How to Cite:
Sridevi, V., & Samath, J. A. (2022). Design and implementation of transfer learned deep CNN with feature fusion for automated mammogram classification. *International Journal of Health Sciences*, *6*(S6), 3033–3047. https://doi.org/10.53730/ijhs.v6nS6.10267

Design and implementation of transfer learned deep CNN with feature fusion for automated mammogram classification

V. Sridevi
Assistant Professor, Department of Computer Science, PSG College of Arts & Science and Research Scholar of Govt. Arts College, Udumalpet, India
Email: vissridevi@gmail.com

J. Abdul Samath
Assistant Professor, Department of Computer Science, Chikkanna Government Arts College, Tirupur, India
Email: abdul_samath@yahoo.com

Abstract---Radiologists frequently struggle to define mammography mass lesions, resulting in unneeded breast biopsies to eliminate suspicions, which adds exorbitant costs to an already overburdened patient and medical system..Existing models have limited capability for feature extraction and representation, as well as cancer classification. Therefore, we built deep Convolution neural networks based Computer-aided Diagnosis system to assist radiologists in classifying mammography mass lesions. Here, Two-view LASSO regression feature fusion and fine-tuned transfer learning network model VGG16 were applied for identification of mammogram cancer. First, two independent CNN branches are utilized to extract mammography characteristics from two different perspectives. Feature Extraction is performed by fine-tuning pre-trained deep network models VGG16 which extracts deep convolutional features. Second, the features of the VGG16 models are serially fused using LASSO regression. Lastly, the fused features are entered into the Fully Connected Layer for mammogram classification. The high accuracy of 95.24, senstitivity of 96.11% and AUC score of 97.95% of the proposed approach revealed that it should be used to enhance clinical decision-making.

Keywords---transfer learning, CNN, VGG16, LASSO regression, augmentation, mammogram classification

Introduction

In recent years, artificial intelligence has significantly improved pathologists' diagnostic efficiency and accuracy, as well as reduced misdiagnoses due to human fatigue and a lack of clinical experience in the field of medical [1]. Breast cancer can be cured in 40% to 90% of cases if detected and treated early [2]. The most common diagnostic procedure for identifying the breast cancer is Mammography which is the X-ray image of a breast [3,4]. The mammography is used by doctors to find the signs of breast cancer. It may identify calcifications, lumps, dimpling, and so on. These are the most prevalent symptoms of breast cancer in its early stages [5]. Mammographic images are a cost-effective way for detecting breast cancer, and radiologists can use these images to diagnosis. On the other side, radiologists are having more pressure because of producing the vast number of mammographic images every day. In addition the rate of misdiagnosis rate has been gradually increased. As a result, putting together a CAD) system can minimize radiologists' burden while also enhancing diagnosis accuracy. Radiologists can use the CAD to distinguish between normal and diseased tissues and diagnose clinical phases.

In recent years, deep learning [6,7] specifically Convolution Neural Networks (CNNs) [8] has been demonstrated to perform well and produced promising results in mammogram classification [9,10]. Deep learning CADs have been applied to a variety of medical domains, including Interstitial lung disease, Cervical cancer classification, pulmonary Peri-fissural nodule classification [9], and Thoraco-abdominal lymph node classification [10]. The works on breast lesions [11,12] piqued our curiosity in particular. Convolution Neural Networks automatically discover discriminative features and its design is specifically transformed to take use of the input image's 2D structure [10,13]. We require huge annotated datasets to train deep CNNs, which are missing in the medical arena, particularly for breast cancer. Furthermore, training a CNN takes a lot of computing power, memory, and time. We can easily overfit the model with the limited data. One method is to use transfer learning [14] from medical images and then fine-tune the results as stated in [10].

In deep learning applications, transfer learning is often used. When the amount of data is generally limited, it has proven to be particularly useful in the medical arena [10,15]. The applicability of transfer learning to mammography images has not yet thoroughly investigated. Scholars have recently become interested in fine-tuned transfer learning networks for performing cancer classification tasks with substantial classification performance. The goal of this research is to find and classify breast cancer at an early stage into benign, malignant masses, and normal in order to reduce cancer risk and improve treatment effectiveness. In this proposed work, we built a two-view feature fusion network to monitor and discover mammograms. The following is a summary of our contributions: First, identify the Cranio-Caudal (CC) and Medio-Lateral (MLO) views of mammograms for the same breast. We suggested two independent CNN branches make up the two-view feature fusion network, which are utilised to extract mammography characteristics from two different perspectives. Fine-tuning pre-trained deep network models VGG16 extracts deep convolutional features. Second, the features of the VGG16 models are serially fused using LASSO regression. Lastly, the fused

features are entered into the Fully Connected Layer for mammogram classification. The Curated Breast Imaging Subset of the Digital Database for Screening Mammography is used in all of the tests in this study (CBIS-DDSM)

