Middle East Journal of Cancer; January 2024; 15(1): 40-51

An Ensemble Deep Learning Model for the Detection and Classification of Breast Cancer

Joy Christy Antony Sami**, PhD, Umamakeswari Arumugam**, PhD

*Department of Computer Science, School of Computing, SASTRA Deemed to be University, Thanjavur, India **School of Computing, SASTRA Deemed to be University, Thanjavur, India

Abstract

Background: Detecting breast cancer in its early stages remains a significant challenge in the present context and is a leading cause of death among women, primarily due to delayed identification. This paper presents a practical and accurate approach based on deep learning to identify breast cancer in cytology images.

Method: The analytical approach leverages knowledge from a related problem through a technique known as transfer learning. Convolutional neural networks (CNNs) are employed due to their remarkable performance on large datasets. Image classification architectures such as Google network (GoogleNet), Visual geographical group network (VGGNet), residual network (ResNet), and dense convolution network (DenseNet) are utilized in this approach. By applying transfer learning, the images are classified into two categories: those containing cancer cells and those without them. The performance of the proposed ensemble method is evaluated using a breast cytology image dataset.

Results: The results of our proposed ensemble framework outperform conventional CNN models in terms of precision, recall, and F1 measures, achieving an impressive 86% prediction accuracy. Visual representations of validation graphs for each classifier demonstrate that the ensemble framework surpasses the performance of pre-trained CNN architectures.

Conclusion: Combining the outcomes of conventional CNN architectures into an ensemble framework enhances early breast cancer detection, leading to a reduction in mortality through timely medical interventions.

Keywords: Biopsy, Mammography, Machine learning, Cytology, Deep learning

Introduction

In medical research, the study of microscopic images representing human parts and cells plays a significant part in understanding human activities and will make it easy for doctors to diagnose the disease. Significant deaths are seen

Corresponding Author: Joy Christy Antony Sami, PhD Department of Computer Science, School of Computing, SASTRA Deemed to be University, Thanjavur, India Email: joychristy001@gmail.com



Please cite this article as: Antony Sami JC, Arumugam U. An ensemble deep learning model for the detection and classification of breast cancer. Middle East J Cancer. 2024;15(1):40-51. doi: 10. 30476/mejc.2023.97317.1857. in women due to breast cancer, especially between 20 and 59 years old.¹ If it is diagnosed in the earlier stages, the survival rate will increase to 80%.² There are two ways of diagnosing breast cancer: 1) Mammography and 2) Biopsy. In mammography, breast cytology pictures of a particular type are used in the preliminary detection of cancer signs in women, and using mammography in such cases reduces the death rate. The biopsy is a more efficient and accurate diagnosis method when compared with mammography.^{3, 4} In this diagnosis method, a tissue sample is taken from an affected person and is inspected under a microscope by pathologists. Through this approach, doctors can identify different types of breast cancers, such as benign and malignant.⁵

Breast cancer is a rapidly increasing problem in women.^{6,7} Using technology like traditional machine learning, did not help doctors effectively treat patients in the earlier stages.^{8,9} Therefore, deep learning with CNN came into the picture, giving efficient and highly accurate results. The

Method No.	Method name	
1	Flipping	
2	Random rotation	
3	Scaling	
4	Shifting	

main objective of our project is to detect and classify breast cancer using a transfer learning methodology. In this process, the feature extraction is done by different CNN architectures like Inception Net (GoogleNet), Visual geographical group network (VGGNet), residual network (ResNet), and dense convolution network (DenseNet), which are ensembled to increase overall accuracy.

Materials and Methods

In the past few decades, there were several algorithms based on machine learning algorithms,¹⁰⁻¹² which are incapable of predicting the cancer productively in the preliminary stages. Many approaches are constantly being developed



Figure 1. This figure shows the proposed ensemble framework for breast cancer prediction.

CNN: Convolution neural network; VGGNet: Visual geographical group network; GoogleNet: Google network; ResNet: Residual network; DenseNet: Dense convolution network

	Precision	Recall	F1-score	
0	0.84	0.96	0.90	
1	0.83	0.54	0.66	
Micro average	0.84	0.84	0.84	
Macro average	0.84	0.75	0.78	
Weighted average	0.84	0.84	0.83	
Samples average	0.84	0.84	0.84	

in the medical field but fail to reach accurate results. Complex actions like noise removal from the pictures, performing mathematical operations on the pictures, and doing the traditional techniques may not be beneficial and will reduce

the accuracy and classification efficiency.
However, it is still in high demand to have an automatic system to increase efficiency.

