Al-Qadisiyah JournalFor Engineering Sciences 19 (2026) 059 – 070

Contents lists available at: http://qu.edu.iq

Al-Qadisiyah Journal for Engineering Sciences

Journal homepage: https://qjes.qu.edu.iq

Research Paper

Breast cancer detection using deep learning techniques: An investigation using the CBIS-DDSM dataset and customized neural network model, ResNetV2 and YOLO

Qutaiba Humadi Mohammed [1] ✉, ***Noor Alhuda Kh Ibrahim*** [2], ***Anupama Namburu*** [3] ***and Chaitanya Konda*** [4]

[1]*Media Technology and Communications Eng. Department, College of Engineering, University of Information Technology and Communications, Baghdad, Iraq.*
[2]*Department of Basic Sciences, College of Nursing, University of Baghdad, Baghdad, Iraq.*
[3]*School of Engineering, Jawaharlal Nehru University, New Delhi 110067, India.*
[4]*Department of CSE, College of Engineering and Technology, Acharya Nagarjuna University, Nagarjuna Nagar, Guntur, 522510, India.*

ARTICLE INFO

Article history:
Received 26 December 2024
Received in revised form 09 April 2025
Accepted 08 October 2025

keyword:
Breast Cancer
CNN
Data augmentation
Deep Learning
ResNetV2
YOLO

ABSTRACT

The current work investigates the application of deep learning methodologies for detecting and classifying tumors in mammograms. A comprehensive analysis was conducted by evaluating various pre-trained neural network architectures to identify the model that delivers optimal performance. Among the tested architectures, Residual Neural Networks and Densely Connected Convolutional Networks were explored extensively. Data preprocessing, including data augmentation, was a critical step due to the limited availability of public medical imaging datasets. This process ensured diversity in the data and improved model robustness. This paper evaluates deep-learning models for mammogram tumor detection. Experiments without the AdamW optimizer and horizontal flip showed overfitting and low precision (below 40%). Densely Connected Convolutional Networks achieved high precision but exhibited overfitting with noisy validation loss curves. Adding AdamW and horizontal flip reduced overfitting but lowered overall performance. In classification, the model detected tumors in 9 out of 16 images, showing potential but requiring improvement. The model's ability to detect calcification tumors enhances robustness. YOLO network metrics were modest, reflecting the task's complexity, but results were acceptable for tumor classification challenges.

1. Introduction

Breast cancer is the cancer with the highest mortality rate in the female population worldwide [1]. According to the WHO data in 2023, Iraqi women were diagnosed with this disease, thus becoming the second cancer with the highest incidence in our country [2]. It is estimated that 1 in 8 Iraq women will be diagnosed at some point in their lives [3] Survival depends largely on the stage of the tumor. According to Rawla et al (2019) [4], survival in stage I is more than 98%, however in stage IV survival drops to 24%. Therefore, early detection is essential for a good prognosis, thus using less invasive therapies and achieving a better quality of life for patients. The main medical test for detecting this cancer is mammography. As presented in the literature [5], more than 65% of the female population over 40 years of age have had at least one mammogram in the two years following 2017. Numerous clinical trials [6] have shown that this diagnostic technique reduces the mortality rate by 20% to 35%. However, it is not a test with 100% reliability. According to the work of Mao et al. (2024) [7], 20% of patients who underwent a mammogram received a false positive, which means having to perform more invasive medical tests to confirm the initial diagnosis. On the other hand, as presented in Moshina et al (2024) [8], around 35% of breast cancer cases are false negatives. This allows the disease to progress without being treated earlier, thus increasing the patient's risk of mortality. In short, mammography is a very valuable tool in the early detection of breast cancer, but it has several limitations.

1.1 Review of literature

In the work Tasnim et al. (202) [9] carried out, a Convolutional Neural Network (CNN) model has been developed using the MobileNetV2 classifier to detect cancer cells in the colon. In the article presented by Adyasha Sahu, In the article proposed by Kavyashree et al. (2022) [10], work is presented in which different pre-trained Densely Connected Convolutional Networks (DenseNet) models are used to detect and classify oral cancer in images. In addition, a CNN is developed that has four convolutional layers, and three maximum pooling layers, and uses the ReLU activation function, the binary Crosentropy loss function, and the Adam optimization algorithm. The work presented by Ankita Sinha et al. (2023) [11] proposes a breast cancer detection model based on the regions of interest in the images. In it, they use the Residual Neural Network (ResNet) architecture, specifically the pre-trained models ResNet18 and ResNet50. In addition, transfer learning is used to obtain more precise results. To do this, the additional layers of ResNet are eliminated. Instead, a fully connected layer and a dropout are included where parameters are randomly removed. A ReLU activation function follows this. In the following work, Samantha et al. (2023) [12] carried out in which different YOLO models were used to detect tumors. In developing the work, The work with the pre-trained models $YOLO_{V6}$, $YOLO_{V7}$, and $YOLO_{V8}$.

*Corresponding Author.
E-mail address: Qutaiba.humadi@uoitc.edu.iq ; Tel: (+964) 790-120 3976 (Qutaiba Mohammed)

https://doi.org/10.30772/qjes.2025.155626.1429

Nomenclature

CAD	Computer Assisted Diagnosis	*DL*	Deep Learning
CBIS	Curated Breast Imaging Subset of DDSM	*MIAS*	Mammographic Image Analysis Society
CNN	Convolutional Neural Network	*MLO*	Medio Lateral Oblique
CSV	Comma-Separated Values	*ResNet*	Residual Neural Network
DDSM	Digital Database for Screening Mammography	*RGB*	Red, Green, Blue
DenseNet	Densely Connected Convolutional Networks	*ROI*	Region of Interest

The latter, in particular, has high performance thanks to its advanced architecture, high efficiency, computing time, and the variety of pre-trained models it offers. Regarding development, the Roboflow library is used to generate data augmentation. Adyasha Sahu et al. (2024) [13], work is presented they use three pre-trained network models for detecting breast cancer using clinical images. Specifically, they combine the AlexNet, ResNet and MobileNetV2 models to achieve optimal results. Recent evolutions in DL techniques [14] are improving image classification accuracy. Algorithms based on CNN [15] applied to tumor detection are achieving great success in this area. Similarly, they present a more efficient method with a lower monetary cost. CNNs can replace the learning algorithms used to locate abnormal regions in Computer Assisted Diagnosis (CAD) images [16]. Their ability to maintain the spatial characteristics of the image makes CNNs ideal for mammography analysis [17]. The novelty of this paper resides in its Region of Interest (ROI)-focused analysis, which shows that training on clipped tumor patches improves model precision, with ResNet surpassing DenseNet. It combats data scarcity by utilizing data augmentation and preprocessing to ensure diverse and robust training. The entire experimental framework, which includes four scenarios with 100 epochs each, demonstrates the impact of design decisions and overfitting. The paper underlines the importance of task-specific architecture design and suggests further research to extend datasets, integrate other medical pictures, and use transfer learning to improve tumor detection in mammography. Based on the above importance, the work aims to identify which architecture provides the most optimal results for this task. Despite not achieving the best performance, YOLO was included as part of the broader analysis to evaluate its potential in tumor detection.

2. Methodology

The following section delves into the different resources used throughout the development of the work. These include both the software tools and the dataset and the different architectures used in the experiments.

CNN: Convolutional neural networks (CNNs) are the most commonly utilized architectures in DL [18, 19]. This is because of how well it finds and learns features from the network input photos. The Convolutional Layer, Grouping Layer, Activation Layer, Flattening Layer, and Fully Connected Layer are the layers the network architecture builds upon and applies to its internal structure.

DenseNet: Another alternative within CNNs that does not present connectivity between sequential layers is the DenseNet architecture. Zhu and Newsam (2022) [20] propose a dense connection between the layers of the so-called dense blocks. This is achieved by connecting each current layer with the input to the following layers in the block. This way, a model that suffers less overfitting is obtained the compactness.

ResNet: The main problem encountered with high-depth neural networks is the gradient's disappearance [21]. This problem is based on the fact that as the network training progresses, the values of the gradients decrease significantly. This decrease occurs in the backpropagation algorithms' weight update phase. All of this results in the deeper layers of the network being unable to update the weights correctly. This is what causes most deep networks not to obtain the expected results. On the other hand, another problem when the layers of neural networks are increased is that the precision stops rising and may even decrease. All this has shown that increasing the depth of the network does not imply a better model obtained Fig. 1 and Fig. 2. Consequently, the ResNet model is presented in the article proposed by He et al. (2022) [22]. This is a ResNet model. In it, the connections of the layers are designed without following linearity. In this way, this concept abandons the traditional connection of CNN layers presented in Fig. 1. The residual connections proposed by this architecture are based on adding the input of the residual block with the output of the block itself, this process is seen in Fig. 2. The formula for this operation is set out in Eq. 1.

YOLO: YOLO is a region-type neural network which is capable of classifying and detecting objects in the image in a single pass. To do this, the image is divided into a series of grids [23]. For each cell, the probability that each object present belongs to each class of the problem is predicted. Its architecture is formulated from a simple CNN. For the detection phase, convolutional layers are applied, which are responsible for extracting the most significant features of the image. These layers are followed by the fully connected layers, which calculate the classification. The network architecture is presented in Fig. 3. YOLO stands out above all for its ability to detect objects of different sizes and shapes. Likewise, it stands out that its computing time decreases compared to other types of CNN.

Figure 1. Linear connections of the layers of a CNN.

