



ORIGINAL

## A Progressive UNDML Framework Model for Breast Cancer Diagnosis and Classification

### Un modelo marco progresivo UNDML para el diagnóstico y clasificación del cáncer de mama

G. Meenalochini<sup>1</sup> , D. Amutha Guka<sup>1</sup> , Ramkumar Sivasakthivel<sup>2</sup> , Manikandan Rajagopal<sup>3</sup> 

<sup>1</sup>School of Computing, Kalasalingam Academy of Research and Education, Krishnankoil, Tamil Nadu.

<sup>2</sup>Department of Computer Science, School of Sciences, CHRIST (Deemed to be University), Bangalore, India

<sup>3</sup>Lean Operations and Systems, School of Business and Management, CHRIST (Deemed to be University), Bangalore, India.

**Cite as:** Meenalochini G, Guka DA, Sivasakthivel R, Rajagopal M. A Progressive UNDML Framework Model for Breast Cancer Diagnosis and Classification. Data and Metadata 2024;3:198. <https://doi.org/10.56294/dm2024198>.

Submitted: 13-10-2023

Revised: 01-12-2023

Accepted: 06-02-2024

Published: 07-02-2024

Editor: Prof. Dr. Javier González Argote 

#### ABSTRACT

According to recent research, it is studied that the second most common cause of death for women worldwide is breast cancer. Since it can be incredibly difficult to determine the true cause of breast cancer, early diagnosis is crucial to lowering the disease's fatality rate. Early cancer detection raises the chance of survival by up to 8 %. Radiologists look for irregularities in breast images collected from mammograms, X-rays, or MRI scans. Radiologists of all levels struggle to identify features like lumps, masses, and microcalcifications, which leads to high false-positive and false-negative rates. Recent developments in deep learning and image processing give rise to some optimism for the creation of improved applications for the early diagnosis of breast cancer. A methodological study was carried out in which a new Deep U-Net Segmentation based Convolutional Neural Network, named UNDML framework is developed for identifying and categorizing breast anomalies. This framework involves the operations of preprocessing, quality enhancement, feature extraction, segmentation, and classification. Preprocessing is carried out in this case to enhance the quality of the breast picture input. Consequently, the Deep U-net segmentation methodology is applied to accurately segment the breast image for improving the cancer detection rate. Finally, the CNN mechanism is utilized to categorize the class of breast cancer. To validate the performance of this method, an extensive simulation and comparative analysis have been performed in this work. The obtained results demonstrate that the UNDML mechanism outperforms the other models with increased tumor detection rate and accuracy.

**Keywords:** Breast Cancer Detection; Preprocessing; Feature Extraction; Mammograms; Segmentation; CNN.

#### RESUMEN

Según investigaciones recientes, se estudia que la segunda causa de muerte entre mujeres a nivel mundial es el cáncer de mama. Dado que puede resultar increíblemente difícil determinar la verdadera causa del cáncer de mama, el diagnóstico temprano es crucial para reducir la tasa de mortalidad de la enfermedad. La detección temprana del cáncer aumenta las posibilidades de supervivencia hasta en un 8 %. Los radiólogos buscan irregularidades en las imágenes de los senos obtenidas de mamografías, radiografías o resonancias magnéticas. Los radiólogos de todos los niveles luchan por identificar características como bultos, masas y microcalcificaciones, lo que conduce a altas tasas de falsos positivos y falsos negativos. Los recientes avances en el aprendizaje profundo y el procesamiento de imágenes generan cierto optimismo en cuanto a la creación de aplicaciones mejoradas para el diagnóstico precoz del cáncer de mama. Se llevó a cabo un estudio metodológico en el que se desarrolló una nueva red neuronal convolucional basada en segmentación

profunda U-Net, denominada marco UNDML, para identificar y categorizar anomalías mamarias. Este marco implica las operaciones de preprocesamiento, mejora de la calidad, extracción de características, segmentación y clasificación. En este caso se lleva a cabo un preprocesamiento para mejorar la calidad de la imagen de la mama. En consecuencia, se aplica la metodología de segmentación Deep U-net para segmentar con precisión la imagen de la mama y mejorar la tasa de detección del cáncer. Finalmente, el mecanismo CNN se utiliza para categorizar la clase de cáncer de mama. Para validar el rendimiento de este método, en este trabajo se ha realizado una simulación exhaustiva y un análisis comparativo. Los resultados obtenidos demuestran que el mecanismo UNDML supera a los otros modelos con una mayor tasa de detección de tumores y precisión.

**Palabras clave:** Detección del Cáncer de Mama; Preprocesamiento; Extracción de características; Mamografías; Segmentación; CNN.