• Doctors were unable to predict the presence of cancer in the early stages, and according to the research, if it is found in the earlier stages, there will be a chance to give proper treatment, and the death rate can be decreased rapidly.

To overcome the problems faced by traditional approaches and bring the idea of using deep learning models to the forefront, the approach aims to extract feature knowledge from cytology pictures and utilize it for classification. Recently, this deep learning approach based on CNN has seen upward growth in the accurate identification of cancer.^{13,14,15,16} Still, the model was created using a single CNN impact accuracy.

Methodology of proposed ensemble breast cancer classifier (EBCC)

In this analytical study, combining different CNN architectures with transfer learning enhances overall performance and can readily replace older approaches. The combination of various architectures, namely GoogleNet, ResNet, VGGNet, and DenseNet, has already been pretrained on ImageNet, leading to improved results.

In this approach, various attributes of the images are extracted by the architectures, such



Figure 2. This figure shows the methods used for preprocessing and augmentation of breast cytology images.

as GoogleNet, VGGNet, ResNet, and DenseNet. These attributes are merged using a fully connected layer, aiding doctors in early-stage disease prediction and treatment planning. This approach consistently yields superior accuracy compared with existing frameworks.

The methodology of the project consists of several steps, which include:

Figure 3. This figure shows the architecture of pre-trained visual geographical group convolutional neural network.

Figure 4. This figure shows the architecture of pre-trained Google convolutional neural network.

Figure 5. This figure shows the architecture of pre-trained residual convolutional neural network.

Figure 6. This figure shows the architecture of pre-trained dense neural network architecture.

	Precision	Recall	F1-score
0	0.82	0.97	0.89
1	0.87	0.46	0.60
Micro average	0.83	0.83	0.83
Macro average	0.85	0.72	0.75
Weighted average	0.84	0.83	0.81
Samples average	0.83	0.83	0.83

1. Dataset collection from KAGGLE

2. Preprocessing and data augmentation in the first phase

3. Implementation of trained CNN architectures

4. Classification with CNN using transfer learning

An illustrative depiction of the methodology is presented in figure 1.

a) Dataset collection

The dataset was sourced from KAGGLE, with a total size of 3.1GB. It comprises 198,738 IDC (-) image patches and 78,768 IDC (+) image patches. The training and testing dataset has been split into a 70:30 ratio. The file-name of each image contains meta-information about the dataset, including patient ID (e.g., 8863_idx5_x51_y 1251_class0), x-coordinate for patch cropping (x51), y-coordinate for patch cropping (y1251), and class (0 for Non-IDC/Benign and 1 for IDC/Malignant).

b) Image data augmentation and preprocessing

The initial objective of this approach is to detect and eliminate noise in input images. To achieve high accuracy, CNN requires substantial datasets. The performance of CNN deteriorates when smaller datasets are used, increasing the risk of over fitting. Image data augmentation is employed to expand the dataset by applying mathematical operations to the images to address this. Figure 2 illustrates the flow of image preprocessing and augmentation.

c) Feature extraction using trained architectures

The features extracted by the imageclassification architectures include:

- 1. VGGNet
- 2. GoogleNet
- 3. ResNet
- 4. DenseNet

Visual geometric group CNN (VGGNet)

VGGNet is similar to AlexNet in all aspects, with one notable exception: its additional convolutional layers. It incorporates 13 rectification, convolution, pooling, and fully connected layers. VGGNet employs 3×3 window size filters and 2×2 pooling networks. VGGNet exhibits exceptional efficiency compared to other CNN architectures, owing to its straightforward structure. To enhance model precision, VGGNet

	Precision	Recall	F1-score	
0	0.74	0.99	0.84	
1	0.74	0.10	0.17	
Micro average	0.74	0.74	0.74	
Macro average	0.74	0.54	0.51	
Weighted average	0.74	0.74	0.65	
Samples average	0.74	0.74	0.74	

adds 16 layers of weights. The architecture comprises 13 convolutional layers, three dense layers, and five pooling layers. Figure 3 illustrates the VGGNet architecture.

