ORIGINAL ARTICLE

Deep-CNN for Disease Classification Using Enhanced Mammographic Images

Ramesh Vaishya * (iD) **, Praveen Kumar Shukla**

Department of Computer Science and Engineering, Babu Banarasi Das University, Lucknow, India

*Corresponding Author: Ramesh Vaishya
Email: ramesh.rv.lko@gmail.com

Abstract

Purpose: Breast cancer has become one of the most common diseases that women face today as a result of poor nutrition and other environmental factors. A mammogram image of the breast will help detect breast cancer, but still, sometimes doctors and radiologists are unable to detect it due to poor image quality or abnormal region that appears to be normal.

Materials and Methods: In this paper, a deep CNN-based classification model is proposed that classifies the mammogram image as normal, masses, and micro-calcification. Firstly, the PSNR values of the mammogram images is improved using a median filter with the Local Contrast Modification (LCM) method. It is further enhanced by Adaptive-CLAHE in con junction with the Wiener filter. After image enhancement, the region of interest is segmented through morphological feature extraction and the Otsu thresholding method.

Results: In order to increase the number of samples in the mammogram image dataset, image data augmentation is applied to segmented images.

Conclusion: Finally, a pre-trained ResNet model is used for the classification of mammogram images. The proposed model has shown improved PSNR for mammogram images and achieved a higher classification accuracy of 98.91%, thus outperforming other existing methods. Additionally, the explainability and causality of the proposed model are also discussed to show the learning process of the model.

Keywords: Mammogram Image; Breast Cancer; Micro Calcification; Transfer Learning.

1. Introduction

The breast is a dynamic organ, with histological variations overlaid by long-term changes in aging processes throughout a menstrual cycle [1]. Breast cancer is a disease that gets worse over time and can become a serious malignant tumor if it is not treated early [2]. According to the amount of calcium deposited, the size of calcifications can be classified as macro-calcifications (bigger deposits) or micro-calcifications (smaller deposits). The mammography images belonging to normal, mass, and micro-calcification class are shown in Figure 1.
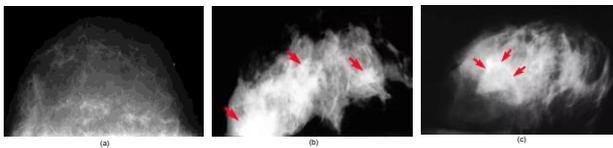
Figure 1. Mammography images of a) normal, b) mass and c) calcification classes [3]

The earlier stages of breast cancer are referred to as micro-calcification, which is very important to detect correctly. Various Computer-Aided Diagnostic (CAD) systems have been employed for breast cancer detection in the past. However, in the last few years, deep learning-based classification methods have evolved for cancer detection using mammogram images because they are more accurate than traditional and self-diagnosed methods [4].

Apart from mammogram images, ultrasonography, Computerized Tomography (CT) scans, Magnetic Resonance Imaging (MRI) can also be used for breast cancer detection.

Ultrasonography works over the fundamental principle of the reflection of ultrasound from the breast [5]. It has been demonstrated to be a reliable method of distinguishing solids from cyst lesions. However, this technique is limited by its low sensitivity and specificity. The utility of ultrasound in detecting individuals with thick breasts who are predisposed to breast cancer is also being researched. MRI is the imaging approach that does not use ionizing radiation. It can be used to detect breast cancer, but is unable to classify abnormalities such as biopsies and is extremely expensive. Among these, mammogram images are widely used as they utilize low-dose X-rays and have high sensitivity and specificity in detecting small tumors or micro-calcifications [6]. It is determined by the absorption of X-rays by various breast components such as blood vessels, fats, and ligaments, among others [7].

On a mammogram, the different grey levels represent the densities of the various tissues being imaged. A typical mammogram will reveal fibro-glandular tissue as well as vascular converging patterns. If the regular pattern deviates, it is usually regarded with suspicion and thoroughly investigated. In the present scenario, mammography is a non-invasive imaging modality that is the most effective method for detecting breast cancer early. However, one of the major challenges of interpreting screening mammograms is the low sensitivity of screening mammography ionizing radiation [8].

This paper proposes a model for detecting and classifying cancer in mammogram images exploiting the advantages of transfer learning using a pre-trained Deep Convolutional Neural Network (DCNN). The proposed method takes mammogram images and performs three steps such as image enhancement, segmentation, and classification. The image enhancement step improves the image contrast and the efficiency of segmentation algorithms [9]. Using segmentation, the micro-calcification which looks round and bigger or represented as tiny dots around the breast region can be easily located [10]. Further, to better accommodate the intraclass variations among images of the same class, various transformations are applied to segmented images. Thus, data augmentation is applied to segmented images to prepare a robust dataset. After image augmentation, an architecture i.e., RESNET-18 is utilized for the classification of mammogram images [11]. It provides faster training and testing compared to a custom-built CNN since it is pre-trained on millions of images and has existing learning for image patterns. In order to explain each particular prediction, a discussion on the interpretability of black box machine learning models is also included. The Local Interpretable Model-Agostic Explanations (LIME) [12] method is utilized to demonstrate how the ResNet model learns the intrinsic characteristics of the mammogram images. Thus, the salient contributions of the proposed model for classifying cancer in mammogram images can be summarized as follows.

