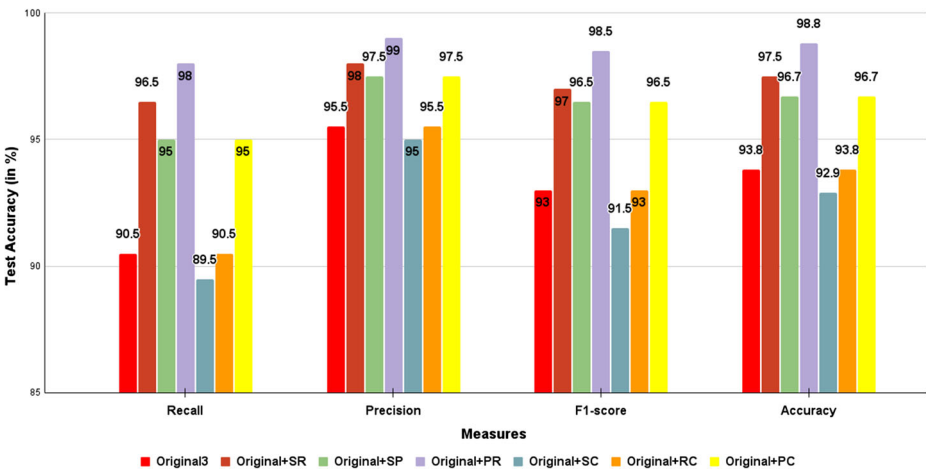

Fig. 6 Performance comparison of different 3-channel image preparation techniques. The classifier in use is built on the pre-trained model of DenseNet121 as feature extractor
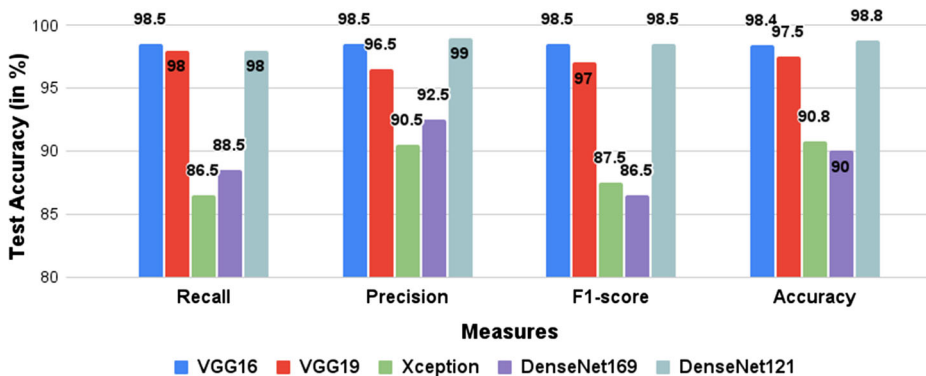
Fig. 7 Performance Comparison pf the present classifier when different pre-trained models are used to extract features from 3-channel edge-prominent images. Here we use the edge detector combination of Prewitt and Roberts

DenseNet121 for detection of breast cancer. It is confirmed from the figure that our classifier focuses on the regions which are very important for detection of breast cancer. From the figure it can be seen that in the case of the original healthy breast image, there is a small square which accidentally got captured in the image, still our model has the ability to ignore it. Same is the case in Fig. 10, where the Grad-CAM analysis of the healthy breast images is done, which are fed into the classifier using pre-trained VGG19 model as feature extractor.