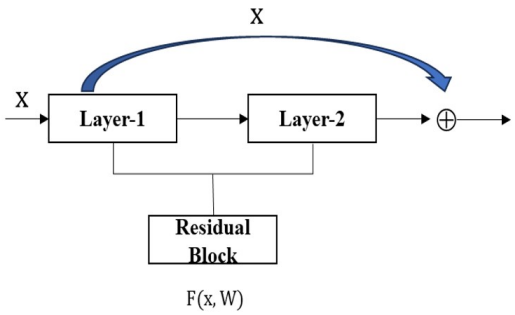


Figure 2. ResNet Residual Block Structure [24].

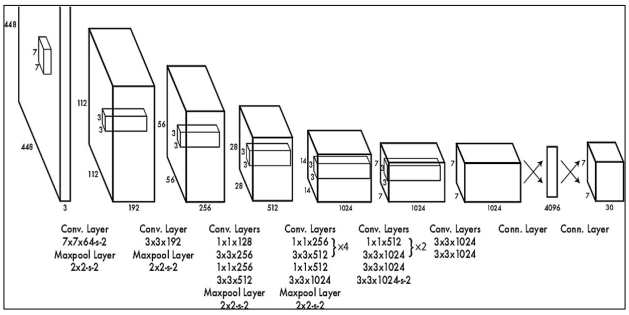


Figure 3. Architecture of the YOLO neural network [25].

2.1 Dataset

Regarding the dataset used for the development of the work, the Curated Breast Imaging Subset of DDSM (CBIS-DDSM): Breast Cancer Image Dataset [26]. This is an extended version of the Digital Database for Screening Mammography (DDSM). In this dataset, there are a total of 10,237 mammography images, obtained from 1,566 participants. Three types are distinguished among them; the complete mammograms, the cropped images in the area of interest and finally the segmented ROI images. You can see the different images in Fig. 4. Therefore, both a cropped image and an ROI image are obtained from each complete mammogram. Thus, the total number of images of each type is expressed in Table 1. In this way, the dataset has 3 classes for tumor classification Malignant (48.3%), Benign (43.8%) and Benign without call back (7.9%). For greater simplification of the data, in this work, the benign and the benign without callback will be unified into the same class. Furthermore, in the dataset the tumors are separated by the type of pathology, these being tumor calcifications and tumor masses. However, in this work, all the experiments carried out have been carried out with data belonging to tumors originating

from a mass. Therefore, the data set provides two different types of Comma-Separated Values (CSV) files, those referring to information on breast tumor masses and those on calcifications.

Table 1. Number of images of each type of class.

Image type	No. of samples
Complete mammogram	2856
Cropped	3557
ROI	3237
Invalid	576

Regarding the information provided by the CSV files of the tumors coming from the mass, the following data is provided:

2.1.1 Breast Density

They have 4 types; those with fatty density, disseminated fibro angular, heterogeneously dense and extremely dense.

2.1.2 Side of the Breast

Provides information about which breast is the one in the mammogram.

2.1.3 Image View

Referring to the two types of breast projections performed in the medical test. The Medio Lateral Oblique (MLO) is where the image is captured from an oblique view of the chest [27]. On the other hand, there is the Craniocaudal (CC) where the chest is compressed between two plates.

2.1.4 Type of Anomaly

mass or calcification. In each CSV it will correspond to a type, the ones used in this work have been the mass CSV so this field is always mass.

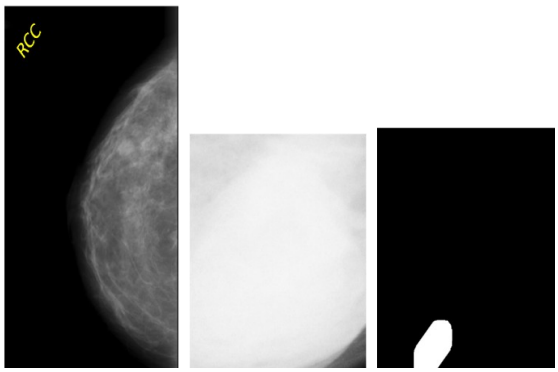
2.1.5 Shape of the Mass

The mass classified as prickly occurs where the edges of the mass are irregular. The circumscribed, on the other hand, are the masses that have defined and rounded edges. The poorly defined type where the edges cannot be seen correctly. Those are hidden either by breast tissue or by the mammogram itself. Finally, the micro-lobed ones have lobes on the edges of the mass.

2.1.6 Pathology

malignant, benign and benign without a callback.

In this work, by only focusing on the task of tumor classification for the detection of breast cancer, only information regarding the pathology has been used. Thus, this field has been used as the labels of each image with which the network has been trained. Therefore, as explained previously, the dataset used in this work covers various information regarding the tumors present in mammograms. This gives it great potential since it can be used for different purposes than those chosen in this work. Although the dataset is made up in most cases of images from the CBIS-DDSM, in some experiments other images have been included. These are images extracted from the Mammographic Image Analysis Society (MIAS) data set, Suckling et al. (2015) [28]. This has a total of 322 images with tumor masses, calcifications and benign masses. The highlight of this work is that images are also provided where the breast does not have any tumor.



(a) Complete Mammography (b) Cropped mammogram (c) Segmented ROI

Figure 4. Dataset images of three set.

3. Development of the work

In the following section, all the steps carried out in the development of the work will be explained. Fig. 5 shows graphically the stages that make up the work. As presented, first the detailed work of the dataset is carried out. When training neural networks, a thorough understanding of the data with which you are going to work is essential. In this way, more suitable models are obtained. Subsequently, the preprocessing of said data is carried out. This guarantees that all images that enter the network have the same format characteristics. In addition, the training of the model is optimized. Once the appropriate data is available, and move on to training the neural network. This training provides a resulting model to which an evaluation must be carried out in order to obtain information about its operation.

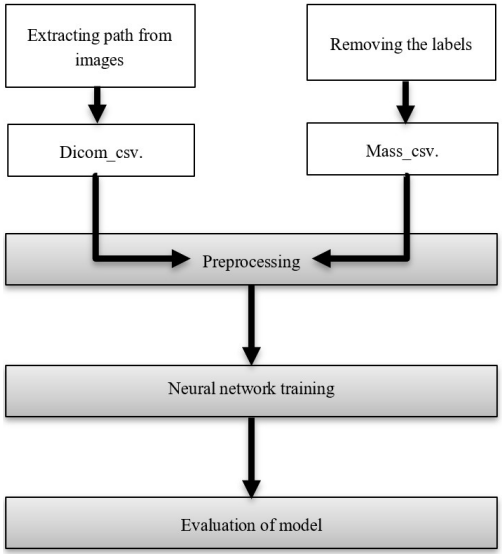


Figure 5. Diagram of the work stages.

3.1 Extraction of paths and labels from images

This process has been expensive due to different issues. Firstly, it was necessary to carry out exhaustive work on the different CSVs provided in the dataset. This had 5 types of files:

- Dicom: in this CSV, all the information regarding the medical image is presented, such as the paths to the images in Dicom format. In this case, the path to the images in jpeg format has also been considered significant.
- Mass Test: the same information explained in the CSV mass train is provided.
- Calc Train: information is provided regarding the breast density, breast side, image view, and type of anomaly fields; it has personalized information on the type of tumor due to calcification. Information is provided on the type of calcification and its distribution.
- Calc Test: the same information is provided as in the CSV calc train.

In this case, only the Dicom, Mass Train and Mass Test files have been used since, as mentioned above, only the images of mass tumors have been used. The path to the JPEG images is extracted from the Dicom file for further processing. To obtain the labels of each image, the information located in the Mass Train and Mass Test CSV is needed. Therefore, the difficulty lies in knowing which label should be chosen for each image. To this end, the problem has been addressed in the following way. A certain part of the path to the images in dicom format has been extracted. This part of the chain is also present in the path to the images in JPEG format. Subsequently, the dicom file was searched for the correspondence between this specific string and the paths to the JPEG images. In this way, the relationship between each row of the dicom and that of the mass is established. Likewise, it is possible to obtain, on the one hand, the path to the image, from the dicom file, and the label of the image from the mass file. This entire process has been automated using a Python program and its procedure can be seen in Fig. 6. On the other hand, it should be said that the dataset proposes a distribution over the files to be used in both training and testing. However, it has been decided to merge the images from both CSVs and subsequently perform a different fragmentation of the data.

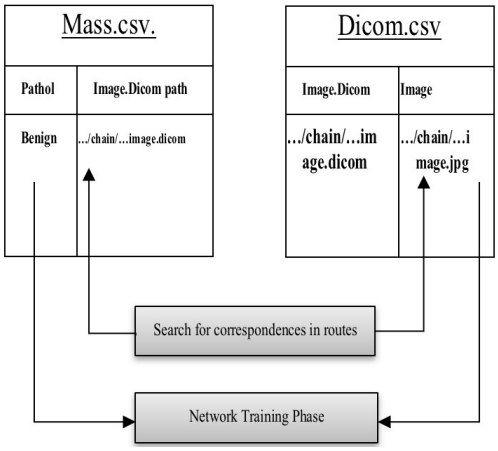


Figure 6. Diagram of the stages of route extraction.

3.2 Pre-processing

The quality of the data used to train neural networks directly influences the classification capacity of the model. Likewise, correct pre-processing of the data is capable of reducing the errors obtained in the model, as well as reducing training time. At the same time, it is possible to generate more robust models which allows them to be able to work adequately under a variety of conditions, such as the diversity that exists in the morphology of the breasts. In relation to the pre-processing of the images of this work, in all the experiments carried out the following techniques have been used to prepare the input data to the network.

3.2.1 Conversion to Red, Green, Blue (RGB) color space [29, 30]

Although mammograms are images extracted under a gray scale, it has been decided to apply this conversion. The reason is that there are a large number of neural network architectures that are implemented to work with RGB images.