## INTRODUCTION

According to research, breast cancer is the second most common cause of death for women worldwide. Generally, breast cancer is caused by the abnormal growth of a mass of tissues made up of malignant cells. Breast cancer is the most prevalent type of cancer among women, with over 2,1 million new cases being identified each year, according to the World Health Organization (WHO). According to estimates, 15 % of all female cancer fatalities—or 627,000—in 2020 were attributable to breast cancer.<sup>(1,2,3)</sup>

Mammography is currently the most popular and well approved method of breast cancer screening. It is less effective, less likely to cause microscopic tumors, and does not signal breast cancer in women under 40 with dense breasts. In cases of large breasts, contrast-enhanced (CE) digital mammography, which is less common due to the high radiation exposure expenses, provides a better level of diagnostic accuracy than mammography and ultrasonography.<sup>(4,5,6)</sup> Magnetic Resonance Imaging (MRI) can detect small lesions that mammography was unable to detect. Its high cost and low level of specificity, however, raise the possibility of overdiagnosis.<sup>(7,8,9)</sup> The diagnosis and classification of breast tumors using various conventional techniques have been studied and worked on, but no specific type of machine learning has been employed.

For deep learning networks to be trained efficiently and with good performance, a large number of tagged images are frequently required.<sup>(10,11,12)</sup> Unfortunately, due to the limited patient load, medical databases frequently only include a few images. The little amount of the breast imaging datasets that are currently available makes the training effort difficult. Recent studies have shown that Computer-Aided Diagnosis systems are quite effective at identifying and categorizing breast cancer in mammography pictures. They can sometimes do better than humans, lowering the mortality rate for both men and women with breast cancer.

The stage of extracting and selecting key features is crucial for classical Machine Learning (ML) classifiers including Support Vector Machine (SVM), K- Nearest Neighbor (KNN), Naive Bayes (NB), Linear Regression (LR), Decision Tree (DT), and others.<sup>(13,14)</sup> However, if this phase is not carried out correctly, performance and accuracy as a whole could decrease.<sup>(15,16,17)</sup> However, in order to complete a classification assignment, these classifiers need to be fed purposely created information. Additionally, recent studies have shown that CNNs can perform several image processing tasks, such as picture analysis, object classification, segmentation, and fragmentation, with incredibly accurate results.<sup>(18,19,20)</sup> Yet, the transferability of a number of layers in pre-trained deep CNNs, such as AlexNet, GoogLeNet, OverFeat, and ResNet, and their reuse for a new job, were shown to be a potential solution to the problem of a lack of training data and high performance outcomes.

The following are the research's main contributions and goals:

- The input breast image is preprocessed to enhance its quality by reducing noise using a novel Contrast Limited Adaptive Histogram Equalization (CLAHE) mechanism.
- To determine the extent and density of breast cancer, the texture feature extraction model is used.
- To improve the accuracy of breast cancer prediction rates, the breast region is carefully segmented using a deep U-Net segmentation technology.
- Less overfitting is achieved by using a Convolutional Neural Network (CNN) mechanism to accurately detect and classify the kind of breast cancer.

The effectiveness and results of the suggested UNDML mechanism were tested and evaluated in this work by a detailed simulation and comparative analysis. The following sections are created from the remaining parts of this article: The comprehensive literature evaluation for the methods currently in use for the diagnosis and classification of breast cancer is provided in Section 2. The proposed UNDML framework is well explained in Section 3 together with its appropriate workflow and descriptions. Additionally, Section 4 compares and validates the outcomes of the proposed UNDML process using a variety of evaluation criteria. Finally, Section 5 summarizes the entire study along with the conclusions and suggested next steps.

### Related works

This section, which looks at the list of available ML and DL models used for breast cancer diagnosis and classification, presents a comprehensive literature review. It also looks at the advantages and restrictions of applying each model.

A transfer learning technique was used by Saber *et al.* to identify and categorize breast cancer.<sup>(21)</sup> By appropriately pre-training the classifier, this effort aimed to save training time and improve classifier performance. Here, the raw photos' quality has been enhanced using morphological techniques and histogram equalization. Additionally, it aims to apply a transfer learning approach to address the overfitting issue. In this framework, the data preprocessing was mainly concentrated on attaining better classification results, which includes the operations of noise elimination, histogram equalization, morphological operation, segmentation and resizing. However, it could be very difficult to understand the system model.

In order to validate the various mammogram-based breast cancer detection and classification approaches, Gardezi *et al.* conducted a systematic review.<sup>(22)</sup> The goal of this research was to look at the list of machine learning and deep learning methods utilized for breast cancer classification and detection. Preprocessing, segmentation, feature extraction, and classification are all included in this system. Using picture enhancement techniques to boost contrast and legibility enhances the mammography's quality.<sup>(23)</sup> The technique may identify mammographic lesions with poor visibility and contrast by amplifying them. Raising the picture quality of low contrast mammograms would be the primary goal of mammography enhancement.<sup>(24)</sup> Tissues adjacent frequently concealed minor anomalies in low-contrast regions, leading to an inaccurate diagnosis.

The following are some of the segmentation approaches covered in this study: Local thresholding

- Global thresholding
- Region growing & clustering
- Template matching
- Edge detection

Classification is the last stage in evaluating whether a lesion is benign or malignant. If it is found that the area has cancer, further classification is done to determine whether it is benign or malignant. The classification phase itself is significantly influenced by other intermediate procedures, particularly segmentation and feature extraction. Some of the widely used classifiers in the classification of breast cancer include SVM, KNN, NB, and DT.<sup>(25)</sup> The performance of the classifier can be improved by eliminating the extraneous traits and maintaining only the most discriminative ones. Many scientists have used DL techniques to examine medical photos. The usability of many training samples to obtain the pertinent feature maps of the DL is what determines the majority of its effectiveness.