Related Work

Rakhlin et al. (2018) [16] offered a deep learning-based method in which the features were extracted using VGG16, Inception V3 and Resnet-50. The classification is driven by LightGBM classifier with ten-fold cross-validation. Their approach had an average accuracy of 87.2% across the board. Chougrad et al. (2018) [17] developed a mammogram-diagnosing deep convolutional neural network. The impact of transfer learning was investigated in training CNN model to find optimum fine-tuning approach. While picture patches can improve feature extraction efficiency to some extent, they also miss some global diagnostic information. On the DDSM database, they achieved 97.35% accuracy and 0.98 AUC; and on the INbreast database, they achieved 95.50% accuracy and 0.97 AUC.

Wang et al. (2016) [18] used stacked auto-encoder which is a deep learning model to diagnose breast lesions. The combined study of microcalcifications and breast masses resulted in satisfactory accuracy. Li et al. (2018) [19] used CNN to detect anomalies and found that they had a high sensitivity for determining whether the anomalies were benign or malignant. Huynh et al. (2016) [20] employed transfer learning to obtain deep features automatically for breast lesions and found that they performed well when compared to characteristics extracted analytically. V. Sridevi et al (2020) [21] used the KMC-GF feature extraction. The KMC-GF approach has the advantage of obtaining fine texture data from the clustered region and providing shape and orientation features. When compared to GLCM features, the KMC-GF feature performs well in terms of early breast cancer identification

To diagnose breast microcalcification, Cai et al. (2019) [22] suggested a CNN model. Furthermore, in order to fully exploit the benefits of handcrafted features, they combined handcrafted and deep features in that study to increase the model's performance. Carneiro et al. (2015) [23] demonstrated the benefits of pre-trained deep learning models in medical images. With datasets INBreast and DDSM, an AUC of 0.90 was achieved in multiple mammography. V. Sridevi et al [24] used gray-scale based FCM clustering for segmentation which performs well, as it can expose pixels in terms of grey value level, allowing it to reveal hierarchical position of linked defects by grey value. Wang et al. (2019) [25] investigated a breast cancer based on feature fusion and CNN deep features which includes both texture features and density features. He et al. (2018) [26] developed a classification model for testing diagnostic performance on discriminating malignant tumours based on deep CNN features. They demonstrated that a deep learning classification model based on picture texture characteristics can successfully distinguish benign and malignant masses.

Research Aim and Scope

The goal of this research is to develop an effective automated deep learning approach to assist radiologists in accurately and quickly classifying lesions into benign, malignant and normal image. The workflow of the method is shown in Algorithm 1. The schematic representation of workflow is depicted in Figure 1.

(1) Input: MLO view mammography mass image $M_i, i = 1,2, \dots N$ and CC View mammographic mass image $C_i, i = 1,2, \dots N$

(2) Output: Diagnosis resultant matrix $Y\ \varepsilon\ R^N$

(3) Initially, pre-processing techniques are applied to remove artefacts and noise in raw mammograms, and then improve the quality of image and followed by data enhancement methods to highlighting the breast lesion.

(4) To deal with overfitting difficulties, data augmentation approaches are used to increase the number of mammograms, and the dataset is spitted into train, validation, and test sets.

(5) Computation of deep feature matrix using pre-trained model from each view of image $X_{mlo}\ \varepsilon\ R^{N\,X\,m}$ and $X_{cc}\ \varepsilon\ R^{N\,X\,n}$

(6) Calculate Fusion feature matrix using LASSO regression with L_1 – penalty loss function $X_f = [X_{mlo}, X_{cc}], X_f \varepsilon R^{N\,x\,(m+n)}$

(7) Train softmax classifier with fully connected layer and classify the input image as benign, malignant and normal X_f and predict M_i and C_i

(8) Return resultant matrix Y

Algorithm 1. Workflow of Proposed Model

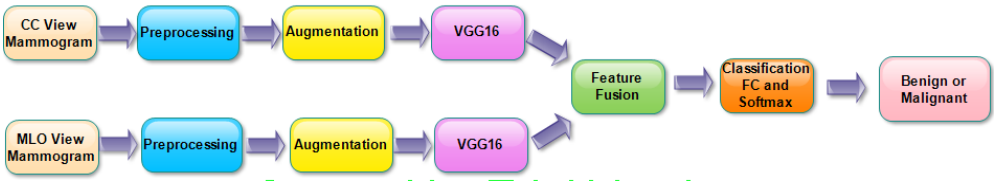

Figure 1. Diagrammatic Representation of workflow

Materials and Methods

Data sets

Our proposed model is trained and assessed with the Curated Breast Imaging Subset of DDSM (CBIS-DDSM) [27]. The dataset is divided into three categories: normal, benign and malignant. It contains mediolateral oblique and craniocaudal views of 861 mass lesion images which includes 912 benign and 784 malignant lesions