• The proposed method uses a segmentation block that enhances the accuracy for detecting micro-calcification that reports significant variations in their patterns.

• Data augmentation performed on training images provides bias-free learning of the proposed model.

• The explainability of the proposed model is demonstrated which ensures the trustworthiness of the model.

The rest of the paper is organized as follows. Sec. II explains about existing methods of mammographic image classification. The proposed methodology including different blocks of image enhancement, segmentation, augmentation, and the learning method is described in Sec. III. A detailed result demonstration, explainability of the proposed architecture, and comparative analysis are presented in Sec. IV. Finally, the conclusion and future scope of the proposed work are discussed in Sec. V.

1.1. Previous Work

There is a vast literature on the classification of mammogram images. Shelda Mohan presented a histogram-based contrast enhancement approach for mammography images [13, 14]. The author used an entropy and edge detection method based on traditional Contrast Limited Adaptive Histogram Equalization (CLAHE). A stack-based machine learning classifier for the detection of micro-calcification in a mammography image is proposed by Alam *et al.* [15]. Initially, the morphological operations are performed for better feature extraction and contrast. The authors tested the proposed method on various methods on the MIAS mammogram dataset and achieved an accuracy of around 85%.

Al-juboori proposed a strategy for improving the contrast of mammography images [16]. The author used the combination of HE, and CLAHE along with morphological operations. Combining all the methods is used for contrast enhancement. Also, the retinex method was applied to images for sharpness and better clarity. Akila *et al.* had done a comparative study on various existing direct or indirect contrast enhancement methods for mammogram images [17]. The goal of this work is to identify the most efficient

method for contrast enhancement of the original image. MIAS dataset is used for comparative study and various existing methods have shown the best results for the PSNR value.

Tot *et al.* provided a review of mammogram micro-calcification detection based on clinical value [18]. In earlier stages of breast cancer, the cancer is difficult to detect and even doctors also cannot predict and locate cancer in the beginning stages. This results in the loss of a patient life. The author analyzed existing approaches and determined that around 10- 12% of breast abnormalities are duct-centric, making detection challenging. Kamra *et al.* provided an overview of ROI texture classification using an SVM classifier [19]. For ground truth ROIs, the proposed strategy employs the public domain method of Digital Database for Screening Mammogram (DDSM) with an accuracy of 95.34% and Mammographic Image Analysis Society Image (MIAS) with an accuracy of 92.34% in AD versus non-AD (normal) patients.

Quinlan *et al.* created a fully automated computer-based method for detecting and characterizing micro-calcification clusters in digital mammograms [20]. Their approach consisted of three steps: cluster identification, feature extraction, and classification. In the final stage, they used a rule-based system, an ANN, and a Support Vector Machine (SVM). They started with a pre-processing step to remove unusable radiological marks and the image's background. Later, they attempted to reveal hidden micro-calcifications using background correction and contrast enhancement. They estimated numerous discriminative morphological and textural features for all of the objects and clusters to use as input to the false positive reduction method, and they included four new rule-based features. Those who then identified the critical properties of each cluster using feature extraction, and then classified the abnormal regions as benign or malignant using classification algorithms. Even though their method produced satisfactory results when compared to the existing automated methods in the literature, more research with larger datasets is required.

Yousef *et al.* presented temporal processes based on automated breast micro-calcification detection [21]. Machine learning is used with a temporal reduction in this work to boost accuracy and detection rate. The proposed method is fully image-based, with source

and result pictures processed and classified as bulk or micro-calcification using an SVM.

Zou *et al.* work on image noise removal to eliminate impulse noise [22]. After detecting corrupted or noisy pixels, a healing algorithm begins to correct the pixels. With impulse noise, bidirectional decomposing pixels are easily identified. Based on the Peak Signal-to-Noise Ratio (PSNR), the suggested technique is compared to several existing methods, and the PSNR value for the new approach is larger than the PSNR value for the existing approaches.

Guo *et al.* suggested an efficient technique for mammography image segmentation and classification utilizing the k-mean and SVM learning methods [23]. Using machine learning, the author classified the images into many groupings. Several activities are performed during pre-processing to raise the PSNR of the images for improved classification. A cancer image segmentation technique based on CAD was proposed by Zhu *et al.* [24]. The proposed method was implemented on the MIAS dataset using MATLAB 2017. When the noise reduction approach is paired with bit-wise segmentation, the resulting image has a higher PSNR value than earlier strategies. The finished image had sharp edges, making it very clear and simple to analyze.

Sechopoulos and Mann investigated the present state-of-the-art in the identification of mammography abnormalities [25]. The author states that traditionally many CAD methods are used for breast abnormality detection, but still, the accuracy of detection is quite low, and many times patients have to suffer due to non-diagnosis of abnormality. However, in the last decade, Artificial Intelligence (AI) and Machine Learning (ML)-based detection technologies have taken control and are now extensively deployed. So, obviously, in the future, there is a need to explore this further for better accuracy.