A thorough examination of the various segmentation and classification approaches utilized for breast cancer detection was published by Krithiga, *et al.* in their article.<sup>(2)</sup> This work's original contribution was to investigate several ML/DL methods for creating a reliable breast cancer diagnosis system. Typically, one of the most important methods for dimensionality reduction, which is essential in applications involving medical imaging, is feature extraction. For breast cancer prediction algorithms, shape-based, color-intensity-based, and texture-based features were progressively utilized. In earlier research, breast masses from already-identified mammograms were categorized using transfer learning. The CNN does both the feature extraction and the classification because the model is an end-to-end deep learning model. This point needs to be emphasized. Also, the mammography image's area of interest was subjected to the feature extraction technique. In order to discriminate between benign and malignant masses as well as cases of calcifications in breast mammograms, the RBF-Based SVM classifier was then given these features in this work. Characteristics from the complete breast image, however, were recovered. A pre-trained OverFeat CNN that was taught for object detection in real-world pictures was utilized in a different study to identify nodules in computed tomography scans.<sup>(26)</sup> It is still up for debate as to how accurate the CNN's classification will be when the area of interest diverges from the original domain in which it was trained. This study aims to demonstrate the significance of domain similarity for CNN transfer learning. Hence, a CNN was trained in the same domain as the study's area of interest. The authors' neural network model uses BPNN and multilayered perceptron networks.<sup>(27)</sup> This model categorizes the simulation's result as benign or malignant. Included are the parameters for weight adjustment and bias values.

Karthick, *et al.* trained the network using a multilayered perceptron network and four back-propagation training methods, including quasi-Newton, gradient descent with momentum and adaptive learning, Levenberg-Marquardt, and robust back propagation.<sup>(28)</sup> The steepest descent back propagation is used to assess the effectiveness of other neural networks. The best accuracy rate was achieved by the Levenberg-Marquardt method utilizing MLP, which was 94,11 %. The Wisconsin Diagnostic Breast Cancer (WDBC) Dataset was subjected to the SVM with a Recursive Feature Elimination (RFE) approach. SVM is used to categorize the dataset in this instance, and independent principal component analysis (PCA) on the same dataset is used to reduce the dimensionality.

Anooj, *et al.* developed a Clinical Decision Support System using weighted fuzzy rules (CDSS).<sup>(29)</sup> The two key

processes in this approach are creating fuzzy rules and creating a decision support system based on fuzzy rules. Fuzzy-rule-based decision support systems rely their decisions on the fuzzy rules that the systems generate in order to ensure better projections. Fuzzy rules are developed based on historical data to improve learning. It made adjustments to a weighted fuzzy rule in accordance with how significant the attributes were.

Mafarja, et al. developed and presented a separate fuzzy-rough nearest neighbor based classification model.<sup>(30)</sup> Instance selection, feature selection, and categorization make up the three stages of this system. To eliminate erroneous and ambiguous instances, the fuzzy-rough instance selection method is combined with weak gamma evaluators. A re-ranking technique and a consistency-based feature selection strategy are used to quickly scan the search space for appropriate enumerations. The classification method is based on the fuzzy-rough closest neighbor. This method functioned better than the previous fuzzy ones.

## METHODS

This research is a methodological article whose objective is the development of a unique framework for accurate diagnosis of breast cancer with increased detection rates and decreased false positives. Deep learning-based segmentation and classification techniques are employed in this work to maximize computing efficiency.

A mammogram in the Dataset typically consists of the label-containing backdrop region and the mammary region. The breast area typically contains the segmented target region (breast masses), whereas the background region typically contains noise information. Both the segmentation of the breast masses and the network's ability to train may be affected by these disturbances. Hence, pre-processing is required to separate the breast region from the background region and remove the label before the image is delivered to the network. The input dataset's quality is improved in this study by reducing noise and artifacts using the CLAHE normalization approach. The size and density of the tumor are then calculated by extracting the textures using the feature extraction model. Additionally, to improve classifier performance, the breast image is carefully segmented using the U-Net segmentation technique. Eventually, the CNN is employed to accurately predict and classify the tumor category. The proposed UNDML framework's overall workflow is depicted in figure1, and it contains the procedures listed below:

- CLAHE based preprocessing
- Texture feature extraction
- Deep U-Net segmentation
- CNN based classification

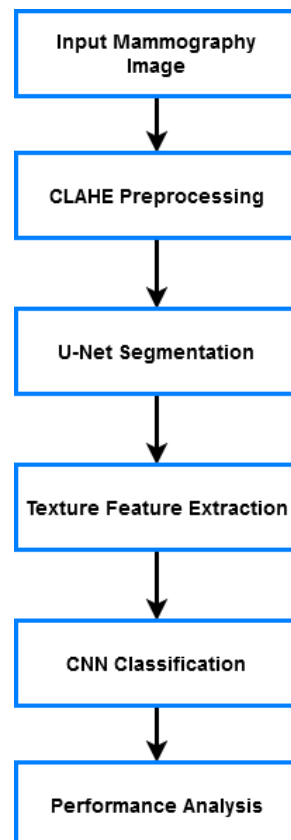


Figure 1. Workflow of the proposed UNDML framework

### Preprocessing

In tissue imaging, the pre-processing stage is essential for eliminating various noise patterns. For CNN to work more accurately, large data sets are required. Additionally, CNN's performance suffers when dealing with small data sets due to over-fitting. It shows that the network performs poorly on test data while performing admirably on training data.