Data Distribution

We retrieve 1146 normal, 713 benign, and 743 malignant lesion images from the database. Our dataset has totally 5204 mammogram images which consist of CC and MLO view of 2602 breast. All data is divided into two sets: training and testing. There are 890 normal breasts and 1200 abnormal breasts in the training set (585 benign and 615 malignant). There are 256 normal, and 128 abnormal mammograms in the testing set (128 benign and 128 malignant). We proportionally partition the training data into ten subsets so that we may test the model's performance using ten-fold cross-validation. The summary of distribution of the data is shown in Table 1.

Table 1
Dataset Description

Type	Normal	Abnormal		Total
		Benign	Malignant	
Train	890	585	615	2090
Test	256	128	128	512
Total	1146	713	743	2602

Methodology

Pre-processing

Original mammograms are not able to achieve high accuracy in classification performance when used directly as the input of a convolutional neural network due to their high noise and poor contrast. As a result, some well-known pre-processing steps, such as contrast enhancement, bilateral filtering, and image normalisation, must be performed on the original mammograms. Pre-processing is employed to improve the quality and quantity of mammograms. Artefact removal, image improvement, and verification are the major steps in pre-processing. To obtain a more accurate result, artefacts are first removed from mammograms and then binary masking, morphological opening [28], and largest contour detection [29] algorithms are applied to determine edge detection and noise reduction. Second, image enhancement is used to make the malignant lesion more noticeable by enhancing the brightness and contrast of the original images. Gamma correction [30] and CLAHE [31] are the sub processes of this stage. There is a noticeable improvement in visibility after using CLAHE. Finally, assessment techniques such as MSE, PSNR, SSIM, RMSE, and Histogram Analysis are performed to processed images in the verification step to evaluate the obtained output.

Data Augmentation

When we have vast datasets [32], deep learning models perform better. Data augmentation, often known as jittering, is a popular method for increasing the size of our datasets. When training on very little data, data augmentation can raise the dataset's size to ten times or more than the original. This helps minimize overfitting. The method aids in the development of simpler and more robust

models that are more generalizable [33]. In order to augment the images, a sequence of random transformations is performed on the image. We employed random rotation range of 0–40 degrees, a shear range of 0.5, a zoom range of [0.5–1.5] and 0.25 shifts of width and height.

Fine-Tuned Transfer Learning

Transfer learning is the process of applying a previously trained model to a new situation. Transfer learning offers the advantage of reducing neural network model training time and resulting in decreased generalisation error. The following are the two basic techniques of adopting transfer learning: Weight Initialization and Feature Extraction. The following step is the transfer learning of deep learning.

- Create a base model and fill it with pre-trained weights.
- Run your new dataset through it, and make a note of the output of one (or more) layers from the base model. Feature extraction is the term for this process.
- Use that output as the input for a new, more compact model.
- Replace the last completely connected layer with the layer based on the number of classes in the target dataset.
- Initialize the weights in the new fully connected layer at random.
- The layers of the pre-trained network should be unfrozen, and the pre-trained weights should be used to initialize the rest of the weights.
- Retrain the neural network as a whole.
- Perform Fine-tuning that entails unfreezing the entire model and retraining it one new data with suitable learning rate. This has the ability to produce significant improvements by slowly changing the pre-trained features to the new data.

The transfer learning of pre-trained VGG16 models are used in this paper. The deep feature matrix was computed using pre-trained VGG16 model from each MLO and CC view of image. The VGG16 is divided into five blocks, each having two or three convolution layers and a Maxpooling layer, and has a depth size of 23 and 16 layers. By multiplying a set of weights with an input, the convolution layer achieves a linear process.