Google CNN (GoogleNet)

GoogleNet is comparatively smaller than other architectural counterparts, consisting of Pooling Layers, two fully connected layers, Rectified Linear Operation Layers, and three convolution layers for filtering purposes. The GoogleNet image classification architecture introduces an approach that integrates various convolution filters, potentially of random dimensions, into one filter. This action reduces the number of parameters and computational complexity. Inception modules in GoogleNet support feature detection across different scales of convolution with distinct filters. Figure 4 depicts the GoogleNet architecture.

Residual CNN (ResNet)

ResNet stands out as an intense ResNet among other CNN architectures, delivering superior results in image classification. ResNet combines multiple-sized convolution filters, effectively addressing the degradation problem. This approach also mitigates the extended training times associated with its deep structures. Figure 5 illustrates the ResNet architecture.

Dense CNN (DenseNet)

DenseNet shares similarities with ResNet but boasts some fundamental differences. DenseNet was developed to ameliorate the accuracy decline stemming from vanishing gradients often encountered in high-level neural networks. DenseNet is a critical component of classic networks. Figure 6 depicts the DenseNet architecture.

d) Transfer learning

In essence, CNNs demand substantial data for training. However, obtaining large datasets for specific problems can be challenging. Transfer learning has emerged as a powerful technique to provide an optimal solution. This approach is popular due to its knowledge-sharing capacity and adaptability to solving related problems. The proposed transfer learning method encompasses the following steps:

1. Selection of a pre-trained model

Figure 8. This figure shows the accuracy and loss analysis of visual geographical group convolutional neural network. acc: Accuracy

	Precision	Recall	F1-score	
0	0.80	0.98	0.88	
1	0.89	0.39	0.54	
Micro average	0.81	0.81	0.81	
Macro average	0.85	0.69	0.71	
Weighted average	0.83	0.81	0.79	
Samples average	0.81	0.81	0.81	

2. Assessment of problem size and its corresponding similarities

The proposed transfer learning method utilizes an appropriate training method with a suitable amount of target data, closely resembling the source data, thereby mitigating the risk of overfitting. In the proposed ensemble framework, the features generated by four CNN architectures - VGGNet, GoogleNet, ResNet, and DenseNet are aggregated and fine-tuned within the fully connected layer for the classification of benign and malignant breast cancer cells using average pooling classification.

Experimentation

To commence experimentation, the process starts with importing the requisite libraries. Subsequently, the data files are loaded into memory. Following this, Image Augmentation is performed on the data files, categorizing the images into two distinct groups: benign and malignant. Pre-trained architectures, namely VGGNet, GoogleNet, ResNet, DenseNet, and our proposed ensemble framework are then applied to the augmented images. The objective is to implement transfer learning and observe an improvement in efficiency. The overall accuracies of all the architectures are meticulously compared. *Image data preprocessing*

The cytology images undergo a series of preprocessing steps, including global-contrast normalization, histogram equalization, local normalization, and image de-noising.

Data Augmentation

To mitigate the risk of overfitting, Image Data Augmentation techniques are employed. This method leads to an expansion in the volume of images at our disposal. This augmentation process is pivotal in generating a sufficiently extensive training dataset. The significance of adding more images cannot be overstated, as a limited dataset hampers our model's ability to learn. By addressing this concern, image augmentation enhances generalization while minimizing overfitting issues. The techniques employed to augment the image dataset are outlined in table 1, encompassing Random rotation, Flipping, Image Shifting, and Image Pixel Scaling. Subsequently, the pre-trained architecture is applied, with transfer learning as

Figure 9. This figure shows the accuracy and loss analysis of google convolutionl neural network. acc: Accuracy

	Precision	Recall	F1-score	
0	0.84	0.98	0.90	
1	0.89	0.53	0.66	
Micro average	0.85	0.85	0.85	
Macro average	0.87	0.75	0.78	
Weighted average	0.86	0.85	0.84	
Samples average	0.85	0.85	0.85	

a crucial component to foster an upward trajectory in terms of efficiency.