3.2.2 Image resizing [31]

Medical images have a very high resolution due to the need to be able to accurately capture the smallest detail. This generates the need for robust hardware to be able to process the data during network training. However, by reducing the dimensionality of the images, in this case to 224x224 pixels, it is possible to carry out the training in a more computationally efficient way.

3.2.3 Normalization of Pixel Intensity [31]

Normalization of all the pixel intensities of the images has been carried out in order to ensure that all the data with which the work is in the same range so that all the characteristics that are detected in the image influence the training in an equitable way. On the other hand, in certain trained models it has been decided to increase the data to analyse whether better results are obtained in this way. To achieve this, the following image magnification techniques have been applied:

3.2.4 Rotation Transformations

It is useful to rotate the images since in mammograms in most cases the test is performed from different angles. In this way, the model is encouraged to learn to classify tumors regardless of their orientation.

3.2.5 Displacement Along the Vertical and Horizontal Axes

In this case a benefit analogous to that discussed in the previous transformation is provided. Different locations of the breast are proposed in mammography.

3.2.6 Flip Alonge the Vertical and Horizontal Axes

By reflecting the image for the vertical axis, the model obtained may be able to detect the type of tumor regardless of which breast is present. Regarding the horizontal axis, it provides an increase in the diversity of the data.

3.2.7 Image Bias

By applying this type of deformation to the image, the objective is to obtain more robust models against possible alterations in the clinical image. However, in the model evaluation, it is concluded that not all of these transformations are appropriate for the data being worked with in the work. This is because many of them distort the image and/or can cause the tumor to disappear from the area being analysed, which generates an inadequate model for the classification task.

3.2.8 Data Augmentation with Rotation, Panning, Zooming, Skewing and Flipping [32]

The first type of data augmentation that has been proposed is based on applying the transformations and values of extensive Data Augmentation Transformations are Rotation (40o), Vertical and Horizontal Scrolling (20%); Zoom (20%) and Biased (20th).

3.2.9 Data Augmentation with Rotation, Zoom and Flip [32]

In the second type of data augmentation used in the experiments, we chose not to apply vertical and horizontal displacement. This is because this transformation could, where the tumor is located at the edges of the image, discard the tumor from the area of the image to be analysed by the network. On the other hand, it is chosen to suppress the bias to simplify the training process. The selected values for the rest of the applied transformations are Rotation (40º) and zoom (20%).

3.2.10 Image Cropping from the Bounding Box [33]

At the same time, a different pre-processing of the data has been carried out. This consists of extracting the bounding box of the tumor from the ROI images provided in the dataset, which are already binarized. In this way, the tumor is enclosed under the white area, and the rest of the breast is represented by the black region. With this segmentation, image processing is applied to extract the coordinates of the pixels of the corners of the bounding box. These coordinates are used to crop the image of the complete mammogram to focus it on the area of interest and thus observe if better results are obtained. This process is detailed in Fig. 7.

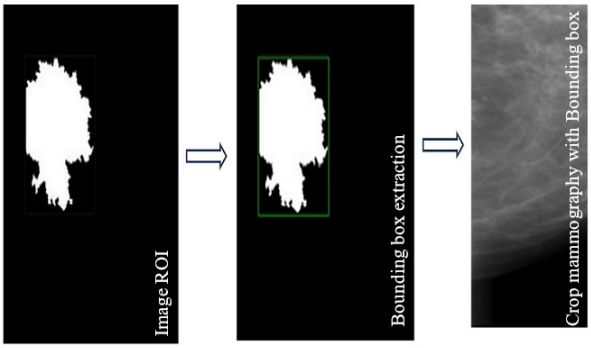


Figure 7. Diagram of the stages of the Trimming process with Bounding Box.

3.3 Training of the neural network

In the development of this work, different experiments have been carried out applying various neural network architectures. In this section the selected architectures will be explained in detail. CNN network architecture and full Mammography images The first experiment that has been carried out is the training of a CNN from the complete mammography images. Since the tumors are small masses, the ResNet network architecture has been used. This architecture does not suffer loss of information in the data which is crucial for the task of tumor classification. Among the different pre-trained models offered by the ResNet architecture, the Inception-ResNetV2 model has been chosen [34]. This presents a series of improvements in its architecture compared to the original ResNet. Inception-ResNetv2 places the ReLU activation function before the sum of the residual connections, unlike ResNet which does it after. This significantly improves the problem of gradient disappearance during training, as well as achieving more stable models. On the other hand, the Nesterov Adam optimizer is used. This is an optimizer that tends to converge faster than others. No other parameters have been established since the already pre-trained model has been used, which is responsible for predefining them in its architecture. Apart from the CNN network architecture used, 3 different types of preprocessing have been applied to the input images. There are 3 types of data augmentation.

3.3.1 network architecture with Mammography images cropped from the ROI

The second type of training carried out has been done from mammography images cropped starting from the ROI. In this way, data is provided to the network that is more focused on the tumor regions of the breast. Regarding the trained network architectures, experiments have been carried out with two types of architectures. The first is the InceptionResNetV2 explained previously. Neither the structure nor the parameters have been modified with respect to what was used in the previous experiment. On the other hand, we have worked

with the DenseNet architecture. Of the models present in said architecture, the DenseNet121 model has been chosen [35]. This is a deeper network model where a total of 121 layers are present. These are divided into 5 convolutional layers, 4 dense blocks of 6, 12, 16 and 24 layers respectively. In turn, each layer of the blocks is made up of two convolutions, which makes a total of 121 layers. A modification has been added to this model by eliminating the classification layer established by the pre-trained model. To do this, a flattening layer is incorporated at the model output to obtain the one-dimensional vector of the extracted features. Next, a dense layer of 128 neurons and ReLU activation function are added. This allows the learning of more complex features based on the output provided by the pre-trained model. Finally, another dense layer is added but in this case with 2 neurons, one for each class of the problem. Therefore, this last layer is where the classification of the input images takes place. In the same way, as it has been used in the ResNet architecture, the Nesterov Adam optimizer is applied in order to obtain a more appropriate model.

3.3.2 YOLO

In the experiments carried out with the YOLO architecture, the latest available pre-trained model, YOLOv8, has been used. Two types of experiments have been applied to this model. In the first, the network has been trained without a customized adjustment of the model's hyperparameters. Once the results were obtained and analysed, it was decided to use an optimizer to make the results more efficient and improved. The chosen optimizer is AdamW which is used to improve the overfitting problem [36]. In the same way, the horizontal flip parameter is adjusted in order to apply said transformation to the set of images to increase the data. This is intended to provide robustness to the model since the tumor can occur in both breasts, which implies that the mammogram can be oriented towards both sides. Another strategy that has been carried out to improve the results obtained in the experimentation is to include background images in the dataset. These are images that contain information related to the context in which the objects themselves will be detected, but without including any object in said image. In this case, the background images that have been used are mammograms in which there is no tumor presence. In this way, the model learns the characteristics of a breast, and thus, when it is going to learn about those with a tumor, it can detect them better.

3.4 Model evaluation

In the development of this work, different architectures have been applied, which leads to analyzing of each model obtained with different metrics. To evaluate the performance of the models obtained from CNN architectures, metrics such as accuracy and loss are used, both for the training set and the validation set [37]. In addition, an analysis of the confusion matrix is carried out to evaluate the model based on the classification it obtains of the different classes. Accuracy represents the proportion of all correct predictions, both positive and negative, over the total predictions made. Furthermore, to carry out the analysis of the evolution of the YOLO network, another metric system is used, apart from those mentioned above. To do this, the following metrics must be explained:

Accuracy: The percentage of instances correctly classified also known as precision, is the percentage of instances that the classifier predicted correctly. Corresponds to the rate of correctly classified positive and negative examples. This metric is calculated according to the following formula, Eq. 1.

$$Accuracy = \frac{TP+TN}{TP} \tag{1}$$

Precision: Precision is a measure that was originally introduced with the aim of measuring the effectiveness of a search engine in returning information considered relevant. In this specific case, precision is the fraction of documents retrieved by a search engine that are equally relevant. In turn, in the evaluation of classifiers, precision is defined as in Eq. 2.

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

Recall: In classifier evaluation, recall, sensitivity and true positive rate (TPR) have in common the fact that they are defined according to the same formula, Eq. 3.

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

F1 score: F1 score measures the effectiveness of a classifier, namely in terms of precision and recall. It is possible to define an F-Measure that assigns arbitrary weight to both precision and recall. This measure is known as F1-Measure, since it assigns equal importance to these two metrics. The formula for F1-score (harmonic mean between precision and recall) is as follows, Eq. 4.

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

Precision-Recall Curve: Presents the behavior of precision as the recall varies Average Precision (mAP): Calculates the area under the Precision-Recall curve. The closer its value is to 1, the better the model will perform. Extracts the average of correct detections that the model is capable of obtaining. The loss is defined as the sum of the errors predicted by the model. The higher this value is, the worse the model performs. On the other hand, the confusion matrix is a comparative matrix between the real classification values and the values predicted by the model. This can be seen in Table 2.

Table 2. Example of confusion matrix.

	Real Values	
Predicted Values	True Positive (TP)	False (FP)
	False Negative (FN)	True Positive (TP)

In the same way, graphs have been obtained that show the evolution of said metrics used during training. This makes it easier to visualize the model's learning process and helps detect possible problems such as overfitting and/or overtraining. Grad-CAM used to CNN models can disclose the particular areas of mammography pictures that the model focuses on while detecting cancers. Grad-CAM (Gradient-weighted Class Activation Mapping) is a visualization technique that identifies image regions that are most important to a CNN prediction.