$$(y * x)\{n\} = \sum_{j=1}^{m} x(j).y(n - j + \frac{m}{2})$$

The input and kernel functions are x() and y(), respectively, and (x*y) signifies the dot product of input and kernel functions on number of variables n. After obtaining the feature matrix from the convolution layer, the maxpooling layer lowers the size of the feature maps by adding filters or kernels. After applying the filter to an input matrix of size (m X m) and a filter size of (y X y), the output size is:

$$(m \times m) * (y \times y) - (m - y + 1) \times (m - y + 1)$$

The input size of an RGB image is $224 \times 224 \times 3$ and the first convolutional layer's input dimension is $224 \times 224 \times 3$. The first block has two convolutional layers with 64 channels and followed by a 2×2 Maxpooling layer with stride 2×2. The second block has two convolution layers with 128 channels and a kernel size of 3×3, followed by a Maxpooling layer with stride 2×2. Three convolutional layers are followed by a Maxpooling layer in the last three blocks. The three convolutional layers in blocks 3, 4, and 5 have channel sizes of 256, 512, and 512, respectively, with the same kernel size of 3×3. In each Maxpooling layer, the initial input image is shrunk to half its original size. The outputted feature map from the last Maxpooling layer is $7 \times 7 \times 512$ pixels after the stack of convolutional and Maxpooling layers. A flatten layer is added to get $1 \times 25{,}088$ feature vector.

Feature Fusion using LASSO regression

In image recognition, feature fusion approaches are commonly employed to provide feature reduction and overcome the inadequacies of a single feature vector [34]. LASSO regression fusion strategies are being developed to combine two-feature vector as single-feature vector. The LASSO (Least Absolute Shrinkage and Selection Operator) regression is the most straightforward approach to merge two sets of feature vectors. The lasso regression does continuous shrinking as well as automatic variable selection. The benefits of backwards-stepwise selection are included into the lasso. The lasso predicted that many coefficients would be zero and just a tiny percentage would be nonzero [35]. Consider the standard linear regression model.

$$y_i = \alpha^* + X_i^T \beta^* + \varepsilon_i, i = 1, \ldots n.$$

Where

$X_j = (x_{1j}, \ldots, x_{nj})^T$ for j = 1,....p are the regressors, α^* is the constant parameter, $\beta^* = (\beta_1^*, \ldots, \beta_p^*)^T$ are the associated coefficients and y_k is the response for the kth observation. Let $X = [x_1, \ldots x_p]$ be the predictor matrix and let $y = (y_1, \ldots y_n)$ the response vector. Some coefficients can be shrunk or set to 0 in order to enhance prediction accuracy. By doing so, we sacrifice a small amount of loss function in order to lower the variance of the projected values and, as a result, enhance overall prediction accuracy. In Lasso Regression, the coefficient's absolute value of magnitude is introduced to the loss function as a penalty term. To provide a sparse solution to the optimization problem, the lasso problem employs the L_1 -penalised least squares criterion.

$$f(\beta) = \frac{1}{2}\sum_{i=1}^{n}(y_i - \lambda \sum_{j=1}^{p} X_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j|$$

Where λ is a non-negative regularization parameter.

L_1-penaly is $\lambda \sum_{j=1}^{p} |\beta_j|$ in equation.

The Lasso continuously reduces the coefficients toward 0 as λ increases, and some coefficients are reduced to exact 0 if λ is sufficiently large. Again, if lambda is 0, we will get OLS (Ordinary Least Square), but very large values will result in zero coefficients, resulting in under-fitting. Lasso reduces the coefficient of the less significant feature to zero, thus deleting it. This works effectively for feature selection if we have a large number of features..

Classification

The classifier is used to anticipate the probable attribution results by determining the link between the sets of attributes [36]. The test data is supplied into the network after the classifier has been trained to forecast the category and evaluate the algorithm's performance. After the weights were transferred, the retraining of the model was carried out with mammography dataset.Next concatenate the features getting from two view image. Finally, the 'Softmax' activation function computes the probability for each of the classes and generates the projected result. The softmax classifier is utilised as the input for direct fusion, and the fused feature is used as the output. The addition of dropout in classification layer is performed which improves the network's robustness. The cross-entropy cost function is minimised via stochastic gradient descent.

Training parameters of the pre-trained model

The parameters of optimizer design, learning rate, loss function, batch size and methods of preventing overfit were used to improve the accuracy. Table 2 gives an overview of the training parameters of the proposed model.