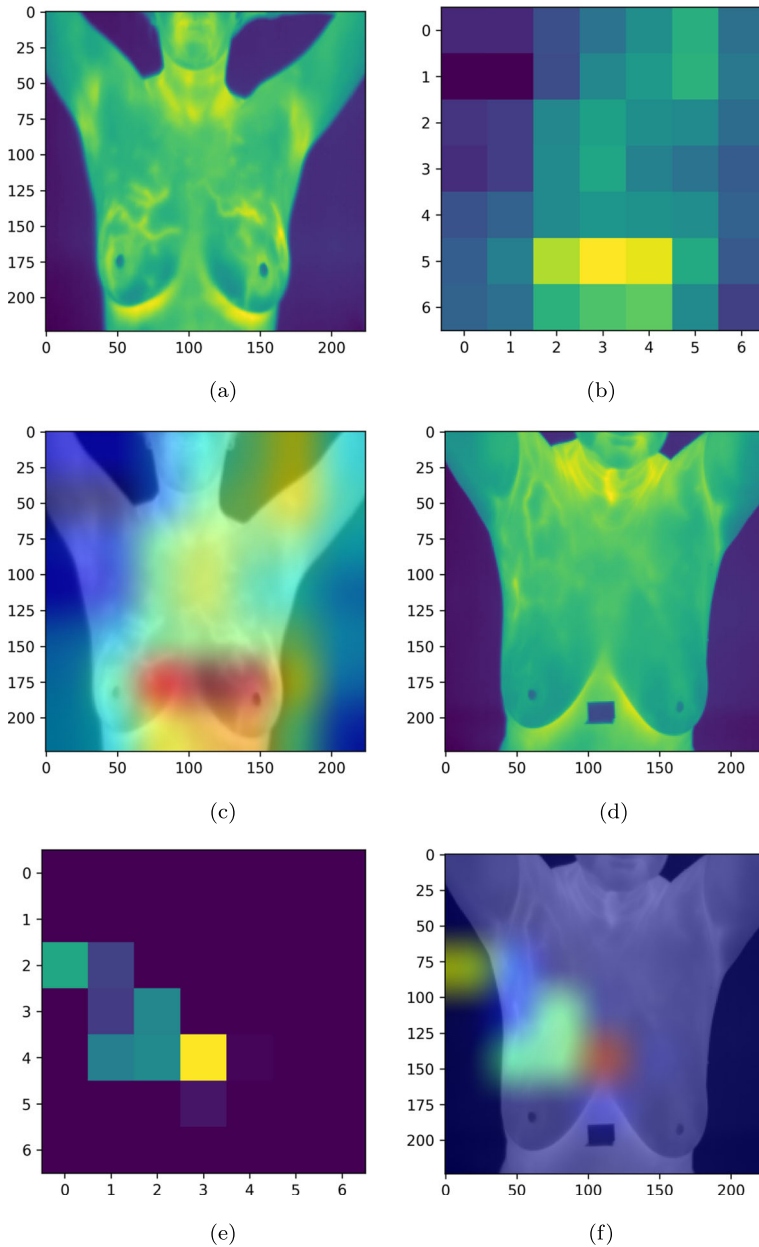**Fig. 8** The confusion matrix of the proposed classification method