4. Results and discussion

An exhaustive analysis of the different architectures applied in the development of the work is presented. The experiments carried out are divided into:

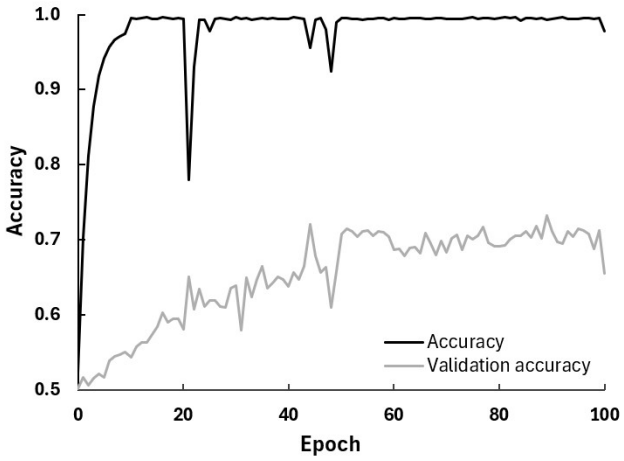
- Model trained from a CNN architecture using complete mammograms
- Model trained from a CNN architecture using the mammograms cropped in the area of interest where the tumor is located.
- Model trained from the YOLO architecture using complete mammograms

4.1 CNN network architecture with complete mammography images

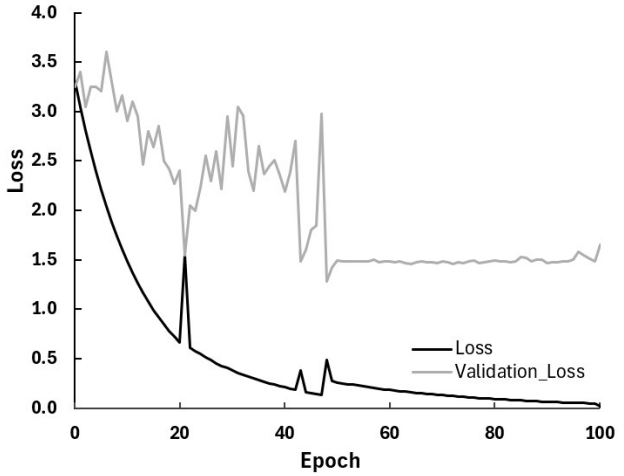
In this instance, three experiments have been conducted to assess the model's functionality in various scenarios. In the first, full mammography images were combined with the ResNetV2 architecture. The identical architecture is used in the second experiment, but the set of photographs first had some data augmented. The third experiment modifies the data augmentation strategies to acquire an alternative approach. Lastly, an experiment was conducted with minimum modifications to the transformations' values to draw conclusions. Training over 100 epochs has been completed in each of the four situations.

4.1.1 Experiment Without Data Augmentation

It has been possible to verify how the model obtained without increasing the data has been the one that has given the best performance, reaching a validation accuracy of 70%. This is set out in Table 3. A high accuracy train value is observed, very close to 1, which reflects that the model has a high classification capacity. The accuracy value is lower than that obtained with the training data set. All this can be observed graphically through Fig. 8, where it is observed that during training the model adjusts the parameters to try to increase the accuracy. It can be seen that the accuracy of the training and the validation have the same trend but lower values are obtained in the validation. This may suggest that the model is slightly overfitting the training data, but in this case, it does not lead to very severe overfitting that affects the performance of the model. In the same way, the model manages to reduce the loss throughout the training periods. To finish contrasting the results obtained from the model, the confusion matrix is presented in Table 4. With this, the classification carried out by the model for each class of the data set is observed in detail. A main diagonal with high values has been obtained, which denotes that the model is making accurate predictions for each class [38].



(a) Model accuracy



(b) Model loss

Figure 8. Analysis of model performance, experiment without data augmentation.

Table 3. Summary of results obtained in CNN architecture and complete mammography.

Experiment	Accuracy	Precision	Recall	F1 score
No data augmentation	0.747	0.761	0.777	0.792
With extensive data augmentation	0.716	0.730	0.745	0.759
With data augmentation with Rotate, Zoom, and Flip	0.688	0.701	0.716	0.730
With data augmentation with low Rotation, Zoom, and Flip	0.716	0.730	0.745	0.759

Table 4. Confusion matrix of the experiment model without data augmentation.

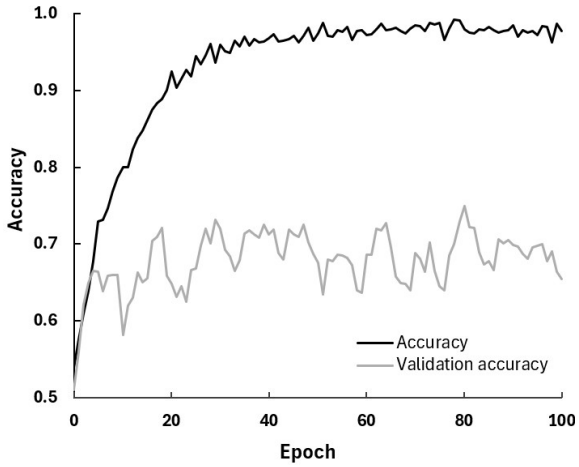
	Real Class: Malignant	Real Class: Benign
Real Class: Malignant	0.74 (TP)	0.24 (FN)
Real Class: Benign	0.36 (FP)	0.64 (TN)

4.1.2 Experiment With Data Augmentation of Rotation, Scrolls, Zoom, Skew and Flip

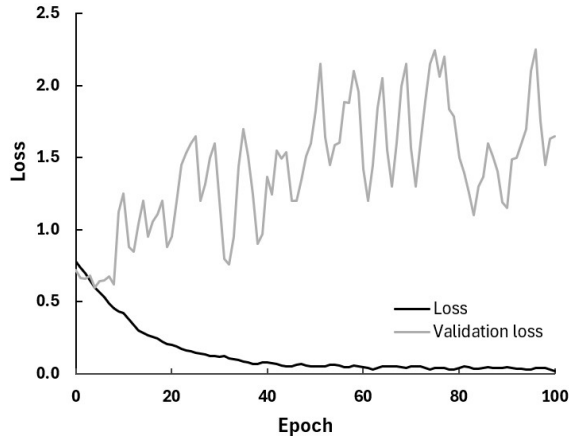
In this second experiment, data augmentation has been carried out in order to try to improve the results obtained. In order to analyse the behavior of the model throughout training, the accuracy and loss graphs are presented in Fig. 9. It shows how the accuracy of the train reaches 100%. However, the actual operation is verified with the validation test, which reaches approximately 70%. This behavior indicates an overfitting of the model to the training data. Furthermore, as can be seen in the confusion matrix presented in Table 5, the results obtained worsen significantly compared to the previous experiment. It can be analysed how the prediction of the Malignant class remains with a high value very close to 1, which indicates that the model is capable of correctly detecting instances of the Malignant class [39]. It is therefore concluded that the model has a tendency to generalize the Malignant class over the other, classifying the majority of tumors as belonging to said class. This behavior of the model is wrong.

Table 5. Confusion matrix of the model with extensive data augmentation.

	Real Class: Malignant	Real Class: Benign
Real Class: Malignant	0.99 (TP)	0.03 (FN)
Real Class: Benign	0.96 (FP)	0.06 (TN)



(a) Model accuracy

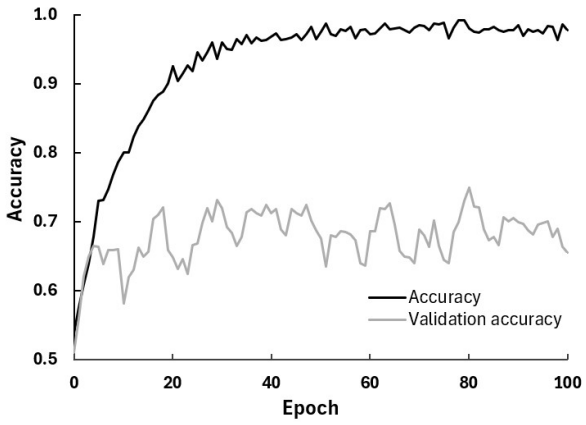


(b) Model loss

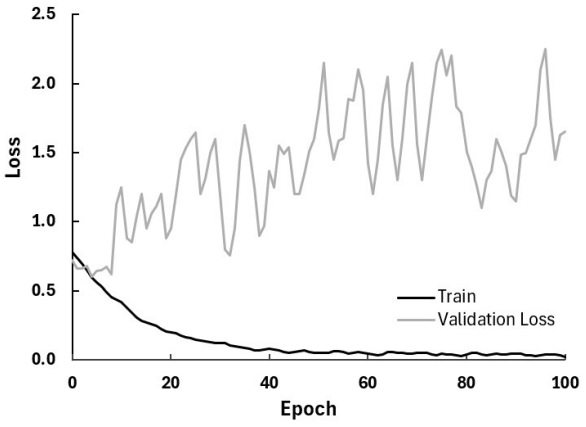
Figure 9. Model results with extensive data augmentation.

4.1.3 Experiment With Data Augmentation Rotation, Zoom and Flip

Below are the results obtained in the third experiment carried out where the network has been trained with a second type of data augmentation. As seen in the confusion matrix in Table 6, the number of instances of the Benign class misclassified decreases. Despite this improvement in the model, an increase in incorrect classification of Benign class tumors is present. In addition to the confusion matrix, the evolution of the accuracy and loss metrics throughout training is also represented in Fig. 10. Both graphs show noisy behavior in the validation metrics, which confirms what was stated by the confusion matrix that the model obtained has difficulty learning correctly [40]. Furthermore, it is observed how the validation loss increases instead of decreasing, just as the training loss does. This behavior of the model denotes an overfitting of the network.