Table 2
Training Parameters of the pre-trained model

Training Options	Value
Learning rate	1.0000e-04
Mini-Batch Size	32
Loss Function	Cross-Entropy Loss
Optimizer	SGD
Activation Function	Softmax
Dropout	0.5
Cross-validation	10
Max-Epochs	90
Momentum	0.9

We chosen a learning rate of 1e-4 as the starting point, and divided it by 10 each time to stop the validation error. The early stopping method was used to monitor validation loss for every 90 epochs. Cross-entropy loss is used for modifying the weights in training which reduce the loss. If lower the loss, the model is the better one to identify cancer accurately [37]. The cross-entropy loss of a perfect model is zero. The cross-entropy is defined as

$$L_{CE}(y) = -\sum_{i=1}^{n} x_i \, log f_i(y)$$

where y indicates the input, n is the expected number of classifications, x_i is the true label, and $f_i(x)$ is the softmax probability of i^{th} class. The optimizer decides the actual output accurately. Stochastic-gradient descent (SGD) algorithm is employed in this research to determine the model parameters that decides the best fit between expected and actual outputs [38].

$$w_{t+1} = w_t - \alpha \frac{\partial L}{\partial w_t}$$

Where w_t *is the* weight to be updated at t time, learning rate *a, and* the *gradient* of loss function is $\partial L/\partial w$. Overfitting is common in transfer learning and has a significant impact on training outcomes. Data augmentation is the most common way to avoid overfitting. After the fully connected, the batch normalization and dropout layer [39] were added. In that, the convolution kernel was regularized in order to increase accuracy. The probability of 0.5 dropout layer switches off the activations during training time |40|. The zero setting of subset of units for the randomly chosen activation units is preventing one layer from relying too heavily on single unit in the preceding layer. We employed L2 regularisation which is equal to the square of the magnitude of coefficients. It combats overfitting by forcing weights to be small but not exactly 0.

$$L2(w) = \Lambda \sum_{1}^{n} w_i^2$$

Results and Discussion

Evaluation Metrics

The metrics of precision, We employed numerous metrics to assess the transfer learning models, including precision, recall, F1-score, accuracy, sensitivity, and specificity. The true positive (TP) , true negative (TN), false positive (FP) and false negative (FN) were calculated using each model of confusion matrix. For statistical examination of the models, the false positive rate, false negative rate, and false discovery rate were determined. Another essential metric for evaluating the performance of diagnostic results is the area under the receiver operating characteristic curve (ROC) (i.e., AUC score). The receiver operating characteristic (ROC) curve is used to compare the true positive rate (TPR) against the false positive rate (FPR) at various threshold levels which yields the AUC value. The following formula is used to calculate metrics:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Sensitivity \ or \ Recall = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

$$F1\ Score = 2\ \frac{Precision * Recall}{Precision + Recall}$$

$$FPR = \frac{FP}{FP + TN}$$

$$FNR = \frac{FN}{FN + TP}$$

Comparative experiment

Usually radiologists combine CC view and MLO view mammograms in order to find accurate decisions. When compared to single-view mammograms, two-view mammograms provide more complementary information which helps to increase the classification performance of the model. As a result, we put up a set of comparative studies to check the performance of feature fusion of two-view features with single view feature only. Figure 2 depicts their accuracy, sensitivity, specificity and AUC in three tasks.

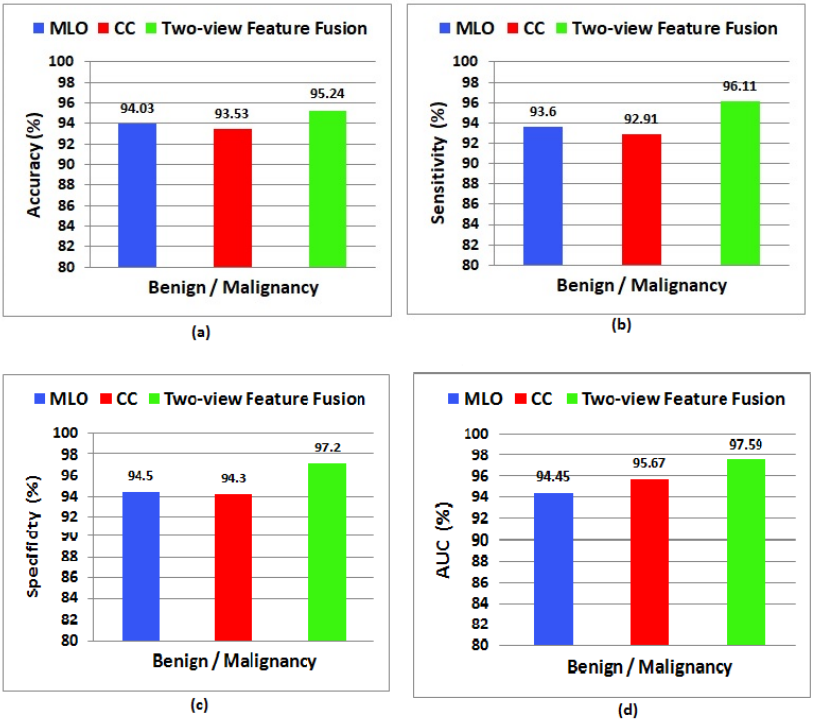

Figure 2. Performance Comparison between single-view and two-view network (a) Accuracy, (b) Sensitivity (c) Specificity (d) AUC

The model is visualised in the normal and pathological classification test to show that it can more precisely focus on the lesion area. The accuracy, sensitivity, and AUC of the suggested two-view feature fusion module are higher than those of existing modules, as demonstrated by the experimental findings. Table 3 summarizes the classification results.