Fig. 9 Grad-CAM images along with original breast images for the classifier using DenseNet121 pre-trained model. **a**, **d** represent the original breast images having breast cancer, **b**, **e** represent the heatmaps showing the focused regions of the classifier for images of (**a**) and (**d**) respectively, **c**, **f** represent the Grad-CAM images of (**a**), (**d**) respectively
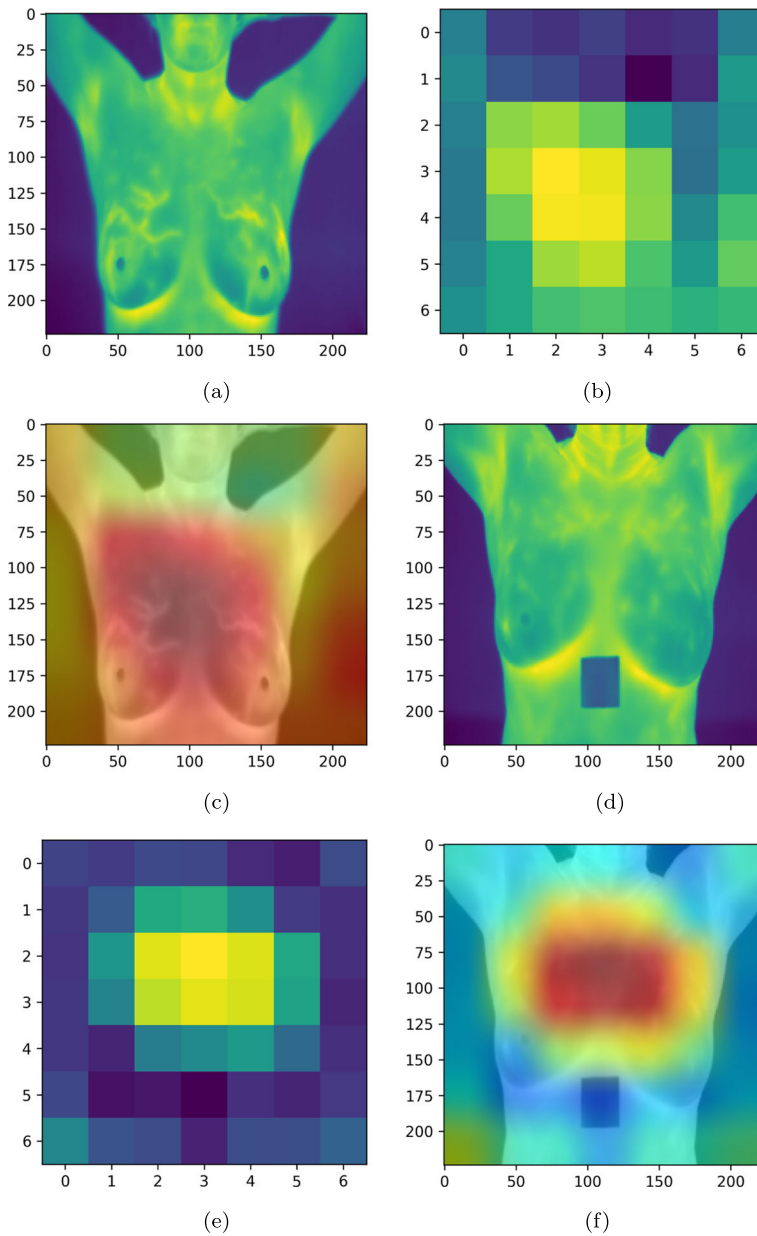
Fig. 10 Grad-CAM images along with original breast images for the classifier using VGG19 pre-trained model. **a**, **d** represent the original breast image having breast cancer, **b**, **e** represent the heatmap showing the focused regions of the classifier for images of (**a**) and (**d**) respectively, **c**, **f** represent the Grad-CAM image of (**a**), (**d**) respectively
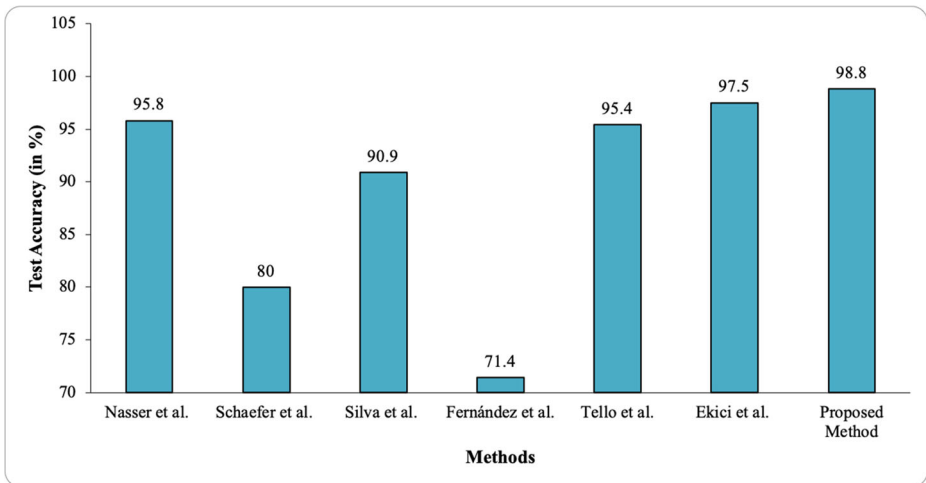
Fig. 11 Comparison of our proposed method with other state-of-the-art methods (Nasser et al. [1], Schaefer et al. [31], Silva et al. [34], Fernandez et al. [13], Tello et al. [36] and Ekici et al. [11]) used for breast cancer classification

4.6 Comparison with state-of-the-art models

In this section, we have compared the performance of present model with the performances of some state-of-the-art methods, proposed by Nasser et al. [1], Schaefer et al. [31] and Silva et al. [34], Fernandez et al. [13], Tello et al. [36] and Ekici et al. [11]) for breast cancer classification using thermogram images. In the work [1], authors have used 56 patients' thermogram images from DMR-IR dataset while the authors of the work [34] have used samples of 22 patients from the same dataset to conduct their experiments. However, Schaefer et al. [31] have used 146 patients' thermogram images from a private dataset. In another work, 500 thermograms of healthy and sick patients each, were used by the authors of [13] for breast cancer classification using the DMR-IR dataset. Also in the work [11] breast thermograms of 140 patients from the same dataset were used for classification. Whereas, in [36], the authors have used 63 thermograms from the dataset [33] for segmentation and breast cancer classification. However, to make the comparison uniform all the models are evaluated on the test set of the present dataset. The comparative results are shown in Fig. 11. From these results, it can be inferred that our proposed method gives the best classification result among the other models with an accuracy of 98.80%.