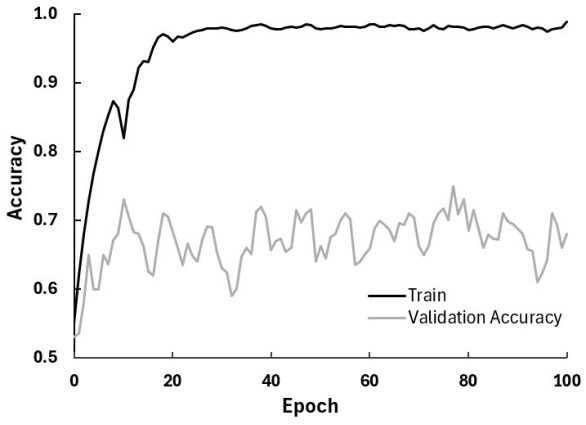


(a) Model accuracy

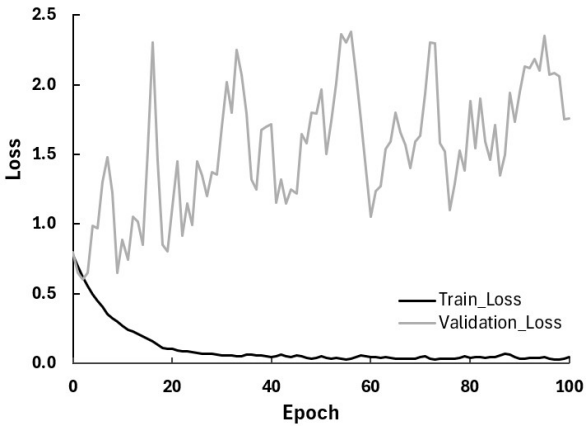


(b) Model loss

Figure 10**.** Model results with data augmentation with Rotate, Zoom and Flip.



(a) Model accuracy



(b) Model loss

Figure 11**.** Model results with data augmentation with low Rotation, Zoom and Flip.

4.1.4 Experiment With Data Augmentation with Low Rotation, Zoom and Flip

The results of the network with the data augmentation are presented below in Fig. 11, and in the confusion matrix as Table 7. It is observed that the correct and incorrect classification of the Malignant class are level, which is a slight improvement compared to the previous experiments where the data augmentation technique was applied. However, there is still ample room to obtain more precise classification metrics [41].

Table 6**.** Confusion matrix of the experiment model with data augmentation with Rotation, Zoom and Flip.

	Real Class: Malignant	**Real Class: Benign**
Real Class: Malignant	0.47 (TP)	0.71 (FN)
Real Class: Benign	0.14 (FP)	0.90 (TN)

Table 7**.** Confusion matrix of the experiment model with data augmentation with low Rotation, Zoom and Flip.

	Real Class: Malignant	**Real Class: Benign**
Real Class: Malignant	00.57 (TP)	0.48 (FN)
Real Class: Benign	0.26 (FP)	0.76 (TN)

Table 8**.** ResNet V2 architecture confusion matrix.

	Real Class: Malignant	**Real Class: Benign**
Real Class: Malignant	0.79 (TP)	0.17 (FN)
Real Class: Benign	0.19 (FP)	0.79 (TN)

4.2 CNN network architecture with mammography images cropped from the ROI

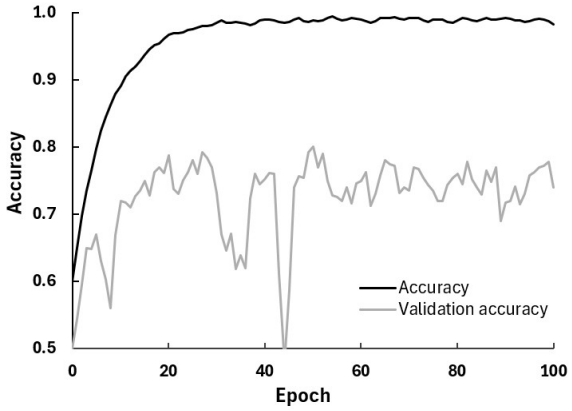
In this case, two different experiments have been carried out to evaluate the performance of the model under various conditions. In the first, the ResNet architecture was used together with the mammogram images cropped from the bounding box of the ROI images. In the second experiment, the DenseNet architecture trained with the same type of data as in the previous experiment has been implemented. In both cases, training of 100 epochs has been carried out.

4.2.1 Experiment applying the ResNet architecture

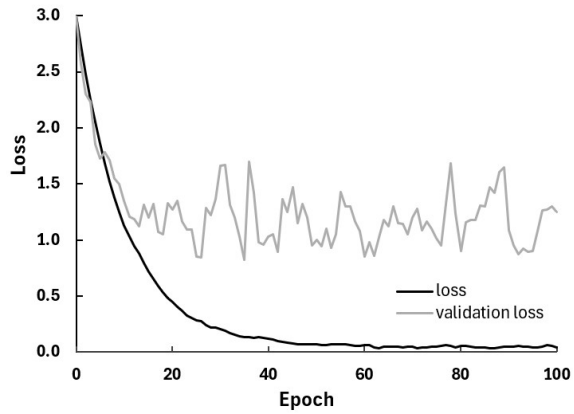
The results of the metrics obtained in both the training and validation sets are observed how the training and validation accuracy values reach close values, which indicates good behavior of the model since there is no evidence that it is overfitting the data. However, further checks on the performance of the model must be carried out before we can conclude how well it is working. To do this, another way to visualize the behavior of the model throughout training is to obtain graphs of the metrics during each epoch of training. Figure 12 shows the detailed behavior of accuracy and loss throughout the training. It can be seen in the graph of the precision model how the metrics remain distant to a certain extent, seeing peaks in the validation curve throughout the training. On the other hand, in the loss model, we see how both curves decrease throughout training but do not remain with close values, thus denoting a slight overfitting of the model [42]. To conclude this analysis, the confusion matrix presented in Table 8 has been obtained. It shows the presence of a main diagonal that takes high values, close to 1. This indicates a high capacity of the model to correctly classify both classes belonging to the data with which we are working. Therefore, even if a small overfitting was obtained during training, the precision values obtained by the model are considered adequate for this work.

4.2.2 Experiment with applying the DenseNet architecture

To carry out the analysis of this model, the same graphs and parameters have been obtained as with the ResNetV2 model.

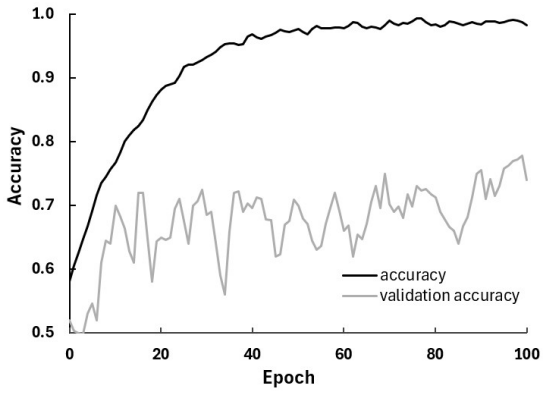


(a) Model accuracy

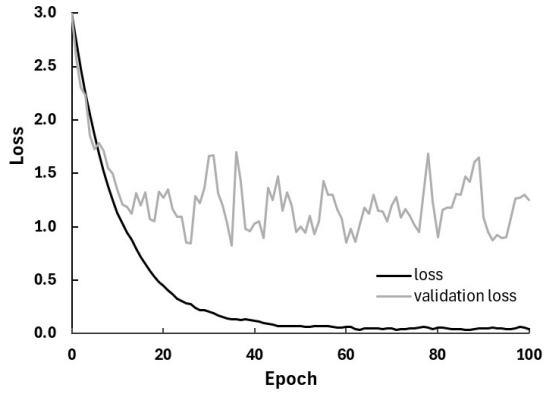


(b) Model loss

Figure 12**.** Performance analysis of ResNetV2 model with cropped mammography.



(a) Model accuracy



(b) Model loss

Figure 13**.** DenseNet model performance analysis.

The final values achieved by the metrics of both the training set and the validation set are observed. High precision values are detected, although these differ by approximately 0.2, which indicates that the perfect behavior of the model is not being obtained. In order to observe if there is overfitting, the graphs of the precision and loss model are displayed to exhaustively analyze the behavior of the network over the epochs. These graphs are presented in Fig. 13. In the precision model you can see how the training curve. In the same way, comparing it with the training curve, which obtains adequate behavior, it is determined that there is overfitting in the model [43]. The same happens with the loss model, a noisy behavior is obtained in the validation curve while the training curve follows an acceptable trend for the model. In order to complete the analysis of the model's behavior, the confusion matrix is presented in Table 9. It confirms the operation of the network, obtaining a low value for the correct predictions of the Benign class. In this way, it is concluded that the model obtained has room for improvement for the tumor classification task [44].

Table 9**.** Confusion matrix of the DenseNet architecture.

	Real Class: Malignant	**Real Class: Benign**
Real Class: Malignant	0.77 (TP)	0.25 (FN)
Real Class: Benign	0.34 (FP)	0.69 (TN)

Table 10**.** Summary of results obtained in CNN architecture and cropped mammography with ROI.

Experiment	**Accuracy**	**Precision**	**Recall**	**F1 score**
Cropped mammography + ResNetV2	0.852	0.823	0.82	0.872
Cropped mammography + DenseNetV2	0.757	0.796	0.783	0.797

4.3 YOLO architecture with full mammography images

YOLO uses the recall [45], precision and mAP metrics to evaluate the performance of the models obtained. The results achieved in the experiments are presented below.