2. Materials and Methods

In this paper, a mammography image classification and detection model is proposed. It includes three basic blocks of image enhancement, segmentation, and learning, as depicted in Figure 2. The working of each block is explained in the following subsections.

2.1. Image Enhancement

The enhancement block is used to enhance the original images in terms of contrast to improve the PSNR value. Despite the usefulness of the global image contrast enhancement in certain contexts, there are occasions when it is essential to enhance the local features of mammography images. The amount of pixels in this region may have a minor impact on the global transformation calculation. Therefore, a transformation function based on the grey level distribution or other image attributes in the vicinity of each pixel is required i.e., known to be Local Contrast Modification (LCM) [26]. Initially, the median filter is applied with the Local Contrast Modification (LCM) method to reduce the induced noise in the image. Further, an adaptive CLAHE in conjunction with the Wiener filter is used to enhance the images obtained after applying LCM. It helps circumvent the issue of noise amplification in homogenous regions of an image. In contrast to traditional HE, the CLAHE algorithm works on tile-sized chunks of an image to disperse the image's brightness levels [27].
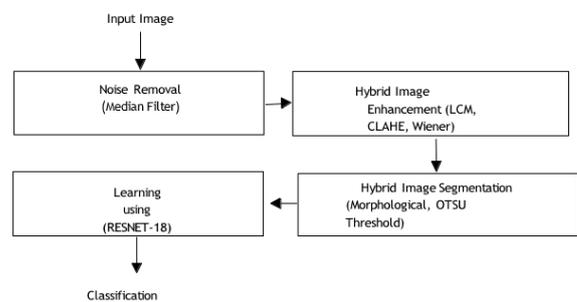
Figure 2. Block diagram of the proposed method

2.2. Image Segmentation

The segmentation block is used to segment the enhanced image from the abnormal areas so that cancer can be detected clearly. In this study, segmentation is accomplished through the use of a hybrid approach that combines the morphological features extraction method with the Otsu thresholding method [14]. Otsu presented a non-parametric, unsupervised, optimum global thresholding approach for image segmentation [28]. It uses just the zeroth and first-order grey-level cumulative moments. This technique selects the appropriate threshold based on the histogram's global characteristics in such a way

that it distributes pixels evenly across the image and sharpens image boundaries so that any radiologist can notice them.

2.3.   Image Data Augmentation

It is difficult to train a model using a deep neural network without access to a huge dataset [29]. A model may suffer from under-fitting if the dataset has fewer samples per class. As a result, we are unable to train the model to its full potential. The random transformation criteria may be used to increase the data set size and variety to solve this problem. Data augmentation refers to the practice of changing data randomly. Rotation, shearing, zooming, cropping, flipping, and so on are typical examples of data transformation criteria. As intraclass variances are introduced throughout the data augmentation process, the model has to be trained to account for them. Over-fitting may arise when there is a significant difference between a number of samples per class in a dataset, however, the data augmentation approaches address this problem. For instance, training may favor classes with more data over those with fewer data points. As a result, the data augmentation method helps ensure that each class has a sufficient amount of samples for accurate learning. Over-fitting and under-fitting are prevented using a real-time data augmentation method in this work. The mammography image dataset is randomly divided into training and test datasets in the ratio of 80:20. With transformations like shift, shear, standard normalization, and zoom, five additional images are added to each mammogram image in the training dataset [30]. At each epoch during training, the raw data is passed through a set of transformation operations. Therefore, model learnability increases with the diversity of the training set.

2.4.   Transfer Learning

Classifiers based on Neural Networks (NN) are capable of learning complicated correlations between training datasets. In the literature, Bayesian network, Artificial Neural Network (ANN), Back-Propagation Neural Network (BPNN), Multi-Layer Perceptron (MLP), and Recurrent Neural Network (RNN) classifiers are used for image classification. CNN is the NN classifier that has received the most attention. The notion of transfer learning is used to circumvent the under-fitting that often happens when training a

model with a smaller dataset [31]. Previously-trained model's weights are used to create a new model using a different dataset. Thus, learning begins from a higher vantage point to avoid insufficient data. Moreover, transfer learning does not need training and test datasets to have similar distributions.

Several state-of-the-art pre-trained architectures are available to exploit transfer learning. In this paper, various pre- trained architectures such as ResNet, DenseNet, VGG, MobilNet, and NasNet are evaluated on the prepared dataset. Among these, ResNet-18 architecture was found to be most effective in terms of performance for classifying mammography images. The architecture of ResNet-18 is shown in Figure 3. Each mammogram image goes through successive convolution with different filters and 512