Table 3
Comparison of CC View features, MLO View features, and Two View Feature Fusion with CBIS-DDSM

Data Set	Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
CBIS - DDSM	CC View Features	94.03	93.6	94.5	95.67
CBIS - DDSM	MLO View Features	93.53	92.91	94.3	94.45
CBIS - DDSM	Two View Feature Fusion	95.24	96.11	97.2	97.59

The two-view feature fusion's ROC curve is shown in Figure 3. The area under the green ROC curve, which indicates the classification result utilising fusion characteristics, is the largest, as shown by the three different colour curves. The curves also show that two-view feature fusion may successfully integrate the benefits of the two features, and addition features selected from two-view help in benign and malignant classification

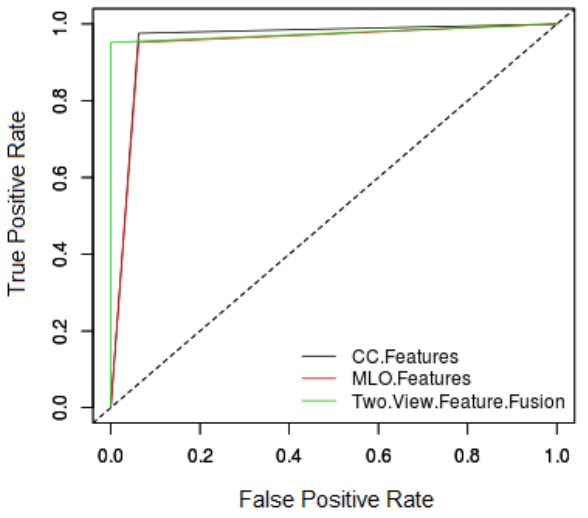

Figure 3. ROC curves of Feature Fusion

Experimental Discussions

The proposed model has a high degree of accuracy when it comes to detecting complicated breast masses or calcification, as well as breast density diversification. Transfer learning has been used to address the shortcomings of previous systems in detecting and categorising breast cancer masses in the framework provided here. Detecting and correctly labelling specific breast masses, such as spiculated and ill-defined lesions, is difficult. The proposed models have appropriately categorised these breast masses, which include a variety of types, edges, and diameters. The proposed technique eliminates the requirement for manual mass segmentation. Recognized masses or calcification are directly entered into the classifier, reducing complexity, decreasing computing time and increasing efficiency. The greatest test accuracy of 95.24, senstitivity of 96.11% and AUC score of 97.95% of the proposed approach revealed that it should be used to enhance clinical decision-making. The provided models were trained and tested using datasets obtained, producing consistent results, and proving the robustness of grading under various imaging settings. Another contribution to achieving the best configuration for the suggested model is parameter optimization. Various hyperparameters were manually tuned to arrive at the optimal classification model for this purpose.

Conclusion

Transfer learning is proposed in this research for the automatic detection of mammography breast masses or calcification. Our technique uses deep networks, which have more complex structures but fewer limitations, requiring less computer resources but providing more consistency. Image pre-processing techniques were employed to remove noise and artefacts from the images utilised in this investigation. To increase the quality of the raw mammograms, background removal and picture enhancement techniques were used. Data augmentation methods were used to overcome overfitting difficulties in mammography image processing, which occur when data is insufficient. The model's robustness and effectiveness were tested through a series of experiments. The experiments demonstrated its efficacy. Our study revealed that fine-tuned transdfer learning and feature fusion provides the best rate of correct classification. We will adjust and fine-tune other pre-trained models in the detection stage in the future to improve the system's classification performance. Furthermore, the classification accuracy will be improved by combining hand-crafted and CNN characteristics.