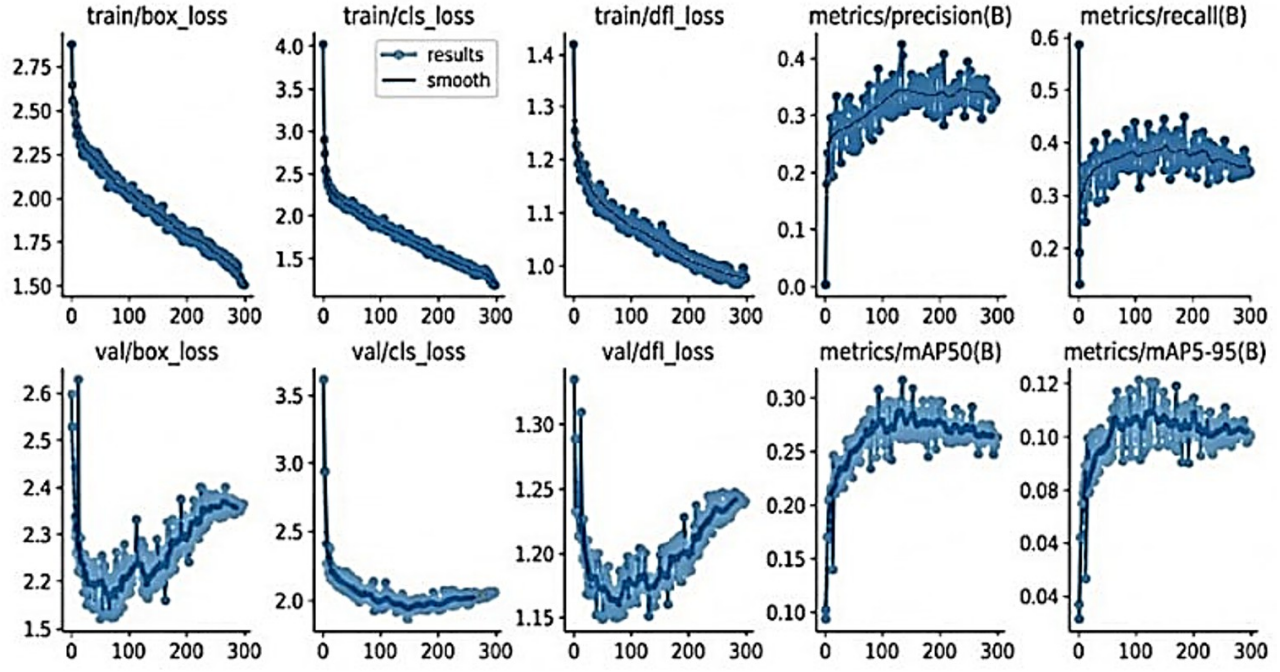
4.3.1 Experiment without AdamW optimizer and horizontal flip

It can be seen in Fig. 14a how the graphs of the model training in the classification and location of the box are not suitable [46]. They denote the presence of overtraining starting around age 70. On the other hand, it is observed that the precision does not exceed 40%, which is a low value for the detection of tumors. The recall is capable of minimally exceeding this 40% but shows that the model still does not correctly classify a high number of images [47]. The mAP graphs show that the model is not robust when it comes to detecting tumors in mammograms.

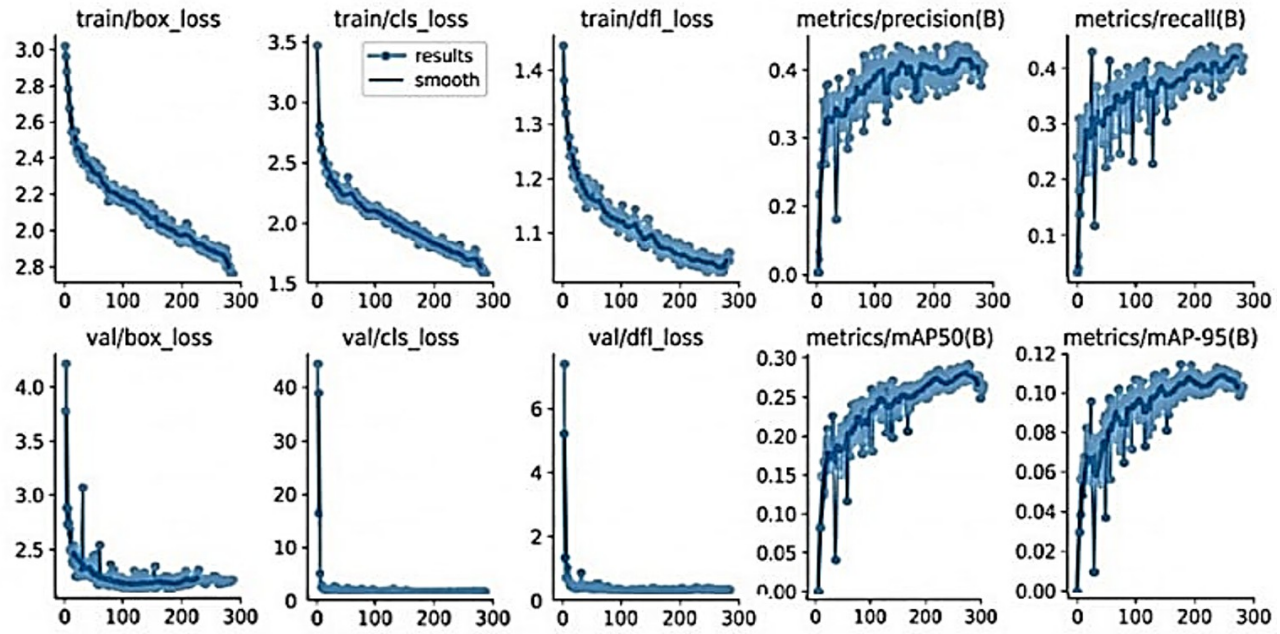
4.3.2 Experiment with AdamW optimizer and horizontal flip

The strategy of adding optimizer and a transformation has not provided better results than those already obtained. As can be seen in Fig. 14b, the model graphs improve slightly, reducing the overtraining present in the previous model. However, it is seen how precision decreases as well as recall and mAP worsen [48].

On the other hand, examples of images classified by the obtained model are presented in Fig. 15. It is observed that of 16 images the model is able to detect tumor in 9 of them. Of those 9 images, it correctly classifies 8. With this, the results are considered to have a range of improvement but are valued as suitable results. It should also be noted that the model in some classified samples is capable of detecting more than one tumor in each image. After a series of checks, it is concluded that it is because it is detecting calcification tumors present in the mammogram. This is considered a strong point of the model that provides robustness to it. On the other hand, Table 10 presents the summary of results obtained in the training developed with the images cropped from the ROI. It is observed that the ResNet architecture has been able to increase its precision compared to the best result obtained with complete mammograms. Therefore, it is evident that providing images focused on the tumor regions improves the results. At the same time, in the experiment carried out applying



(a) Without



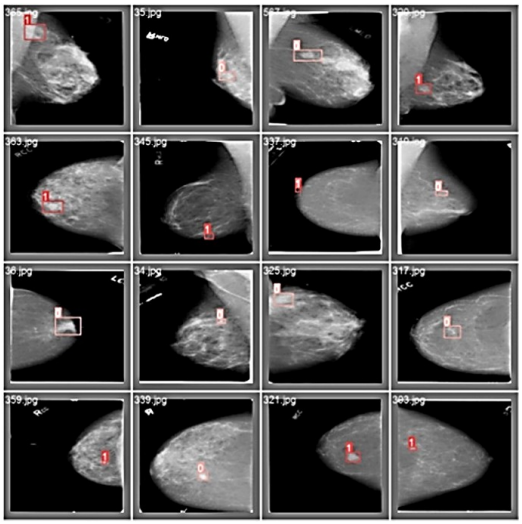
(b) AdamW Optimizer and horizontal flip

Figure 14**.** Yolo model results (a) without and (b) AdamW Optimizer and horizontal flip.

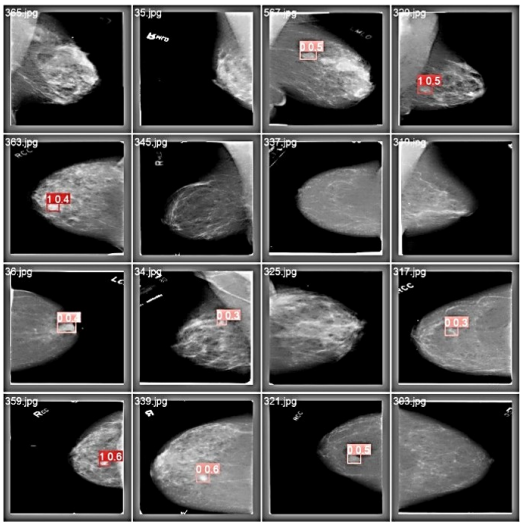
DenseNet, a decrease in precision is observed with respect to what was obtained with ResNet. Therefore, in this case, the ResNet model is considered more appropriate. Finally, Table 11 presents the results of the YOLO network [49] in which notable metrics have not been achieved. However, the final results presented for the classified images are acceptable within the difficulty of the task. During the training process, YOLO has been refined to precisely outline the borders of anomalies. The work also explored detection utilizing YOLO, a cutting-edge object detection technique. YOLO utilizes a singular neural network to concurrently predict several bounding boxes and their associated class probabilities within an image, yielding results that align well with prior studies conducted by other researchers [50–52]. By using Grad-CAM to the best-performing CNN architectures, researchers can create heatmaps that emphasize active regions in breast images, improving the interpretability and reliability of AI-driven breast cancer detection models. The Grad-CAM plot visualizes the active regions on the breast images, highlighting the areas that the CNN model focuses on for tumor detection. The heatmap overlayed on the mammography depicts regions of interest, with warm colors indicating strong model attention and likely tumor locations as shown Fig. 16. This offers information about the model's focus during the detection process.