5 Conclusion

To reduce the high mortality of women having breast cancer, researchers all around the world are now trying to develop ways of screening them at the early stages. However, such early breast cancer screening methods should be accurate and cost effective. To this end, deep learning based breast cancer detection using thermography has become popular among the researchers. Following this research trend, we have proposed a method which uses the pre-trained model of DenseNet121 as a feature extractor to build a classifier for breast cancer detection. In doing so, we have converted the input gray-scale thermal breast images into

3 channel edge-prominent input images by using two edge detectors namely, Roberts and Prewitt. After performing a number of experiments, we have compared the performance of our model with other classifiers which also use pre-trained models like VGG16, VGG19, DenseNet169 and Xception, and shown that our model has outperformed all the other models. Besides, it is to be noted that the proposed method even though has been trained on a very small dataset still achieves state-of-the-art classification accuracy.

Although the present work performs well in comparison with state-of-the-art methods yet the results can be improved to make it useful in the practical field, as a small mistake can result in wrong diagnosis or even the death of the patient. Therefore, the system needs to be 100% accurate. We can use data augmentation techniques or GAN based synthesized data preparation to improve the performance further. Also, the DMR-IR dataset suffers the class imbalance problem which is true for most of the medical domain datasets. Hence, in future some class imbalance handling techniques may be considered to increase the model performance for minority classes also. Besides, use of some feature selection techniques on the extracted features may improve the results further.

Declarations

Conflict of Interest　Authors declare that there is no conflict of interest.

References

1. Abdel-Nasser M, Moreno A, Puig D (2019) Breast cancer detection in thermal infrared images using representation learning and texture analysis methods. Electronics 8(1):100
2. Acharya UR, Ng EYK, Tan JH, Sree SV (2012) Thermography based breast cancer detection using texture features and support vector machine. J Med Sys 36(3):1503–1510
3. Araújo MC, Lima RC, De Souza RM (2014) Interval symbolic feature extraction for thermography breast cancer detection. Expert Syst Appl 41(15):6728–6737
4. Borchartt TB, Conci A, Lima RC, Resmini R, Sanchez A (2013) Breast thermography from an image processing viewpoint: a survey. Signal Process 93(10):2785–2803
5. Chaib S, Yao H, Gu Y, Amrani M (2017) Deep feature extraction and combination for remote sensing image classification based on pre-trained cnn models. In: Ninth international conference on digital image processing (ICDIP 2017), vol 10420. International Society for Optics and Photonics, p 104203D
6. Cho N, Han W, Han BK, Bae MS, Ko ES, Nam SJ, Chae EY, Lee JW, Kim SH, Kang BJ, et al. (2017) Breast cancer screening with mammography plus ultrasonography or magnetic resonance imaging in women 50 years or younger at diagnosis and treated with breast conservation therapy. JAMA Oncol 3(11):1495–1502
7. Das S, Roy SD, Malakar S, Velásquez JD, Sarkar R (2021) Bi-level prediction model for screening covid-19 patients using chest x-ray images. Big Data Res 25:100233
8. Deepak A (2020) Thermal images for breast cancer diagnosis dmr-ir. https://www.kaggle.com/asdeepak/thermal-images-for-breast-cancer-diagnosis-dmrir
9. Dey S, Bhattacharya R, Malakar S, Mirjalili S, Sarkar R (2021) Choquet fuzzy integral-based classifier ensemble technique for covid-19 detection. Comput Biol Med :104585
10. Dorafshan S, Thomas RJ, Maguire M (2018) Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. Constr Build Mater 186:1031–1045
11. Ekici S, Jawzal H (2020) Breast cancer diagnosis using thermography and convolutional neural networks. Med Hypotheses 137:109542
12. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639):115–118

13. Fernández-Ovies FJ, Alférez-Baquero ES, de Andrés-Galiana EJ, Cernea A, Fernández-Muñiz Z, Fernández-Martínez JL (2019) Detection of breast cancer using infrared thermography and deep neural networks. In: International work-conference on bioinformatics and biomedical engineering. Springer, pp 514–523

14. Francis SV, Sasikala M, Saranya S (2014) Detection of breast abnormality from thermograms using curvelet transform based feature extraction. J Med Sys 38(4):1–9

15. Gao F, Wu T, Li J, Zheng B, Ruan L, Shang D, Patel B (2018) Sd-cnn: a shallow-deep cnn for improved breast cancer diagnosis. Comput Med Imaging Graph 70:53–62

16. Hela B, Hela M, Kamel H, Sana B, Najla M (2013) Breast cancer detection: A review on mammograms analysis techniques. In: 10th international multi-conferences on systems, signals & devices 2013 (SSD13). IEEE, pp 1–6