Different CNN designs are used for feature extraction, and for classification tasks, they are combined into a fully connected layer. The combined features may contain many features that have been retrieved from a single descriptor; these features may represent shape descriptors like roundness, sphericity, density, etc. The goal of this procedure is to boost the possibility that an anomaly will be correctly diagnosed while also enhancing the image quality. The breast tissue is highlighted more clearly because of improved object separation and the elimination of noise from the image capturing process. To choose the best filtering method for this situation, several filtering approaches have been examined and applied to the images.

In fact, the Contrast Limited Adaptive Histogram Equalization (CLAHE) based preprocessing model has been selected for this work since it operates on a small area of the breast image rather than the entire image and applies normalizing to each of them. The CLAHE is a variation on adaptive histogram equalization, a technique used in image processing to enhance local contrast by adjusting the brightness of each pixel in the image and applying local histogram equalization to surrounding pixel regions. The CLAHE algorithm tackles the noise over amplification problem of the histogram equalization by restricting the cumulative density function slope when computing histogram equalization.

### Segmentation

The U-Net segmentation, which consists of a decoder (a channel that grows on the right side) and an encoder, is typically employed to enhance the network's capability to segment medical images (a path that contracts on the left). The representational architecture of a CNN, which also serves as the network's feature extractor and oversees downsampling, is utilized by the encoder. Each downsampling module consists of two convolutional layers (3X3, exclude padding) and a maximum pool layer (2X2, stride D 2).

The number of feature channels grows as the size of the feature maps reduces with each downsampling. The up-sampling process is handled by the decoder. A de-convolutional layer (2X2) and two convolution layers are present in each up-sampling module (3X3, excluding padding). Using skip connections, the shallow layer and deep layer feature maps are joined. The number of feature channels is reduced in half while the image size increases for each up-sampling. The final convolutional layer (1X1) maps a 64-channel feature vector to the necessary classification outcome and then forecasts the outcomes for each individual pixel. After each convolutional layer, Rectified Linear Unit (ReLU) activation functions are used. The network can simultaneously account for the contribution of shallow information and deep information at the final output because it concatenates the output of the shallow layer with the output of the deep layer.

#### Algorithm 1 - Deep U-Net Segmentation

Input: Mammography image, no of layer, window size and training set;

Output: Segmented cancer area;

Step 1: Input the mammography image;

Step 2: Set the window size to choose the pixel;

Step 3: Set the no of layer of CNN;

Step 4: Insert the training set;

Step 5: Perform the following layer operations;

- Get the pixel from the input image based on the window size
- Decoding and downsample the image based on the layer
- ReLU function for conventional layer
- Polling the Conventional layer

Step 6: Segment the cancer area

### Classification

The science of computer vision has significantly advanced owing to the convolutional neural network (CNN), which consists of numerous layers of neural computing connections with minimum systematic processing. The CNN global learning strategy mimics how the mammal visual system is set up. A convolutional layer's primary objective is to recognize edges, curves, and other visual components like local patterns. The system has been taught how to configure convolutions, which are complex filter operators. Neighboring pixels from a particular pixel are multiplied in this process using a certain kernel array.

This method also mimics how the kernels extract visual elements like borders and hue. Deep learning may dynamically learn a hierarchical representation of models, from low to high level functions, and then select

the most important features for a particular model by using a deep CNN architecture to imitate the naturally occurring artificial neural multilayered network. It may be essential to estimate millions of weight parameters because the deep CNN architecture frequently uses numerous layers in the neural network, which demands a large number of data samples for model development and parameter setup.

The sequential model, on which this architecture is built, enables the model to develop sequential network layers from input to output in the proper sequence. Convolution operations are carried out by filters in the convolution layer while they scan the input image's dimensions. The filter size and stride, which describe the distance between succeeding receptive filters, are two of its hyper-parameters. The output is referred to as an activation map or feature map. A first 2D convolutional layer is implemented in order to process the input breast images.

#### Algorithm 2 - CNN based Classification

Input: Segmented cancer area, no layer, window size, no of hidden layer, no of output layer;

Output: training set, Classification result, performance graph;

Step 1: Input the segmented image;

Step 2: Set the no of layer;

Step 3: Get the pixel based on the window size;

Step 4: Find the ReLU for the conventional layer and extract the polling layer;

Step 5: Convert to the vector as input layer of CNN;

Step 6: Set the no of hidden layer;

Step 7: Set the no of output layer based on the output class

Step 8: Get the training set for classification;

Step 9: Compare the input to the training set to find the classification result;

Step 10: Find the class value for all in the test data to compare to generate the performance graph of the algorithm;

#### RESULTS

The most popular evaluation methodologies in the field of medical imaging were utilized to evaluate the effectiveness of the proposed hybrid method for distinguishing benign from malignant instances. These evaluation methods include confusion matrix, sensitivity, specificity, and classification accuracy.