Mammogram Images

Convolution 3×3, filters 64, S= [2 2], P= [3 3 3 3]
Batch Normalization
ReLU
Max pooling 3×3, S= [2 2], P= [1 1 1 1]
Residual Block-2A Convolution 3×3, filters 64, S= [1 1], P= [1 1 1 1]
ReLU
Residual Block-2B Convolution 3×3, filters 64, S= [1 1], P= [1 1 1 1]
ReLU
Residual Block-3A Convolution 3×3, filters 128, S= [2 2], P= [1 1 1 1] Convolution 1×1, filters 128, S= [2 2], P= [0 0 0 0]
ReLU
Residual Block-3B Convolution 3×3, filters 128, S= [1 1], P= [1 1 1 1]
ReLU
Residual Block-4A Convolution 3×3, filters 256, S= [2 2], P= [1 1 1 1] Convolution 1×1, filters 256, S= [2 2], P= [0 0 0 0]
ReLU
Residual Block-4B Convolution 3×3, filters 256, S= [1 1], P= [1 1 1 1]
ReLU
Residual Block-5A Convolution 3×3, filters 512, S= [2 2], P= [1 1 1 1] Convolution 1×1, filters 512, S= [2 2], P= [0 0 0 0]
ReLU
Residual Block-5B Convolution 3×3, filters 512, S= [1 1], P= [1 1 1 1]
ReLU
Average-pooling 7×7, S= [7 7], P= [0 0 0 0]
Fully connected, 3 neurons
Softmax

Normal	masses	microcalcification

Figure 3. ResNet-18 architecture

characteristics are retrieved using Convolution5X layers. All the features retrieved by Convolution5X are given to FC1000, a fully connected layer that begins training based on the extracted features. The training rate is set at 0.001 as the standard rate of learning. Further, the model starts training itself to detect the image as normal, masses, or micro-calcification. Some parameters defined for FC1000 layer for the process of learning are shown in Table 1.

Table 1. Network training parameters

Parameter	Value
Network Layer	FC1000
Hardware Source	Single GPU
Learning Rate	0.001
Learning Rate Schedule	Constant Frequency
Max iterations	1560
Epochs	15

3. Results

In this paper, a framework is designed for mammography image enhancement, segmentation, and classification. The experiment was performed on AMD Ryzen 5 3600 CPU, 32 GB RAM, RTX 2060 Super GPU using Python language at Anaconda 1.10.0, Jupyter Notebook 6.2.0. In this section, the details of the dataset, experimental methodology, results, architecture explainability, and comparative analysis with existing methods are presented.

3.1. Dataset

The proposed method is evaluated on Mammographic Image Analysis Society (MIAS) dataset [32]. It includes 322 original images (161 subjects) in Portable Gray Map (PGM) format with 50-micron resolution and associated truth data. For example, the size of data for normal, masses and microcalcification is 204, 93, and 25, respectively. Further, the dataset has information about breast type such as fatty (106 images), fatty-glandular (104 images), and dense (112 images). Also, information about the severity of abnormalities like benign (66 images) and malignant (52 images) is given in the dataset. The size of all the images is 1024 1024 pixels. Further, the x, y image-coordinates of center of abnormality, and the approximate radius (in pixels) of a circle enclosing the abnormality in case of masses and microcalcification are also given in the dataset.

All the images (322) from the MIAS dataset are utilized after enhancement and segmentation of the region of interest. After image segmentation, the dataset is divided into 80:20 ratios for training and testing, respectively. Therefore, 257 images belong to the training set and the rest images are used for testing.

3.2. Results

The proposed approach is utilized to detect abnormalities in mammography images in terms of masses and micro- calcification, by employing three major blocks of enhancement, segmentation, and classification. The images are enhanced to increase the PSNR of the original image using noise removal and LCM, adaptive CLAHE with Wiener filter. The resultant image is better in contrast and clarity with a better PSNR value. Further, the improved images are segmented using the morphology method with the Otsu threshold to locate the region of interest i.e., regions where abnormality is found. The achieved values of PSNR for random images from the used dataset are shown in Figure 4. Thus, hybrid enhancement and segmentation methods are used. The confusion matrix, its parameters (TP, TN, FP, and FN), and various performance measures such as precision, recall, F1-score, accuracy, and Dice similarity coefficient (DSC) of segmentation block for

Table 2. Confusion matrix and values for various parameters and performance measures of segmentation block

		Normal	Micro-calc	Masses
Confusion Matrix	Normal	192	5	7
	Micro-calc	0	24	1
	Masses	1	3	89
Parameters	TP	192	24	89
	TN	117	289	221
	FP	12	1	4
	FN	1	8	8
Performance Measures	Precision	0.941	0.960	0.957
	Recall	0.995	0.750	0.957
	F1-Score	0.967	0.842	0.957
	Accuracy	0.960	0.972	0.963
	DSC	0.652	0.457	0.484

Precision=$[TP/(TP + FP)]$; Recall=$[TP/(TP + FN)]$; F1-Score=[2 Precision Recall/(Precision + Recall)];

Accuracy= $[(TP + TN)/(TP + TN + FP + FN)$;

DSC=$[2 \times TP/(2 \times TP + FP + FN)]$.

different class of mammogram images is shown in Table 2.