Results

The proposed ensemble work is trained using four ideal CNN architectures: VGGNet, GoogleNet, ResNet, and DenseNet. It leverages transfer learning to combine the features extracted by each CNN architecture. The final decision is reached by comparing the results obtained from the single CNN with the combined features. Performance metrics, including recall, precision, and F-1 score, are computed for each classifier, and these metrics are then compared with the ensemble framework for breast cancer prediction. The results are detailed in tables 2-6 and figures 7-12.

Discussion

According to the current study, cytology image features' contribution to breast cancer prediction is less significant in conventional CNN models such as VGGNet, GoogleNet, ResNet, and DenseNet. However, the ensemble cytology features of these CNN architectures improve the classifier's performance concerning precision. recall, and F1 measures. Table 7 compares the ensemble framework's precision, recall, and F1 measures with pre-trained CNN architectures. The ensemble framework showed the highest precision (86%), recall (85%), and F1 measures (84%) values compared to the conventional CNN architectures. The various CNN architectures exhibited different accuracy rates in the classification and detection of breast cancer. Accuracy analysis of the four CNN architectures with the proposed ensemble framework is presented in table 8, and the corresponding pictorial representation is depicted in figure 13. As indicated in table 8, VGGNet, GoogleNet, ResNet, and DenseNet achieved classification accuracies of 84%, 84%, 74%, and 83%, respectively, while the proposed ensemble framework achieved an accuracy of 86%. Thus, the results demonstrate a significant improvement in detecting and classifying breast cancer using the ensemble framework compared with the other four CNN architectures.

Figure 10. This figure shows the accuracy and loss analysis of residual convolutional neural network. acc: Accuracy

Table 7. Comparative analysis of precision, recall and F1-measures of the ensemble framework with pretrained CNN architectures				
Conventional CNN models	Precision	Recall	F1-measure	
VGGNet	0.84	0.84	0.83	
GoogleNet	0.84	0.83	0.81	
ResNet	0.74	0.74	0.65	
DenseNet	0.83	0.81	0.79	
Proposed ensemble framework	0.86	0.85	0.84	

Table 7. Comparative analysis of precision	sion, recall and F1-measure	s of the ensemble framework v	with pretrained CNN architectures	
Conventional CNN models	Precision	Recall	F1-measure	
TIG GDI -	0.04	0.04	0.00	

CNN: Convolution neural network; ResNet: Residual network; VGGNet: Visual geographical group network; GoogleNet: Google network; ResNet: Residual network; DenseNet: Dense convolution network

This study encourages the use of 1) transfer learning with combined features, and 2) an ensemble of pre-trained models with appropriate adjustments of classification parameters to enhance the performance of the breast cancer classifier. However, the results indicate a notable improvement in the performance of the ensemble method in terms of prediction accuracy, precision, recall, and F1 measures, but time complexity emerges as an essential concern. This is because the need arises to train individual models and then ensemble the results for classification. Fortunately, this issue is mitigated by employing transfer learning, as the model utilizes the weights trained by other models rather than starting from scratch.

Xue et al.,¹⁷ proposed GAN-based augmentation for cancer prediction and used ResNet18 to classify benign and malignant cancerous images. However, the GAN-based augmentation method does not control the quality of synthetic images, and their CNN model could obtain only 71% prediction accuracy. This work employed four types of augmentation methods, including flipping, random rotation, scaling, and

shifting, to ensure the quality of augmented images. An ensemble of pre-trained models, such as VGGNet, GoogleNet, ResNet, and DenseNet, was utilized. The model achieves an 86% accuracy in breast cancer prediction.

Duggento et al.,18 have designed and validated an ad hoc CNN architecture for breast cancer lesion classification. The authors have explored 260 model architectures with different trainvalidation-test combinations to design a CNN model with reduced false negatives and a palatable accuracy of 71%.

Ragab et al.,¹⁹ have proposed a computeraided detection system for classifying benign and malignant breast tumors from mammography images. The authors used manual and automated tumor region selection methods and passed the processed images in two AlexNets. The authors have claimed that the performance of the automated threshold-based region selection method with AlexNet has obtained 80.5% accuracy than the manual region selection method, which has only produced 79% accuracy.