#### Confusion Matrix

Information regarding the actual classes and anticipated classes of the hybrid approach categorization are shown in this matrix. Differences between actual and projected classes may be positive or negative.

figure 2 shows the confusion matrix for the suggested hybrid method. A true negative (TN) situation materializes when both the anticipated and actual cases are benign. False negative (FN) situations occur when the expected case is benign but the actual case is cancerous. When a benign condition is found even though a malignant condition was expected, this is known as a false positive (FP) circumstance. When a malignant case is predicted and one occurs, this is referred to as a true positive (TP) case.

#### Sensitivity

Also known as the TP rate, which is calculating the proportion of accurate positive forecasts to actual positive cases. It is estimated by using the following equation:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (1)$$

#### Specificity

It is also referred to as the TN rate, which determines the percentage of cases that were projected to be negative but turned out to be positive. This is estimated by using the following equation:

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (2)$$

#### Accuracy

It indicates the proportion of correctly predicted cases among all the cases, which is computed by using the following model:

$$Accuracy = \frac{TN+TP}{TN+TP+FP+FN} \quad (3)$$

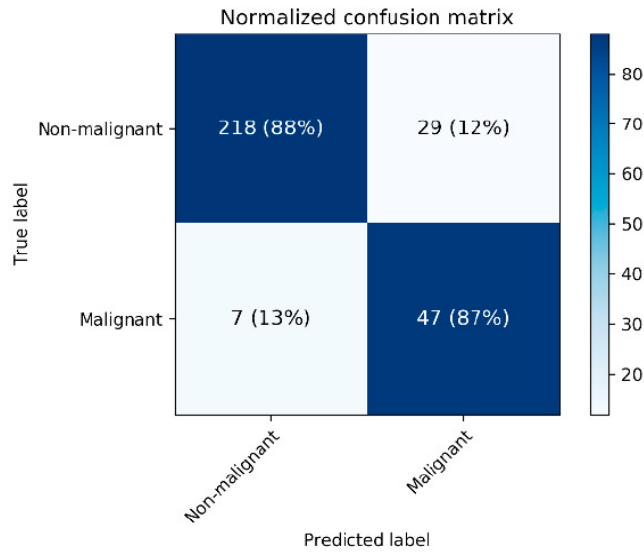


Figure 2. Confusion matrix

Figure 3 and figure 4 are graphical representations of table 1, which compares the accuracy and error rate of existing and proposed classification models used for breast cancer detection and classification. The research shows that the proposed model has a greatly reduced error rate and a higher degree of precision. The training and testing processes of the classifier are greatly improved using a deep U-Net segmentation model, resulting in improved accuracy and decreased error rate. Additionally, the proposed UNDML model produces better outcomes than the ones currently in use.

Algorithm	Accuracy	Error rate
Proposed	98,25	1,75
Hybrid of K-means Gaussian Mixture Model	95,50	4,5
Gaussian mixture Model	93,80	6,2
K-means	71,00	29,00
Growth region hand selection	63,00	37,00
Growth region FCM-GA selection	71,00	29,00