Table 11**.** Summary of results obtained in YOLO architecture and complete mammography.

Experiment	Accuracy	Precision	Recall	F1 score
YOLO without AdamW Optimizer and horizontal flip	0.343	0.326	0.394	0.372
YOLO with AdamW Optimizer and horizontal flip	0.324	0.312	0.371	0.329



(a) Without AdamW



(b) horizontal flip

Figure 15**.** Results obtained in the YOLO architecture with AdamW optimizer and horizontal flip.

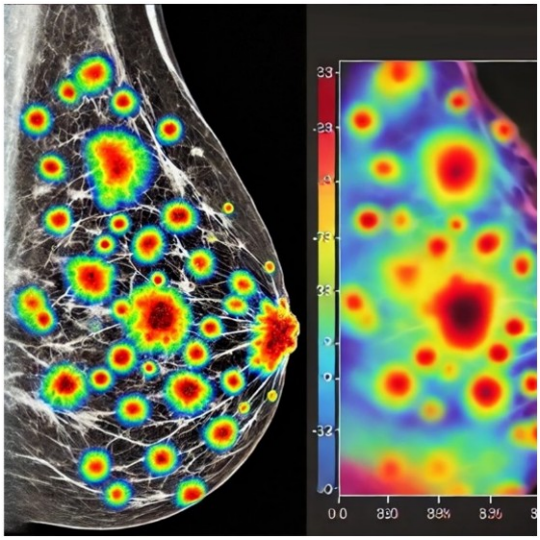


Figure 16**.** Gradient-weighted Class Activation Mapping (GRAM-CAM) image.

5. Conclusions

In the development of this work, different DL methodologies [53–55] have been implemented for the detection and classification of tumors in mammograms. To achieve this, various experiments have been carried out applying different neural network architectures. To achieve the main objective, work on the current applications of DL in tumor detection has been carried out. In this way, those neural network architectures that a priori provide the best results have been chosen. One of the great challenges of this work has been obtaining and pre-processing the data set. The scarcity of public medical images has been a great handicap when collecting data. In turn, this process has included the use of data augmentation techniques in order to improve the robustness of the models. This entire process has taken a remarkable amount of time. However, it has been essential to guarantee the diversity of the data used. After the preprocessing phase, the training of various neural network models has been carried out. The strategy of using multiple architectures has allowed us to compare the performance of each one and determine the most appropriate model for this task. After having exhaustively evaluated the results obtained in the different experiments, it is concluded that there is room for improvement. The existence of overfitting has been present in several of the models obtained. The precision of the results in the detection of tumors should not give rise to error since this would lead to errors in the diagnoses. In order to obtain an improvement in the results achieved, the following lines of future work are proposed.

- Increasing the dataset & the use of other types of medical images other than mammograms is also being considered.
- The training of the models is proposed with the data provided by the dataset of tumors from calcifications.
- It is also proposed to be able to apply different architectures to those already exposed in this work.
- Transfer Learning is proposed to apply transfer learning techniques, thus increasing the performance provided by the pre-trained models.

Authors' contribution

Qutaiba Humadi Mohammed: Proposed the problem statement, developed the theory and performed the computations. Noor Alhuda Kh Ibrahim: Data collecting and determine the objective. Anupama Namburu & Chaitanya Konda: Verified the analytical methods and investigated and supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

Declaration of competing interest

The authors declare no conflicts of interest.

Funding source

This study didn't receive any specific funds.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] M. Arnolda, E. Morgana, H. Rumgay, and et al., "Current and future burden of breast cancer: Global statistics for 2020 and 2040," *The Breast*, vol. 66, pp. 15–23, 2022. [Online]. Available: https://doi.org/10.1016/j.breast.2022.08.010

[2] H. M. Ahmed and K. Khoshnood, "Breast cancer stigma in iraq," *Current Breast Cancer Reports*, vol. 15, no. 4, p. 345–349, 2023. [Online]. Available: https://doi.org/10.1007/s12609-023-00513-2

[3] N. N. R. Alrawi, "A review on breast cancer in iraq and future therapies insights," *Chemical Engineering Journal*, vol. 3, no. 01, p. 4–16, 2022. [Online]. Available: https://doi.org/10.47419/bjbabs.v3i01.64

[4] P. Rawla, T. Sunkara, and A. Barsouk, "Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors," *Prz Gastroentero*, vol. 14, no. 2, pp. 89–103, 2019. [Online]. Available: :https://doi.org/10.5114/pg.2018.81072

[5] H. Swaminathan, K. Saravanamurali, and A. Y. Sangilimuthu, "Extensive review on breast cancer its etiology, progression, prognostic markers, and treatment," *Medical Oncology*, vol. 40, no. 8, p. 238, 2023. [Online]. Available: https://doi.org/10.1007/s12032-023-02111-9

[6] M. Cardoso, P. Poortmans, E. Senkus, D. Gentilini, and N. Houssami, "Breast cancer highlights from 2023: Knowledge to guide practice and future research," *Breast*, vol. 74, p. 103674, 2024. [Online]. Available: https://doi.org/10.1016/j.breast.2024.103674

[7] X. Mao and et al., "Breast cancer incidence after a false-positive mammography result," *JAMA Oncology*, vol. 10, no. 1, pp. 63–70, 2024. [Online]. Available: https://doi.org/10.1001/jamaoncol.2023.4519

[8] N. Moshina, A. Gräwingholt, K. Lång, and et al., "Digital breast tomosynthesis in mammographic screening: false negative cancer cases in the to-be 1 trial," *Insights Imaging*, vol. 15, p. 38, 2024. [Online]. Available: https://doi.org/10.1186/s13244-023-01604-5

[9] Z. Tasnim, S. Chakraborty, F. M. J. M. Shamrat, A. Newaz, and et al, "Deep learning predictive model for colon cancer patient using cnn-based classification," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 8, 2021. [Online]. Available: https://doi.org/10.14569/ijacsa.2021.0120880

[10] C. Kavyashree, H. Vimala, and J. Shreyas, "Improving oral cancer detection using pretrained model," *2022 IEEE 6th Conference on Information and Communication Technology (CICT)*, pp. 1–5, 2022. [Online]. Available: https://doi.org/10.1109/cict56698.2022.9997897

[11] A. Sinha, M. Pandey, M. Nazma, J. Naskar, and S. S. Rautaray, "Segmentation-based classification deep learning model for breast cancer detection using mammogram images," *2023 IEEE 3rd Mysore Sub Section International Conference (MysuruCon), HASSAN, India*, pp. 1–8, 2023. [Online]. Available: https://doi.org/10.1109/mysurucon59703.2023.10397015

[12] U. Sirisha, S. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, "Statistical analysis of design aspects of various yolo-based deep learning models for object detection," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, p. 126, 2023. [Online]. Available: https://doi.org/10.1007/s44196-023-00302-w

[13] M. Panda, A. Abraham, B. Gopi, and R. Ajith, "Computational intelligence for oncology and neurological disorders," *Informa,CRC Press*, vol. 35, no. 1, p. 292, 2024. [Online]. Available: https://doi.org/10.1201/9781003450153

[14] A. Carriero, L. Groenhoff, E. Vologina, P. Basile, and M. Albera, "Deep learning in breast cancer imaging: State of the art and recent advancements in early 2024," *Diagnostics*, vol. 14, no. 8, p. 848, 2024. [Online]. Available: https://doi.org/10.3390/diagnostics14080848

[15] C. Rajeshkumar, K. R. Soundar, M. Sneha, S. S. Maheswari, M. S. Lakshmi, and R. Priyanka, "Convolutional neural networks (cnn) based brain tumor detection in mri images," *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 976–979, 2023. [Online]. Available: https://doi.org/10.1109/icssit55814.2023.10060968

[16] Y. Kadhim, U. Khan, and A. Mishra, "Deep learning-based computer-aided diagnosis (cad): Applications for medical image datasets," *Sensors*, vol. 22, no. 22, p. 8999, 2022. [Online]. Available: https://doi.org/10.3390/s22228999

[17] R. Yamashita, M. Nishio, K. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, p. 611–629, 2018. [Online]. Available: https://doi.org/10.1007/s13244-018-0639-9

[18] L. Alzubaidi and et al., "Review of deep learning: concepts, cnn architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, p. 53, 2021. [Online]. Available: https://doi.org/10.1186/s40537-021-00444-8

[19] O. Olayemi, O. Salako, A. Jinadu, A. M. Obalalu, , and B. E. Anyaegbuna, "Aerodynamic lift coefficient prediction of supercritical airfoils at transonic flow regime using convolutional neural networks (cnns) and multi-layer perceptions (MLPs)," *Al-Qadisiyah Journal for Engineering Sciences*, vol. 16, no. 2, p. 108–115, 2023. [Online]. Available: https://doi.org/10.30772/qjes.v16i2.955

[20] T. Zhou, X. Ye, H. Lu, X. Zheng, S. Qiu, and Y. Liu, "Dense convolutional network and its application in medical image analysis," *BioMed Research International*, vol. 2022, no. 1, p. 2384830, 202. [Online]. Available: https://doi.org/10.1155/2022/2384830

[21] M. Shafiq and Z. Gu, "Deep residual learning for image recognition: A survey," *Applied Sciences*, vol. 12, no. 18, p. 8972, 2022. [Online]. Available: https://doi.org/10.3390/app12188972

[22] L. He, P. Tiwari, C. Lv, W. Wu, and L. Guo, "Reducing noisy annotations for depression estimation from facial images," *Neural Networks*, vol. 153, p. 120–129, 2022. [Online]. Available: https://doi.org/10.1016/j.neunet.2022.05.025