References

1. The American College of Obstetricians and Gynecologists. Cervical Cancer Screening and Prevention. Obstet. Gynecol. (2016), 168, 1–20.
2. Yu, X., Pang, W., Xu, Q. et al. Mammographic image classification with deep fusion learning. Sci Rep 10, 14361 (2020). https://doi.org/10.1038/s41598-020-71431-x
3. Bharati, S., Podder, P., & Mondal, M. R. H. (2020a). Artificial neural network based breast cancer screening: A comprehensive review. International

Journal of Computer Information Systems and Industrial Management Applications, 12, 125–137.
4. Thanh, D., & Surya, P. (2019). A review on CT and X-ray images denoising methods. Informatica, 43(2),151–159.
5. Khamparia, A., Bharati, S., Podder, P. et al. Diagnosis of breast cancer based on modern mammography using hybrid transfer learning. Multidim Syst Sign Process 32, 747–765 (2021). https://doi.org/10.1007/s11045-020-00756-7
6. Y. Bengio, Learning deep architectures for AI, found, Trends®Mach. Learn 2 (2009) 1–127, doi: 10.1561/2200000006.
7. J. Schmidhuber, Deep learning in neural networks: an overview, Neural Netw. 61 (2015) 85–117, doi: 10.1016/j.neunet.2014.09.003.
8. Y. LeCun, K. Kavukcuoglu, C. Farabet, Convolutional networks and applications in vision, in: Proc. 2010 IEEE Int. Symp. Circuits Syst., 2010, pp. 253–256, doi: 10.1109/ISCAS.2010.5537907.
9. F. Ciompi, B. de Hoop, S.J. van Riel, K. Chung, E.T. Scholten, M. Oudkerk, P.A. de Jong, M. Prokop, B. van Ginneken, Automatic classification of pulmonary peri- fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box, Med. Image Anal. 26 (2015) 195–202, doi: 10.1016/j.media.2015.08.001.
10. H.C. Shin, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R.M. Summers, Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, IEEE Trans. Med. Imaging 35 (2016) 1285–1298, doi: 10.1109/TMI.2016.2528162.
11. W. Peng, R.V. Mayorga, E.M.A. Hussein, An automated confirmatory system for analysis of mammograms, Comput. Methods Programs Biomed. 125 (2016) 134–144, doi: 10.1016/j.cmpb.2015.09.019.
12. Z. Jiao, X. Gao, Y. Wang, J. Li, A deep feature based framework for breast masses classification, Neurocomputing 197 (2016) 221–231, doi: 10.1016/j. neucom.2016.02.060.
13. G. Carneiro, J. Nascimento, A.P. Bradley, Unregistered multiview mammogram analysis with pre-trained deep learning models, in: Int. Conf. Med. Image Com- put. Comput.-Assist. Interv., Springer, 2015, pp. 652–660. http://link.springer. com/chapter/10.1007/978- 3- 319- 24574- 4 _ 78.
14. M. Oquab, L. Bottou, I. Laptev, J. Sivic, Learning and transferring mid-level image representations using convolutional neural networks, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit, 2014, pp. 1717–1724. http://www.cv-foundation.org/openaccess/content _ cvpr _ 2014/html/ Oquab _ Learning _ and _ Transferring _ 2014 _ CVPR _ paper.html . (accessedJanuary 27, 2017).
15. H. Chen, D. Ni, J. Qin, S. Li, X. Yang, T. Wang, P.A. Heng, Standard plane local- ization in fetal ultrasound via domain transferred deep neural networks, IEEE J. Biomed. Health Inf. 19 (2015) 1627–1636, doi: 10.1109/JBHI.2015.2425041.
16. Rakhlin, A., Shvets, A., Iglovikov, V., & Kalinin, A. A. (2018). Deep convolutional neural networks for breast cancer histology image analysis. In Paper presented at the international conference image analysis and recognition
17. Chougrad H, Zouaki H, Alheyane O. Deep convolutional neural networks for breast cancer screening. Computer Methods and Programs in Biomedicine.