An example of original, enhanced, and segmented images is shown in Figure 5. After image segmentation and splitting the dataset into training and test, the training dataset is augmented.
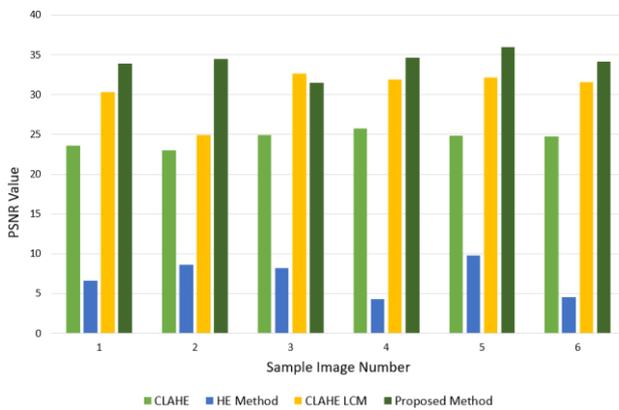
Figure 4. PSNR values for different image enhancement methods
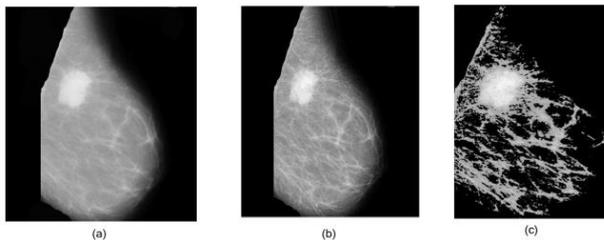
Figure 5. Examples of a) original, b) enhanced, and c) segmented images

The effect of segmentation block and image augmentation is shown in Table 3. It can be observed that the segmentation block significantly increases the classification accuracies of various architectures used for mammographic image classification. Further, the segmented image is given as input to the proposed transfer learning block which is trained to classify the mammogram images.

The proposed transfer learning model is trained to identify calcium depositions around the breast and classify the image as normal, mass, or microcalcification. Various pre-trained architectures are utilized and evaluated for mammography image classification such as ResNet, DenseNet, VGG, MobilNet, and NasNet. For the evaluation of these architectures, the dataset is divided into the ratio of

80:20. The experimental results for different architectures on various performance metrics is shown in Table 4. Several variants of these architectures were tested and ResNet-18 was found to be most effective for mammography image classification. For example, the values of precision, recall, F1-score, and accuracy for ResNet are reported as 98.89%, 98.89%, 98.9%, and 98.91%, respectively.

Table 3. Effect of segmentation block and image augmentation on classification accuracy of various architectures

Pre-trained architectures	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
VGG	95.6	95.65	95.65	95.69
ReseNet-18	98.89	98.89	98.9	98.91
NasNet	94.78	94.78	94.78	94.87
MobileNet–V2	89.81	89.82	89.82	89.9
DenseNet	96.28	96.27	96.28	96.56

SNA=No Segmentation No Augmentation, NSWA=No Segmentation With Augmentation, WSNA=With Segmentation No Augmentation, WSWA=With Segmentation With Augmentation

Table 4. Values of different performance metrics for transfer learning architecture on different pre-trained models

Architecture	Accuracy (%)			
	NSNA	NSWA	WSNA	WSWA
VGG	93.13	93.9	94.34	95.69
ReseNet-18	96.83	97.78	98.2	98.91
NasNet	92.41	93	93.85	94.87
MobileNet–V2	86.62	87.9	88.72	89.9
DenseNet	93.89	94.86	95.43	96.56

The testing accuracies of evaluated architectures at different epochs are shown in Figure 6. The accuracies of all architectures start from the elevated point due to their existing learning. For example, accuracies of ResNet, DenseNet, VGG, MobilNet, and NasNet architectures at the very first epoch are reported as 46.67%, 45.59%, 35.07%, 37.21%, and 36.81%, respectively. All these architectures achieve higher accuracies within 15 epochs and ResNet-18 reports the highest classification accuracy. The proposed model reported a classification accuracy of 98.91% using ResNet-18, which is better than several existing techniques.

A confusion matrix effectively visualizes True Positive (TP), True Negative (TN), False Positive
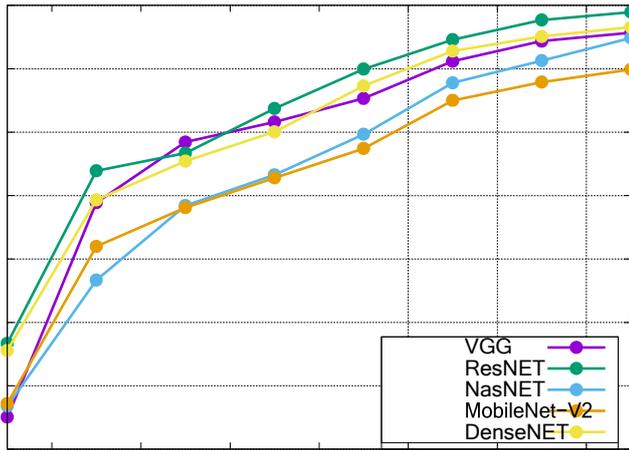
Figure 6. Testing accuracies of various architectures at different epochs
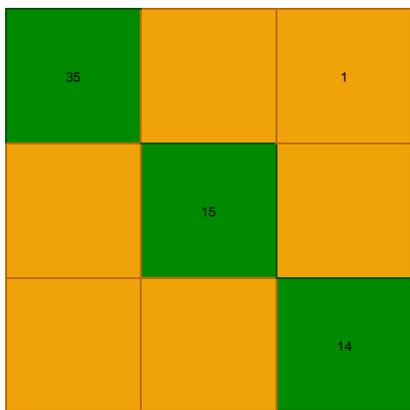
Figure 7. Confusion matrix of the proposed method for image classification