Fang et al.,²⁰ have proposed a restricted Boltzmann machine (RBM) based feature

Figure 11. This figure shows the accuracy and loss analysis of a dense convolutional neural network. acc: Accuracy

extraction method over 2D and 3D CNN architectures for breast cancer prediction. The authors have used magnetic resonance imaging images for the analysis. The authors have claimed that the accuracy (82.5%) obtained by the 3D CNN architecture is better than the accuracy (77%) obtained by 2D CNN as the prior analyses of the spatial structural information of the tumor.

Deniz et al.,²¹ have used transfer learning and deep feature extraction methods on AlexNet and VGG16 architectures to predict breast cancer. The authors have used histopathologic images for the analysis. The authors have passed the obtained features of the CNN architectures as an input to the support vector machine algorithm to classify the benign and malignant tumorous images. The authors have claimed that the combined features of AlexNet and VGG16 have obtained 84.8% accuracy.

The comparative analysis of the results obtained by the well-known CNN architectures discussed in previous studies,^{18,19,20,21} for breast cancer prediction have given accuracies of 71%, 80.5%, 82.5%, and 84.8%, respectively. At the same time, the results obtained by the proposed ensemble framework give an accuracy of 86%, which is higher than the well-known CNN architectures proposed for breast cancer prediction. Thus, the results show the strengthened performance of the proposed ensemble framework compared with similar methods.

Some of the limitations of the current study are considered as follows: The model uses images taken from a single source, and the lack of access

Conventional CININ models	Accuracy
VGGNet	84%
GoogleNet	84%
ResNet	74%
DenseNet	83%
Proposed ensemble framework	86%
CNN: Convolution neural network; ResNet: Residual	network; VGGNet: Visual

geographical group network; GoogleNet: Google network; DenseNet: Dense convolution network

to similar cytology images may impact the study's findings. However, the study uses an adequate number of images to build a robust breast cancer prediction model.

Conclusion

This paper proposes an ensemble deep-learning framework to achieve finer and more accurate results in identifying breast cancer from cytology pictures. The approach begins with applying preprocessing techniques to the dataset, effectively removing noise. Subsequently, image data augmentation is employed to increase the dataset's size, mitigating the limitations of CNNs.

Features are extracted from breast cytology pictures during this process utilizing various CNN architectures, including VGGNet, ResNet, GoogleNet, and DenseNet. These architectures are combined using transfer learning as an ensemble framework, improving accuracy. Finally, the individual accuracies of each architecture and the proposed ensemble framework are assessed, and the results are presented in validation graphs.

Figure 12. This figure shows the accuracy and Loss analysis of the proposed ensemble framework for breast cancer prediction. acc: Accuracy

Model accuracy

Figure 12 This figure shows the comparative analysis of ecourseiss of the managed appendix for broad concernment

Figure 13. This figure shows the comparative analysis of accuracies of the proposed ensemble framework for breast cancer prediction and other CNN models.

CNN: Convolution neural network; VGGNet: Visual geographical group network; GoogleNet: Google network; ResNet: Residual network; DenseNet: Dense convolution network

Conflict of Interest

None declared.

References

- West D, West V. Model selection for a medical diagnostic decision support system: a breast cancer detection case. *Artif Intell Med.* 2000;20(3):183-204. doi: 10.1016/s0933-3657(00)00063-4.
- Barba D, León-Sosa A, Lugo P, Suquillo D, Torres F, Surre F, et al. Breast cancer, screening and diagnostic tools: All you need to know. *Crit Rev Oncol Hematol.* 2021;157:103174. doi: 10.1016/j.critrevonc.2020. 103174.
- Wang F, Zhang S, Henderson LM. Adaptive decisionmaking of breast cancer mammography screening: a heuristic-based regression model. *Omega*. 2018;76(C):70-84. doi: 10.1016/j.omega.2017.05.001.
- 4. Stomper PC, Gelman RS. Mammography in symptomatic and asymptomatic patients. *Hematol Oncol Clin North Am.* 1989;3(4):611-40.
- Sharma D, Kumar R, Jain A. Breast cancer prediction based on neural networks and extra tree classifier using feature ensemble learning. Measurement: *Sensors*. 2022;24:100560. doi: 0.1016/j.measen.2022. 100560.
- Mizzi D, Allely C, Zarb F, Kelly J, Hogg P, McEntee M, et al. Examining the effectiveness of supplementary imaging modalities for breast cancer screening in women with dense breasts: A systematic review and meta-analysis. *Eur J Radiol.* 2022;154:110416. doi: 10.1016/j.ejrad.2022.110416.
- Wilson JM, Colebaugh CA, Flowers KM, Overstreet D, Edwards RR, Maixner W, et al. Applying the Rapid OPPERA Algorithm to Predict Persistent pain outcomes among a cohort of women undergoing breast cancer surgery. *J Pain*. 2022;23(12):2003-12. doi: 10.1016/j.jpain.2022.07.012.
- 8. Hessock M, Brewer T, Hutson S, Anderson J. Use of