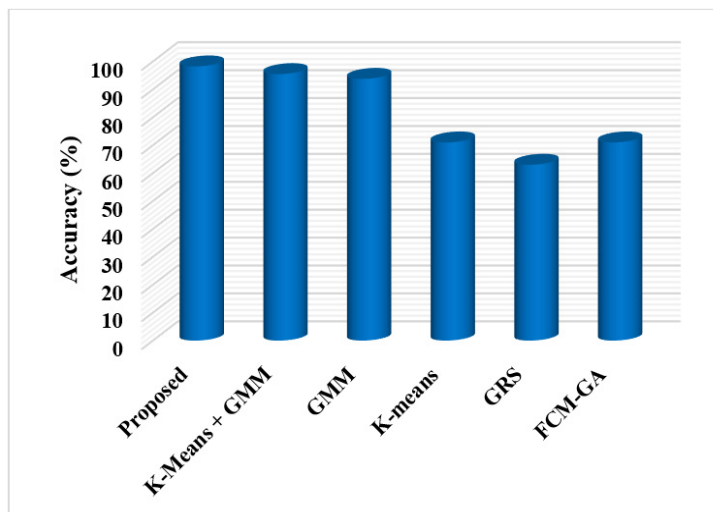


Figure 2. Accuracy

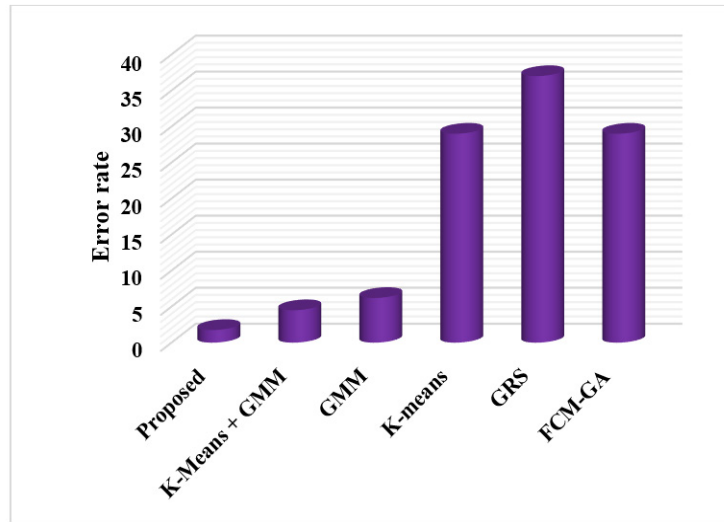


Figure 3. Error rate

Based on the metrics of accuracy, sensitivity, and specificity, table 2 compares the proposed and existing deep learning models. These criteria are mostly used to assess how well the classifier predicts cancer with a high rate of detection. The proposed framework performs significantly better when compared to the other models. As a result, table 3 and figure 5 show a comprehensive comparison of the existing and suggested deep learning-based breast cancer detection systems.

The obtained results suggest that, when compared to the baseline models, the proposed UNDML framework might effectively increase the performance of breast cancer diagnosis.

Techniques	Accuracy	Sensitivity	Specificity
SAE	86	84,3	83,6
DBN	89	88,5	87,5
CNN	95	93,2	94,5
DTCNN	98	96,8	95,7
GA-CNN	98,5	97,5	96,6
UNDML	99,5	98,3	97,2

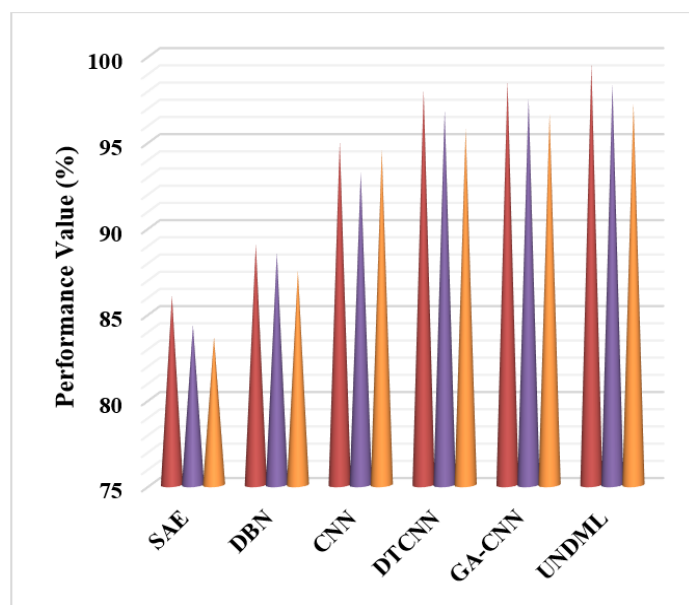


Figure 4. Comparative analysis with the deep learning models



Parameters	AHEE-CDLS-CNN	GA-CNN	UNDML
Sensitivity	0,975	0,9938	0,99
Specificity	0,966	0,984	0,98
Precision	0,9885	0,9876	0,98
F1_Score	0,9952	0,9907	0,99
MCC	0,9802	0,9786	0,98
Accuracy	0,985	0,98	0,985
Kappa Coefficient	0,9821	0,975	0,98
Error rate	0,015	0,02	0,05
FPR	0,012	0,012	0,014