(FP), and False Negative (FN). The confusion matrix is represented using a 3×3 block where the x-axis represents the predicted class and the y-axis represents the true class, as shown in Figure 7. Random images are taken from the dataset for the evaluation to make the network unbiased. It is apparent from the confusion matrix that the matching probability of true and predicted class samples is higher for all true positive cases. For example, out of 36 normal class images, only 1 image is misclassified. All images from masses and micro-calcification classes are correctly identified. It also shows the robustness of the proposed model.

3.3. Explainability of Proposed Architecture

Despite the usefulness of deep learning models, their underlying structures are so complicated that their workings remain a mystery, rendering them black boxes [12]. As a result, visualizing the logic behind the predictions made by deep learning algorithms is challenging. To avoid being labeled a black box, a model should adhere to the quality of explain- ability and causality. In a more general sense, the model's causality determines how practical it is and explainability stresses the openness and accountability of the model.

In this work, the explanability and causality of the proposed model have been evaluated using the LIME method. The LIME method provides a human-comprehensible justification for why the learned model has chosen a certain action for a given input. The explainability of the proposed model is demonstrated in Figure 8. The segmented images are fed to the model in the form depicted in the first image. In the second image, we can see how the model uses the super-pixels to pick out specific areas of interest. Superpixels are a computer vision and image processing approach for grouping pixels intomeaningful areas. They depict an image at a higher level by grouping related pixels based on color, texture, and spatial closeness. A close inspection reveals that only the super-pixels that make up the mass region of the mammography image are seen. The input image is correctly labeled by the model based on these super-pixel segments. In the latter two images, the yellow super-pixels represent features of a certain
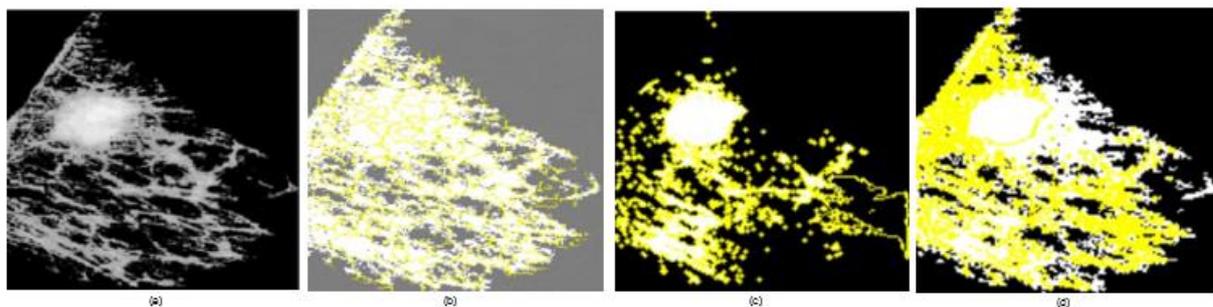
Figure 8. Explainability of architecture representing a) original image, b) superpixels representing regions of interest and c)–d) image interpretation by the model at different hidden layers

class that increase the chances of a successful prediction, while the white super-pixels reflect features of the image that decrease the likelihood of a given classification. Therefore, it provides valuable insights into the model's decision-making process and helps ensure that its predictions are trustworthy and aligned with the intended purpose of the model.

3.4. Comparative Analysis

The performance of the proposed method for mammography image classification is compared with other existing methods utilizing transfer learning. The accuracies of existing and proposed methods with their architecture are shown in Table 5. Jiang *et al.* [33] and Falconi *et al.* [35] reported very lower accuracies using GoogLeNet, AlexNet and MobileNet, and ResNet-50, respectively. Shen *et al.*, have used various combinations of Resnet-50 and VGG-16 architectures by exchanging the fully connected layers of each other. However, they reported only 95% accuracy for each combination [34]. Similarly, the VGG-16 architecture used by Guan *et al.* reported a lower accuracy of 90.1% [40]. Although the dataset-based transfer learning method of Alzubaidi *et al.* has achieved a better accuracy [39], the proposed method outperforms other existing methods reporting 98.91% accuracy.