2018; 157:19–30. https://doi.org/10.1016/j.cmpb.2018.01.011 PMID: 29477427.
18. Wang, J. et al. Discrimination of breast cancer with microcalcifications on mammography by deep learning. Sci. reports 6, 1–9 (2016).
19. Li, B., Ge, Y., Zhao, Y., Guan, E. & Yan, W. Benign and malignant mammographic image classification based on convolutional neural networks. In Proceedings of the 2018 10th International Conference on Machine Learning and Computing, 247–251 (2018).
20. Huynh, B. Q., Li, H. & Giger, M. L. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J. Med. Imaging 3, 034501 (2016).
21. V. Sridevi, Dr. J. Abdul Samath. (May-June 2020). Advancement on Breast Cancer Detection Using Medio-Lateral-Oblique (Mlo) and Cranio-Caudal (CC) Features. Test Engineering and Management, Vol. 83. pp. 85-93.
22. Cai, H. et al. Breast microcalcification diagnosis using deep convolutional neural network from digital mammograms. Comput. Math. Methods Med. 1, 4. https ://doi.org/10.1155/2019/27174 54 (2019).
23. G. Carneiro, J. Nascimento, and A. P. Bradley, Unregistered multiview mammogram analysis with pre-trained deep learning models, in Medical Image Computing And Computer-Assisted Intervention, pp. 652–660, Springer International Publishing Ag, Cham, Switzerland, 2015.
24. V. Sridevi, A. Nirmala. Inspection of Welding Images Using Image Segementation Techniques, International Journal of Engineering Research & Technology, Vol. 2 Issue 3, March - 2013
25. Z. Wang, M. Li, H. Wang et al., Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features, Institute of Electrical and Electronics Engineers Access, vol. 7, pp. 105146–105158, 2019.
26. Z. He, W. Lyu, G. Qin et al., “A feasibility study of building up deep learning classification model based on breast digital breast tomosynthesis image texture feature extraction of the simple mass lesions,” Chinese Journal of Radiology, vol. 52, no. 9, pp. 668–672, 2018.
27. Heath M, Bowyer K, Kopans D, Moore R, Kegelmeyer P. The digital database for screening mammography. Proceedings of the 5th International Workshop on Digital Mammography; 2000 Jun 11–14; Toronto, Canada. Medical Physics Publishing; 2000; pp.212-218.
28. Breuel, T.M. Efficient Binary and Run Length Morphology and its Application to Document Image Processing. arXiv 2007, arXiv:0712.0121.
29. Gong, X.Y.; Su, H.; Xu, D.; Zhang, Z.T.; Shen, F.; Yang, H. Bin an Overview of Contour Detection Approaches. Int. J. Autom. Comput. 2018, 15, 656–672.
30. Dhar, P. A, Method to Detect Breast Cancer Based onMorphological Operation. Int. J. Educ. Manag. Eng. 2021, 11, 25–31.
31. Hassan, N.; Ullah, S.; Bhatti, N.; Mahmood, H.; Zia, M. The Retinex based improved underwater image enhancement. Multimed. Tools Appl. 2021, 80, 1839–1857.
32. Halevy, P. Norvig, F. Pereira, The unreasonable effectiveness of data, IEEE Intell. Syst. 24 (2009) 8–12, doi: 10.1109/MIS.2009.36.
33. S.C. Wong, A. Gatt, V. Stamatescu, M.D. McDonnell, Understanding data aug- mentation for classification: when to warp?, ArXiv160908764 Cs. (2016). http: //arxiv.org/abs/1609.08764.

34. J. Ma, P. C. Yuen, and J.-H. Jian-Huang Lai, Linear dependency modeling for classifier fusion and feature combination, Institute of Electrical and Electronics Engineers Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 5, pp. 1135–1148, 2013.
35. Ogutu, J., Schulz-Streeck, T., & Piepho, H.-P. (2012). Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. BMC Proceedings, 6(Suppl 2), S10.
36. S. Aydin, N. Arica, E. Ergul et al., Classification of obsessive compulsive disorder by EEG complexity and hemispheric dependency measurements, International Journal of Neural Systems, vol. 25, no. 3, 2015
37. Rusiecki, A. Trimmed categorical cross-entropy for deep learning with label noise. Electron. Lett. 2019, 55, 319. [CrossRef]
38. D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, ArXiv14126980 Cs. (2014). http://arxiv.org/abs/1412.6980 (accessed January 31,2017).
39. Xu, T.; Feng, Z.-H.; Wu, X.-J.; Kittler, J. Learning adaptive discrimina-tive correlation _lters via temporal consistency pre-serving spatial feature selection for robust visual object tracking. IEEE Trans. Image Process. 2019, 28, 5596–5609. [CrossRef] [PubMed]
40. Srivastava , G.E. Hinton , A. Krizhevsky , I. Sutskever , R. Salakhutdinov , Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (2014) 1929–1958 .
41. Suryasa, W., Sudipa, I. N., Puspani, I. A. M., & Netra, I. (2019). Towards a Change of Emotion in Translation of Kṛṣṇa Text. Journal of Advanced Research in Dynamical and Control Systems, 11(2), 1221-1231.
42. Suwija, N., Suarta, M., Suparsa, N., Alit Geria, A.A.G., Suryasa, W. (2019). Balinese speech system towards speaker social behavior. Humanities & Social Sciences Reviews, 7(5), 32-40. https://doi.org/10.18510/hssr.2019.754
43. Mustafa, A. R., Ramadany, S., Sanusi, Y., Made, S., Stang, S., & Syarif, S. (2020). Learning media applications for toddler midwifery care about android-based fine motor development in improving midwifery students skills. International Journal of Health & Medical Sciences, 3(1), 130-135. https://doi.org/10.31295/ijhms.v3n1.290