17. Herry CL, Frize M (2004) Quantitative assessment of pain-related thermal dysfunction through clinical digital infrared thermal imaging. Biomed Eng Online 3(1):19

18. Hu J, Zhao Y, Zhang X (2020) Application of transfer learning in infrared pedestrian detection. In: 2020 IEEE 5Th international conference on image, vision and computing (ICIVC). IEEE, pp 1–4

19. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708

20. Jones BF, Plassmann P (2002) Digital infrared thermal imaging of human skin. IEEE Eng Med Biol Mag 21(6):41–48

21. Kennedy DA, Lee T, Seely D (2009) A comparative review of thermography as a breast cancer screening technique. Integrat Cancer Therap 8(1):9–16

22. Kim D, MacKinnon T (2018) Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. Clin Radiol 73(5):439–445

23. LeCun Y, Haffner P, Bottou L, Bengio Y (1999) Object recognition with gradient-based learning. In: Shape, contour and grouping in computer vision. Springer, pp 319–345

24. Lin CJ, Jeng SY, Chen MK (2020) Using 2d cnn with taguchi parametric optimization for lung cancer recognition from ct images. Appl Sci 10(7):2591

25. Mambou SJ, Maresova P, Krejcar O, Selamat A, Kuca K (2018) Breast cancer detection using infrared thermal imaging and a deep learning model. Sensors 18(9):2799

26. Okuniewski R, Nowak RM, Cichosz P, Jagodziński D, Matysiewicz M, Neumann Ł, Oleszkiewicz W (2016) Contour classification in thermographic images for detection of breast cancer. In: Photonics applications in astronomy, communications, industry, and high-energy physics experiments 2016, vol 10031. International Society for Optics and Photonics, p 100312V

27. Pramanik S, Bhattacharjee D, Nasipuri M (2015) Wavelet based thermogram analysis for breast cancer detection. In: 2015 International symposium on advanced computing and communication (ISACC). IEEE, pp 205–212

28. Pramanik S, Bhattacharjee D, Nasipuri M (2016) Texture analysis of breast thermogram for differentiation of malignant and benign breast. In: 2016 International conference on advances in computing, communications and informatics (ICACCI). IEEE, pp 8–14

29. Rouhi R, Jafari M, Kasaei S, Keshavarzian P (2015) Benign and malignant breast tumors classification based on region growing and cnn segmentation. Expert Syst Appl 42(3):990–1002

30. Sayed GI, Soliman M, Hassanien AE (2016) Bio-inspired swarm techniques for thermogram breast cancer detection. In: Medical imaging in clinical applications. Springer, pp 487–506

31. Schaefer G, Závišek M, Nakashima T (2009) Thermography based breast cancer analysis using statistical features and fuzzy classification. Pattern Recogn 42(6):1133–1137

32. Shahari S, Wakankar A (2015) Color analysis of thermograms for breast cancer detection. In: 2015 international conference on industrial instrumentation and control (ICIC). IEEE, pp 1577–1581

33. Silva L, Saade D, Sequeiros G, Silva A, Paiva A, Bravo R, Conci A (2014) A new database for breast research with infrared image. J Med Imag Health Inf 4(1):92–100

34. Silva LF, Sequeiros GO, Santos MLO, Fontes CA, Muchaluat-Saade DC, Conci A (2015) Thermal signal analysis for breast cancer risk verification. In: MedInfo, pp 746–750

35. Silva TAEd, Silva LFd, Muchaluat-Saade DC, Conci A (2020) A computational method to assist the diagnosis of breast disease using dynamic thermography. Sensors 20(14):3866

36. Tello-Mijares S, Woo F, Flores F (2019) Breast cancer identification via thermography image segmentation with a gradient vector flow and a convolutional neural network. J Healthcare Eng :2019

37. Vijayarani S, Vinupriya M (2013) Performance analysis of canny and sobel edge detection algorithms in image mining. Int J Innov Res Comput Commun Eng 1(8):1760–1767

38. Zhang Z, Sabuncu MR (2018) Generalized cross entropy loss for training deep neural networks with noisy labels. In: 32nd conference on neural information processing systems (NeurIPS)

39. Zuluaga-Gomez J, Al Masry Z, Benaggoune K, Meraghni S, Zerhouni N (2020) A cnn-based methodology for breast cancer diagnosis using thermal images. Comput Methods Biomech Biomed Eng Imag Vis :1–15