Table 5. Classification accuracies of existing methods using transfer learning

Methods	Architecture	Accuracy (%)
Jiang *et al.* [33]	GoogLeNet	88
	AlexNet	83
Shen *et al.* [34]	ResNet	95
	Resnet-VGG	95
Kassani *et al.* [38]	DCNN	92.50
Falconi *et al.* [35]	MobileNet	74.30
	ResNet-50	78.40
Alzubaidi *et al.* [39]	Datasets-based Transfer Learning	96.10
Guan *et al.* [40]	VGG-16	90.10
Proposed Model	ResNet-18	98.91

4. Conclusion

The detection of breast abnormalities in the earlier stages is very important as breast cancer is a progressive disease. The detection of breast cancer from mammogram images is even sometimes also difficult for the doctor as the deposits of calcium in the form of lumps are not clear in many cases. As a result, for reliable identification and classification, a CAD-based technique is necessary. A transfer learning-based model for mammography image categorization is provided in this work. The proposed model is built around three main modules such as enhancement, segmentation, and learning. Initially, the original images are enhanced to remove noise and improve contrast. The enhanced images had shown a higher PSNR value than the original and contrast was also found to be better than various existing methods. The filtered image with improved PSNR and contrast seems effective for the segmentation of the region of interest. Finally, the ResNet-18 model is employed where segmented images are fed as input. It effectively classifies the mammography images as normal, masses, and micro-calcification. Further, in order to break the black-box dilemma and build public trust and acceptance, the explainability of deep-CNN architecture is also discussed.

The proposed model will add a valuable contribution to the existing literature due to its higher accuracy, as it can assist radiologists and doctors in detecting breast cancer efficiently. The current work is likely to contribute significantly to the field of computer-aided diagnosis. In the future, the network can be trained to identify abnormalities related to bilateral asymmetry and architecture distortion and the network may be trained for other medical purposes as well.

Acknowledgments

References

1- Russo J, Russo IH, "Development of the human breast", Maturitas (2004) 49:2–15. 10.1016/j.maturitas.2004.04.011.

2- Q. Hadi, I. Masroor, and Z. Hussain, "Mammographic Criteria for Determining the Diagnostic Accuracy of Microcalcifications in the Detection of Malignant Breast Lesions", *Cureus*, 11 (10), (2019).

3- Jaime R. Herndon, MS, MPH, "Normal and Ab- normal Mammogram Images.", online available at: https://www.verywellhealth.com/mammogram-images-descriptions-and-details-4020351. (accessed on Sept 30, 2022.)

4- N. Z. Sonthineni, Chaitra, Namita Mohindra, and Agrawal Vinita, "Correlation of digital mammography and digital breast tomosynthesis features of self detected breast cancers with human epidermal growth factor receptor type 2/neu status", *South Asian J. cancer*, 7 (2), pp. 171– 174., (2018).

5- Becker, Sven, "A historic and scientific review of breast cancer: The next global healthcare challenge", *International Journal of Gynecology, and Obstetrics*. 131: S36–S39, (2015).

6- L. M. B. Hashem, R. H. M. Ali, M. H. Helal, E. El. E. L. Gemeae, and A. F. I. Moustafa, "Characterization of breast masses: a comparative study between automated breast ultrasound (ABUS) and digital breast tomosynthesis (DBT)", *Egypt. J. Radiol. Nucl. Med*., 51 (1), (2020).

7- Bland, Kirby I, Edward M Copeland, V Suzanne Klimberg, and William J Gradishar (2017), The Breast E-Book: Comprehensive Management of Benign and Malignant Diseases (Elsevier Health Sciences).

8- Heywang-Kobrunner SH, Hacker A, Sedlacek S, "Advantages and Disadvantages of Mammography Screening", *Breast Care (Basel)*. 2011;6(3):199-207. doi: 10.1159/000329005. Epub 2011 May 27.

9- S. J. S. Gardezi, A. Elazab, B. Lei, and T. Wang, "Breast cancer detection and diagnosis using mammographic data: Systematic review", *J. Med. Internet Res*., 21 (7), pp. 1–22, (2019).

10- Komen, Susan G, "Breast Calcifications", *American Cancer Society*. pp. 4–5, (2020).

11- K. He, X. Zhang, S. Ren and d J. Sun, "Deep Residual Learning for Image Recognition", *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas*, NV, pp. 770 778, (2016).

12- M. T. Ribeiro, S. Singh and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (2016).

13- M. R. Shelda Mohan, "Optimized Histogram Based Contrast Limited Enhancement for Mammogram Images", *ACEEE Int. J. Inf. Technol*., 3 (1), pp. 66–71, (2013).

14- K. Kipli *et al.*, "Morphological and Otsu's thresholding-based retinal blood vessel segmentation for detection of retinopathy", *Int. J. Eng. Technol*., 7 (3), pp. 16–20, (2018).

15- N. Alam, E. R. E. Denton, and R. Zwiggelaar, "Classification of micro- calcification clusters in digital mammograms using a stack generalization based classifier", *J. Imaging*, 5 (9), (2019).

16- R. A. L. Al-juboori, "Contrast Enhancement of the Mammographic Image Using Retinex with CLAHE methods CLAHE", *Iraqi J. Sci*., 58 (1), pp. 327–336, (2017).

17- K. Akila, L. S. Jayashree, and A. Vasuki, "Mammographic image enhancement using indirect contrast enhancement techniques - A com- parative study", *Procedia Comput. Sci*., vol. 47, no. C, pp. 255– 261, (2015).