[23] M. Anas, I. U. Haq, G. Husnain, and A. Faraz, "Advancing breast cancer detection: Enhancing yolov5 network for accurate classification in mammogram images," *IEEE Access*, vol. 12, pp. 16 474–16 488, 2024. [Online]. Available: https://doi.org/10.1109/access.2024.3358686

[24] S. A. Hasanah, A. A. Pravitasari, A. S. Abdullah, I. N. Yulita, and M. H. Asnawi, "A deep learning review of resnet architecture for lung disease identification in cxr image," *Applied sciences*, vol. 13, no. 24, p. 13111, 2023. [Online]. Available: https://doi.org/10.3390/app132413111

[25] T. Ochiai and A. Fujishima, "Comparison of yolo versions for object detection from aerial images," *International Journal of Environment and Geoinformatics*, vol. 9, no. 2, p. 87–93, 2022. [Online]. Available: https://doi.org/10.30897/ijegeo.1010741

[26] R. Lee, F. Gimenez, A. Hoogi, K. Miyake, M. Gorovoy, and D. Rubin, "A curated mammography data set for use in computer-aided detection and diagnosis research," *Scientific Data*, vol. 4, no. 1, p. 170177, 2017. [Online]. Available: https://doi.org/10.1038/sdata.2017.177

[27] M. Brahim, K. Westerkamp, L. Hempel, R. Lehmann, D. Hempel, and P. Philipp, "Automated assessment of breast positioning quality in screening mammography," *Cancers*, vol. 14, no. 19, p. 4704, 2022. [Online]. Available: https://doi.org/10.3390/cancers14194704

[28] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal, and J. Savage, "Mammographic image analysis society (MIAS) database v1.21," *Apollo - University of Cambridge Repository*, 2015. [Online]. Available: https://doi.org/10.17863/CAM.105113

[29] P. Park, H. Oh, and H. Kim, "Dual-color space network with global priors for photo retouching," *Scientific Reports*, vol. 13, no. 1, p. 19798, 2023. [Online]. Available: https://doi.org/10.1038/s41598-023-47186-6

[30] K. Saa, "Method for detection and diagnosis of the area of skin disease based on color by wavelet transform and artificial neural network," *Al-Qadisiyah Journal for Engineering Sciences*, vol. 2, no. 4, p. 799–829, 2022.

[31] A. Bugeau, R. Giraud, and L. Raad, "Influence of color spaces for deep learning image colorization," *Springer eBooks*, p. 847–878, 2023. [Online]. Available: https://doi.org/10.1007/978-3-030-98661-2_125

[32] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, p. 100258, 2022. [Online]. Available: https://doi.org/10.1016/j.array.2022.100258

[33] S. Ni, F. Shao, X. Chai, H. Chen, and Y.-S. Ho, "Composition-guided neural network for image cropping aesthetic assessment," *IEEE Transactions on Multimedia*, vol. 25, no. 3, p. 6836–6851, 2022. [Online]. Available: https://doi.org/10.1109/tmm.2022.3215003

[34] A. H. Abdulaal, M. Valizadeh, C. Amirani, and M. S. Shah, "A self-learning deep neural network for classification of breast histopathological images," *Biomedical Signal Processing and Control*, vol. 87, no. part B, p. 105418, 2024. [Online]. Available: https://doi.org/10.1016/j.bspc.2023.105418

[35] M. A. Wakili, H. A. Shehu, M. H. Sharif, and et al., "Classification of breast cancer histopathological images using densenet and transfer learning," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 8904768, 2022. [Online]. Available: https://doi.org/10.1155/2022/8904768

[36] X. Jiang, Z. Hu, S. Wang, and Y. Zhang, "Deep learning for medical image-based cancer diagnosis," *Cancers*, vol. 15, no. 14, p. 3608, 2023. [Online]. Available: https://doi.org/10.3390/cancers15143608

[37] M. Nasser and U. K. Yusof, "Deep learning based methods for breast cancer diagnosis: A systematic review and future direction," *Diagnostics*, vol. 13, no. 1, p. 161, 2023. [Online]. Available: https://doi.org/10.3390/diagnostics13010161

[38] K. Uddin, N. Biswas, T. Rikta, and K. Dey, "Machine learning-based diagnosis of breast cancer utilizing feature optimization technique," *Computer Methods and Programs in Biomedicine Update*, vol. 3, p. 100098, 2023. [Online]. Available: https://doi.org/10.1016/j.cmpbup.2023.100098

[39] J. Guo and et al., "A novel breast cancer image classification model based on multiscale texture feature analysis and dynamic learning," *Scientific Reports*, vol. 14, no. 1, p. 7216, 2024. [Online]. Available: https://doi.org/10.1038/s41598-024-57891-5

[40] M. D. Ali, A. Saleem, H. Elahi, M. A. Khan, and et al., "Breast cancer classification through meta-learning ensemble technique using convolution neural networks," *Diagnostics*, vol. 13, no. 13, p. 2242, 2023. [Online]. Available: https://doi.org/10.3390/diagnostics13132242âĂŘ

[41] R. G. Martinez and M. van Dongen, "Deep learning algorithms for the early detection of breast cancer: A comparative study with traditional machine learning," *Informatics in Medicine Unlocked*, vol. 41, p. 101317, 2023. [Online]. Available: https://doi.org/10.1016/j.imu.2023.101317

[42] A. Raza, N. Ullah, J. A. Khan, M. Assam, A. Guzzo, and H. Aljuaid, "Deepbreastcancernet: A novel deep learning model for breast cancer detection using ultrasound images," *Applied sciences*, vol. 13, no. 4, p. 2082, 2023. [Online]. Available: https://doi.org/10.3390/app13042082

[43] A. M. Alhassan, "An improved breast cancer classification with hybrid chaotic sand cat and remora optimization feature selection algorithm," *PLoS ONE*, vol. 19, no. 4, 2024. [Online]. Available: https://doi.org/10.1371/journal.pone.0300622

[44] A. Boulenger, Y. Luo, C. Zhang, and et al, "Deep learning-based system for automatic prediction oftriple-negative breast cancer from ultrasound images," *Medical Biological Engineering Computing*, vol. 61, p. 567–578, 2023. [Online]. Available: https://doi.org/10.1007/s11517-022-02728-4

[45] F. Prinzi, M. Insalaco, A. Orlando, S. Gaglio, and S. Vitabile, "A yolo-based model for breast cancer detection in mammograms," *Cognitive Computation*, vol. 16, no. 1, p. 107–120, 2023. [Online]. Available: https://doi.org/10.1007/s12559-023-10189-6

[46] N. Hassan, S. Hamad, and K. Mahar, "Yolo-based cad framework with vit transformer for breast mass detection and classification in cesm and ffdm images," *Neural computing applications*, vol. 36, p. 6467–6496, 2024. [Online]. Available: https://doi.org/10.1007/s00521-023-09364-5

[47] A. Mohammed and D. Ekmekci, "Breast cancer diagnosis using yolo-based multiscale parallel cnn and flattened threshold swish," *Applied Sciences*, vol. 14, no. 7, p. 2680, 2024. [Online]. Available: https://doi.org/10.3390/app14072680

[48] S. R. Kebede and et al., "Dual view deep learning for enhanced breast cancer screening using mammography," *Scientific Reports*, vol. 14, no. 1, p. 3839, 2024. [Online]. Available: https://doi.org/10.1038/s41598-023-50797-8

[49] G. Aly, M. Marey, S. El-Sayed, and F. Tolba, "Yolo based breast masses detection and classification in full-field digital mammograms," *Computer Methods and Programs in Biomedicine*, vol. 200, p. 105823, 2021. [Online]. Available: https://doi.org/10.1016/j.cmpb.2020.105823

[50] J. Terven, M. Córdova-Esparza, and A. Romero-González, "A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, p. 1680–1716, 2023. [Online]. Available: https://doi.org/10.3390/make5040083

[51] X. Wang, H. Gao, Z. Jia, and Z. Li, "Bl-yolov8: An improved road defect detection model based on yolov8," *Sensors*, vol. 23, no. 20, p. 8361, 2023. [Online]. Available: https://doi.org/10.3390/s23208361

[52] G. Yang, J. Wang, Z. Nie, H. Yang, and S. Yu, "A lightweight yolov8 tomato detection algorithm combining feature enhancement and attention," *Agronomy*, vol. 13, no. 7, p. 1824, 2023. [Online]. Available: https://doi.org/10.3390/agronomy13071824

[53] Q. Hu, M. Whitney, and L. Giger, "A deep learning methodology for improved breast cancer diagnosis using multiparametric mri," *Scientific Reports*, vol. 10, no. 1, p. 10536, 2020. [Online]. Available: https://doi.org/10.1038/s41598-020-67441-4

[54] R. Adam, K. Dell'Aquila, T. M. L. Hodges, and Q. Duong, "Deep learning applications to breast cancer detection by magnetic resonance imaging: a literature review," *Breast Cancer Research*, vol. 25, no. 1, p. 87, 2023. [Online]. Available: https://doi.org/10.1186/s13058-023-01687-4

[55] L. Shen, R. Margolies, H. Rothstein, E. Fluder, R. McBride, and W. Sieh, "Deep learning to improve breast cancer detection on screening mammography," *Scientific Reports*, vol. 9, no. 1, p. 12495, 2019. [Online]. Available: https://doi.org/10.1038/s41598-019-48995-4

How to cite this article:
Qutaiba Humadi Mohammed, Noor Alhuda Kh Ibrahim, Anupama Namburu, and Chaitanya Konda, (2026). 'Breast cancer detection using deep learning techniques: An investigation using the CBIS-DDSM dataset and customized neural network model, ResNetV2 and YOLO', Al-Qadisiyah Journal for Engineering Sciences, 19(1), pp. 059-070. https://doi.org/10.30772/qjes.2025.155626.1429