a standardized tool to identify women at risk for hereditary breast and ovarian. *Nurs Womens Health*. 2021;25(3):187-97. doi: 10.1016/j.nwh.2021.03.008. Erratum in: *Nurs Womens Health*. 2021;25(5):322.

- Shaikh FJ, Rao DS. Prediction of cancer disease using machine learning approach. Materials Today: *Proceedings*. 2022;50:40-7. doi:10.1016/j.matpr.2021. 03.625.
- Sawssen B, Okba T. A novel machine learning approach for breast cancer diagnosis. *Measurement*. 2022;187:110233. doi:10.1016/j.measurement.2021. 110233.
- Gopal VN, Al-Turjman F, Kumar R, Anand L, Rajesh M. Feature selection and classification in breast cancer prediction using IoT and machine learning. *Measurement*. 2021;178:109442. doi:10.1016/j. measurement.2021.109442.
- Naji MA, Filali SE, Aarika K, Benlahmar EH, Abdelouhahid RA, Debauche O. Machine learning algorithms for breast cancer prediction and diagnosis. *Procedia Computer Science*. 2021;191:487-92. doi: 10.1016/j.procs.2021.07.062.
- Lu SY, Wang SH, Zhang YD. SAFNet: A deep spatial attention network with classifier fusion for breast cancer detection. *Comput Biol Med.* 2022;148:105812. doi: 10.1016/j.compbiomed.2022.105812.
- Gao F, Wu T, Li J, Zheng B, Ruan L, Shang D, et al. SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis. *Comput Med Imaging Graph*. 2018;70:53-62. doi: 10.1016/j.compmedimag. 2018.09.004.
- 15. Shadab SA, Ansari MA, Singh N, Verma A, Tripathi P, Mehrotra R. Detection of cancer from histopathology medical image data using ML with CNN ResNet-50 architecture. In: Agrawal R, Ansari MA, Anand RS, et al., editors. Computational intelligence in healthcare applications. Academic Press; 2022.p.237-254. doi:10.1016/B978-0-323-99031-8.00007-7.
- 16. Balkenende L, Teuwen J, Mann RM. Application of

deep learning in breast cancer imaging. *Semin Nucl Med.* 2022;52(5):584-96. doi: 10.1053/j.semnuclmed. 2022.02.003.

- Xue n, Zhou Q, Jiarong Ye, Rodney Long L, Antani S, Cornwell C, et al. Synthetic augmentation and feature-based filtering for improved cervical histopathology image classification. In: Shen, D, et al, editor. Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. MICCAI 2019. Lecture Notes in Computer Science. Springer, Cham; 2019.p.387-396. doi: 10.1007/978-3-030-32239-7_43.
- Duggento A, Aiello M, Cavaliere C, Cascella GL, Cascella D, Conte G, et al. An ad hoc random initialization deep neural network architecture for discriminating malignant breast cancer lesions in mammographic images. *Contrast Media Mol Imaging*. 2019;2019:5982834. doi: 10.1155/2019/5982834.
- Ragab DA, Sharkas M, Marshall S, Ren J. Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ*. 2019;7:e6201. doi: 10.7717/peerj.6201.
- Fang Y, Zhao J, Hu L, Ying X, Pan Y, Wang, X. Image classification toward breast cancer using deeply-learned quality features. *J Vis Commun Image Represent*. 2019;64:102609. doi: 10.1016/j.jvcir.2019.102609.
- Deniz E, Şengür A, Kadiroğlu Z, Guo Y, Bajaj V, Budak Ü. Transfer learning based histopathologic image classification for breast cancer detection. *Health Inf Sci Syst.* 2018;6(1):18. doi: 10.1007/s13755-018-0057-x.