18- T. Tot, M. Gere, S. Hofmeyer, A. Bauer, and U. Pellas, "The clinical value of detecting microcalcifications on a mammogram", *Semin. Cancer Biol*., no. October, (2019).

19- A. Kamra, V. K. Jain, S. Singh, and S. Mittal, "Characterization of Architectural Distortion in Mammograms Based on Texture Analysis Using Support Vector Machine Classifier with Clinical Evaluation", *J. Digit. Imaging*, 29 (1), pp. 104–114, (2016).

20- J. R. Quinlan, "Induction of decision trees", *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, (1986).

21- Yousef, Waleed A *et al.*, "Method and System for Image Analysis to Detect Cancer". *arXiv* 26(19): 1–21, (2019).

22- Zou, L.; Yu, S.; Meng, T.; Zhang, Z.; Liang, X.; Xie, Y., "A Technical Review of Convolutional Neural Network-Based Mammographic Breast Cancer Diagnosis", *Comput. Math. Methods Med*. 2019, vol:2019, pp 1–16, (2019).

23- Guo, R.; Lu, G.; Qin, B.; Fei, B., "Ultrasound Imaging Technologies for Breast Cancer Detection and Management: A Review". *Ultrasound Med. Biol*. 2018, 44, 37–70, (2018).

24- Zhu, Wentao, *et al.*, "Deep Multi-Instance Networks with Sparse La- bel Assignment for Whole Mammogram Classification", *CoRR*, vol. abs/1612.05968. (2016).

25- I. Sechopoulos and R. M. Mann, "Stand-alone artificial intelligence - The future of breast cancer screening?", *Breast*, vol. 49, pp. 254–260, (2020).

26- M. Sundaram, K. Ramar, N. Arumugam, G. Prabin, "Histogram Mod- ified Local Contrast Enhancement for mammogram images", *Applied Soft Computing*, 11 (8), pp. 5809-5816, (2011).

27- Jintasuttisak, T.; Intajag, S., "Color Retinex Image Enhancement by Rayleigh Contrast Limited Histogram Equalization", *International Con- ference on Control, Automation and Systems.*,10,692-697, (2014).

28- N. Otsu, "A threshold selection method from gray-level histogram", *IEEE Trans. Syst. Man Cybern*, vol. 9, no. 1, pp. 62-66, (1997).

29- A. Pal, R. Srivastva, Y. N. Singh, "CardioNet: An Efficient ECG Arrhythmia Classification System Using Transfer Learning", *Big Data Research*, Vol. 26, (2021), 100271.

30- Shorten, C., Khoshgoftaar, T.M., "A survey on Image Data Augmenta- tion for Deep Learning", *Journal of Big Data* 6 (60), (2019).

31- C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, C. Liu, "Survey on deep transfer learning", *in: 27th International Conference on Artificial Neural Networks, Rhodes, Proceedings, Part III*, (2018).

32- Suckling, J.; Parker, J.; Dance, D.; Astley, S.; Hutt, I.; Boggis, C.; Rick- etts, I.; Stamatakis, E.; Cerneaz, N.; Kok, S.; *et al.* Mammographic Image Analysis Society (MIAS) database v1. 21. *Med. Imaging 2014 PACS Imaging Inform. Next Gener. Innov*. 2015, (2015), 9039. Available online: https://www.repository.cam.ac.uk/handle/1810/250394/ (accessed on 30 September 2022).

33- F. Jiang and H. Liu, "Breast Mass Lesion Classification in Mammograms by Transfer Learning", *ICBCB '17 Hong Kong*, no. September (2018).

34- L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. McBride, and W. Sieh, "Deep Learning to Improve Breast Cancer Detection on Screening Mammography," *Sci. Rep.*, 9 (1), pp. 1–12, (2019).

35- L. G. Falconi, M. Perez, and W. G. Aguilar, "Transfer Learning in Breast Mammogram Abnormalities Classification with Mobilenet and Nasnet", *Int. Conf. Syst. Signals, Image Process.*, vol. 2019-June, pp. 109–114, (2019).

36- Lepori, D., "Inflammatory breast disease: The radiologist's role", *Diagn Interv Imaging*, 96: 1045-64, (2015).

37- F. Ma, L. Yu, G. Liu, and Q. Niu, "Computer aided mass detection in mammography with temporal change analysis", *Comput. Sci. Inf. Syst.*, vol. 12 (4), pp. 1255–1272, (2015).

38- S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Breast Cancer Diagnosis with Transfer Learning and Global Pooling", *ICTC 2019- 10th Int. Conf. ICT Converg. ICT Converg. Lead. Auton. Futur.*, no. c, pp. 519–524, (2019).

39- L. Alzubaidi, O. Al-Shamma, M. A. Fadhel, L. Farhan, J. Zhang, and Y. Duan, "Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model", *Electron.*, vol. 9, no. 3, (2020).

40- S. Guan and M. Loew, "Breast cancer detection using transfer learning in convolutional neural networks.," Proc. *Appl. Imag. Pattern Recognit. Work.*, vol. 2017-October, (2018).