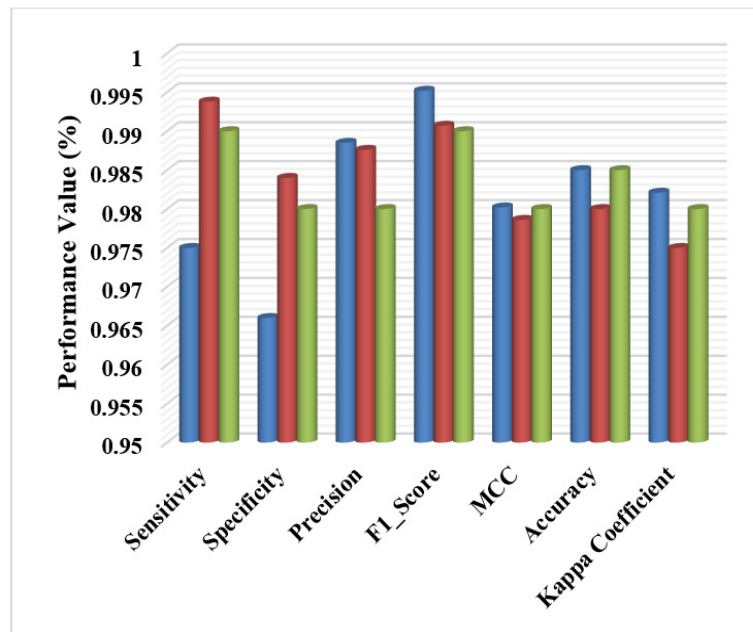


Figure 5. Overall analysis

## CONCLUSION

Today's population is afflicted by numerous modern diseases. Breast cancer is one of the most common and deadly illnesses that has been expanding throughout many nations. Lack of knowledge and delayed disease detection will be the main causes of rising death rates. Computer-aided diagnostics will be the best way to make an accurate diagnosis for all types of patients. The CAD system will be an excellent tool for practitioners to review patient records and make the best judgments possible, but it will not totally replace qualified medical experts. This work suggests a hybrid method to potentially detect breast cancer in mammograms by combining a CNN and U-Net segmentation model. When training a CNN from scratch to do a certain task, especially when there are not enough medical images available, this method takes advantage of the concept of domain adaptation. In this system, the input breast photo is preprocessed to enhance its quality by reducing noise using a sophisticated CLAHE method. The texture feature extraction model is used to estimate the breast cancer's size and density. To improve the accuracy of breast cancer prediction rates, the breast region is carefully segmented using a deep U-Net segmentation method. The type of breast cancer is then accurately predicted and classified using a CNN technique with reduced overfitting. The effectiveness and results of the suggested UNDML mechanism were tested and evaluated in this work by a detailed simulation and comparative analysis. The proposed UNDML outperforms the other model with higher performance values, according to the observed data.

## REFERENCES

1. Zahra Rezaei. A review on image-based approaches for breast cancer detection, segmentation, and classification, *Expert Systems with Applications*, vol. 182, p. 115204, 2021, <https://doi.org/10.1016/j.eswa.2021.115204>.

2. Krithiga R, Geetha P. Breast Cancer Detection, Segmentation and Classification on Histopathology Images Analysis: A Systematic Review. Archives of Computational Methods in Engineering, vol. 28, pp. 2607-2619, 2020, <https://doi.org/10.1007/s11831-020-09470-w>.

3. Soulamy K B, Kaabouch N, Saidi M N, Tamtaoui A. Breast cancer: One-stage automated detection, segmentation, and classification of digital mammograms using UNet model based-semantic segmentation, Biomedical Signal Processing and Control, vol. 66, p. 102481, 2021, <https://doi.org/10.1016/j.bspc.2021.102481>.

4. Zhang G, Zhao K, Hong Y, Qiu X, Zhang K, Wei B. SHA-MTL: soft and hard attention multi-task learning for automated breast cancer ultrasound image segmentation and classification, International Journal of Computer Assisted Radiology and Surgery, vol. 16, pp. 1719-1725, 2021, doi: 10.1007/s11548-021-02445-7.

5. Chanda P B, Sarkar S K. Detection and classification of breast cancer in mammographic images using efficient image segmentation technique, Advances in control, signal processing and energy systems, ed: Springer, pp. 107-117, 2020, doi: 10.1007/978-981-32-9346-5\_9.

6. Jahangeer G S, Rajkumar T D. Early detection of breast cancer using hybrid of series network and VGG-16. Multimedia Tools and Applications, 80, 7853-7886, 2020, doi:10.1007/s11042-020-09914-2.

7. Salama W M, Aly M H. Deep learning in mammography images segmentation and classification: Automated CNN approach, Alexandria Engineering Journal, vol. 60, pp. 4701-4709, 2021, <https://doi.org/10.1016/j.aej.2021.03.048>.

8. Chowdhary CL, Mittal M, Pattanaik P, Marszalek Z. An Efficient Segmentation and Classification System in Medical Images Using Intuitionist Possibilistic Fuzzy C-Mean Clustering and Fuzzy SVM Algorithm. Sensors, vol. 20, p. 3903, 2020, <https://doi.org/10.3390/s20143903>.

9. Zahoor S, Lali IU, Khan MA, Javed K, Mehmood W. Breast Cancer Detection and Classification using Traditional Computer Vision Techniques: A Comprehensive Review. Current Medical Imaging, vol. 16, pp. 1187-1200, 2020, doi:10.2174/1573405616666200406110547.

10. Ilesanmi AE, Chaumrattanakul U, Makhanov SS. Methods for the segmentation and classification of breast ultrasound images: a review. Journal of Ultrasound, vol. 24, pp. 367-382, 2021, doi:10.1007/s40477-020-00557-5.

11. Galli A, Gravina M, Marrone S, Piantadosi G, Sansone M, Sansone C. Evaluating impacts of motion correction on deep learning approaches for breast DCE-MRI segmentation and classification. International Conference on Computer Analysis of Images and Patterns, pp. 294-304, 2019, DOI: 10.1007/978-3-030-29891-3\_26.

12. Zeebaree D Q, Haron H, Abdulazeez A M, Zebari D A. Machine learning and Region Growing for Breast Cancer Segmentation. 2019 International Conference on Advanced Science and Engineering (ICOASE), 88-93, 2019, DOI:10.1109/ICOASE.2019.8723832.

13. Amin J, Sharif M, Fernandes S L, Wang S H, Saba T, Khan A R. Breast microscopic cancer segmentation and classification using unique 4-qubit-quantum model. Microscopy Research and Technique, vol. 85, pp. 1926-1936, 2022, <https://doi.org/10.1002/jemt.24054>.

14. Ramesh S, Sasikala S, Gomathi S, Geetha V, Anbumani V. Segmentation and classification of breast cancer using novel deep learning architecture. Neural Computing and Applications, pp. 1-13, 2022, <https://doi.org/10.1007/s00521-022-07230-4>.

15. Khan S U, Islam N, Jan Z, Haseeb K, Shah S I A, Hanif M. A machine learning-based approach for the segmentation and classification of malignant cells in breast cytology images using gray level co-occurrence matrix (GLCM) and support vector machine (SVM), Neural Computing and Applications, vol. 34, pp. 8365-8372, 2022, DOI: 10.1007/s00521-021-05697-1.

16. Hamed G, Marey M A, Amin S E, Tolba M F. Deep Learning in Breast Cancer Detection and Classification, International Conferences on Artificial Intelligence and Computer Vision, pp. 322-333, 2020, DOI:10.1007/978-3-030-44289-7\_30.

17. Michael E, Ma H, Li H, Kulwa F, Li J. Breast Cancer Segmentation Methods: Current Status and Future Potentials, *BioMed Research International*, 2021, DOI:10.1155/2021/9962109.
18. Zhou Y, Chen H, Li Y, et al. Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images, *Med Image Anal*, vol. 70, p. 101918, 2021, doi:10.1016/j.media.2020.101918
19. Dizaj SB, Valizadeh P. Breast cancer segmentation and classification in ultrasound images using convolutional neural network, *Research Square*, 2021, DOI: 10.21203/rs.3.rs-952669/v1.
20. Tsochatzidis L, Koutla P, Costaridou L, Pratikakis I. Integrating segmentation information into CNN for breast cancer diagnosis of mammographic masses, *Comput Methods Programs Biomed*, vol. 200, p. 105913, 2021, doi:10.1016/j.cmpb.2020.105913.
21. Saber A, Sakr M, Abo-Seida O M, Keshk A, Chen H. A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique, *IEEE Access*, vol. 9, pp. 71194-71209, 2021, DOI: 10.1109/ACCESS.2021.3079204.
22. Gardezi SJS, Elazab A, Lei B, Wang T. Breast Cancer Detection and Diagnosis Using Mammographic Data: Systematic Review, *Journal of medical Internet research*, vol. 21, p. e14464, 2019, doi:10.2196/14464.
23. Karimi Jafarbigloo S, Danyali H. Nuclear atypia grading in breast cancer histopathological images based on CNN feature extraction and LSTM classification, *CAAI Transactions on Intelligence Technology*. 2021 Sep;6(4):426-39. <http://dx.doi.org/10.1049/cit2.12061>.
24. Civilibal S, Çevik K K, Bozkurt A. A deep learning approach for automatic detection, segmentation and classification of breast lesions from thermal images, *Expert Systems with Applications*, vol. 212, p. 118774, 2023, <http://doi.org/10.1016/j.eswa.2022.118774>.
25. Benagoune K, Al Masry Z, Devalland C, Valmary-degano S, Zerhouni N, Mouss L, Data Labeling Impact on Deep Learning Models in Digital Pathology: a Breast Cancer Case Stud, *Intelligent Vision in Healthcare*, ed: Springer, pp. 117-129, 2022, [http://dx.doi.org/10.1007/978-981-16-7771-7\\_10](http://dx.doi.org/10.1007/978-981-16-7771-7_10).
26. Sirazitdinov I, Kholiavchenko M, Mustafaev T, Yuan Y, Kuleev R, Ibragimov B. Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database, *Computers & electrical engineering*, vol. 78, pp. 388-399, 2019, <https://doi.org/10.1016/J.COMPELECENG.2019.08.004>.
27. Sarosa S J A, Utamingrum F, Bachtiar F A. Breast cancer classification using GLCM and BPNN, *Int J Adv Soft Comput Appl*, vol. 11, 2019.
28. Karthik S, Srinivasa Perumal R, Chandra Mouli P. Breast cancer classification using deep neural networks, *Knowledge computing and its applications*, ed: Springer, pp. 227-241, 2018, [https://doi.org/10.1007/978-981-10-6680-1\\_12](https://doi.org/10.1007/978-981-10-6680-1_12).
29. Anooj P. Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules, *Journal of King Saud University-Computer and Information Sciences*, vol. 24, pp. 27-40, 2012, <https://doi.org/10.1016/j.jksuci.2011.09.002>.
30. Mafarja M, Sabar N R. Rank based binary particle swarm optimisation for feature selection in classification, *Proceedings of the 2nd International Conference on Future Networks and Distributed Systems*, pp. 1-6, 2018, <https://doi.org/10.1145/3231053.3231072>.

#### **FINANCING**

The authors did not receive financing for the development of this research.

#### **CONFLICT OF INTEREST**

None.

#### **AUTHORSHIP CONTRIBUTION**

*Conceptualization:* G.Meenalochini, D.Amutha Guka, S.Ramkumar, R.Manikandan.

*Research:* G.Meenalochini, D.Amutha Guka.

*Drafting - original draft:* G.Meenalochini, D.Amutha Guka, S.Ramkumar, R.Manikandan.

*Writing - proofreading and editing:* G.Meenalochini, D.Amutha Guka.