Journal of Cancer Research and Clinical Oncology (2023) 149:14549–14564
https://doi.org/10.1007/s00432-023-05249-1

RESEARCH

Enhanced breast mass mammography classification approach based on pre-processing and hybridization of transfer learning models

Saida Sarra Boudouh[1] · Mustapha Bouakkaz[1]

Received: 25 June 2023 / Accepted: 4 August 2023 / Published online: 12 August 2023

Abstract

Background and objective The second most prevalent cause of death among women is now breast cancer, surpassing heart disease. Mammography images must accurately identify breast masses to diagnose early breast cancer, which can significantly increase the patient's survival percentage. Although, due to the diversity of breast masses and the complexity of their microenvironment, it is still a significant issue. Hence, an issue that researchers need to continue searching into is how to establish a reliable breast mass detection approach in an effective factor application to increase patient survival. Even though several machine and deep learning-based approaches were proposed to address these issues, pre-processing strategies and network architectures were insufficient for breast mass detection in mammogram scans, which directly influences the accuracy of the proposed models.

Methods Aiming to resolve these issues, we propose a two-stage classification method for breast mass mammography scans. First, we introduce a pre-processing stage divided into three sub-strategies, which include several filters for Region Of Interest (ROI) extraction, noise removal, and image enhancements. Secondly, we propose a classification stage based on transfer learning techniques for feature extraction, and global pooling for classification instead of standard machine learning algorithms or fully connected layers. However, instead of using the traditional fine-tuning feature extraction phase, we proposed a hybrid model where we concatenate two recent pre-trained CNNs to assist the feature extraction phase, rather than using one.

Results Using the CBIS-DDSM dataset, we managed to increase mainly each of the accuracy, sensitivity, and specificity reaching the highest accuracy of 98,1% using the Median filter for noise removal. Followed by the Gaussian filter trial with 96% accuracy, meanwhile, the winner filter attained the lowest accuracy of 94.13%. Moreover, the usage of global average pooling as a classifier is suitable in our case better than global max pooling.

Conclusion The experimental findings demonstrate that the suggested strategy of breast Mass detection in mammography can outperform the top-ranked methods currently in use in terms of classification performance.

Keywords Breast cancer · Breast mass detection · Mammography processing · Deep learning

Introduction

Breast cancer is becoming the most frequently diagnosed disease among women (DeSantis et al. 2019). Based on the presence or lack of molecular markers for estrogen, progesterone, and human epidermal growth factor, breast cancer is divided into three main categories (Waks and Winer 2019). With an estimated 2.3 million new cases per year, female breast cancer has now surpassed lung cancer as the most common kind of cancer in the world (Sung et al. 2021). The key components of treating breast cancer include early detection and patient monitoring (Jafari et al. 2018).

Imaging techniques, among many other modalities, have developed into useful tools for determining and monitoring treatment outcomes in primary breast cancer (Jafari et al. 2018). While MRI and mammography have proven the most sensitive imaging techniques for high-risk breast cancer patients, additional imaging techniques may be helpful for different risk groups (Wellings et al. 2016). A low-dose

✉ Saida Sarra Boudouh
s.boudouh@lagh-univ.dz

Mustapha Bouakkaz
m.bouakkaz@lagh-univ.dz

[1] LIM Laboratory, University of Laghouat Amar Telidji, Laghouat, Algeria

X-ray called a mammography can show the interior structure of the breast. Radiologists utilize it to spot worrisome breast abnormalities (Waks and Winer 2019).

One of the concerning symptoms that warrants further research is the emergence of Micro-Calcifications (MCs), tiny calcium deposits that often manifest in the breast (Loizidou et al. 2020). While masses remain the most prevalent abnormality detected by mammography (Divyashree and Kumar 2021), they refer to breast tissues exhibiting uncontrolled growth and displaying an unnatural shape and contour. Micro-Calcifications (MCs) are an intriguing area of interest for researchers due to their potential significance in breast health assessment. These minuscule calcified deposits, when detected during mammographic examinations, can raise red flags and prompt the need for closer investigation. The accurate characterization and understanding of these micro-calcifications are crucial in distinguishing benign formations from potentially malignant ones, which can play a decisive role in early cancer detection and treatment planning (society 2019). On the other hand, mammograms often reveal the presence of masses, which can be a source of concern for both patients and healthcare providers. These masses represent abnormal growth within the breast tissue, and their identification through mammography is paramount in the detection and diagnosis of breast cancer (Loizidou et al. 2020). Detecting masses at an early stage enables timely intervention, which significantly improves the chances of successful treatment and positive patient outcomes. While mammography is a valuable screening tool, it is essential to recognize that the appearance of MCs and masses may not always signify cancer. In some instances, benign conditions can also give rise to similar imaging findings. Hence, accurate interpretation by skilled radiologists and additional diagnostic procedures may be required to ascertain the nature of these abnormalities and devise appropriate management strategies. Therefore, advancing research in the fields of micro-calcifications and mass characterization will contribute significantly to refining breast cancer screening, diagnosis, and treatment approaches. Enhanced awareness and improved imaging technologies can lead to earlier detection, improved accuracy, and ultimately better prognoses for individuals facing breast health concerns (Mahmood et al. 2022).

The detection of breast masses can be challenging due to the wide variation in their intensity, distribution, shape (lobulated, irregular, round, oval), and boundary characteristics (spiculated, ill-defined, confined) within the breast area (Mahmood et al. 2022). As breast cancer masses are frequently small and may not exhibit any initial symptoms, the timely identification of abnormalities becomes paramount for ensuring effective treatment (Mahmood et al. 2022). Nevertheless, early diagnosis faces obstacles in the form of potential misinterpretation of mammograms, influenced by both the radiologists' expertise and the quality of the imaging, which can result in diagnostic errors. The increase in failures to detect abnormalities in mammograms can be attributed to several significant factors. Firstly, poor image quality compromises the clarity of the findings, making it more arduous for radiologists to discern potential issues accurately. Secondly, the demanding nature of the job can lead to eye fatigue, reducing the radiologists' vigilance and increasing the risk of overlooking critical details. Lastly, occasional negligence on the part of radiologists further adds to the complexity of the diagnostic process, potentially leading to missed detection that could have serious consequences for the patient's health. Addressing these challenges and developing better tools and protocols for mammogram analysis is of utmost importance to enhance early detection rates and ultimately improve breast cancer outcomes.

Deep Convolutional Neural Networks (DCNNs), which show exceptional results in image segmentation, classification, and retrieval, are one of the greatest learning techniques for comprehending visual equipment (Khan et al. 2020). However, the scarcity of available datasets, particularly in the healthcare industry, and the substantial quantity of data required for CNN training reduce the models' accuracy. Sinno Jialin Pan and Qiang Yang suggested the transfer learning strategy as a solution to the issue at the beginning of 2009 (Zhou et al. 2020). Due to the dearth and small quantity of available datasets in the medical industry generally, and mammography databases in particular, deep transfer learning is the best approach to address such problems. Therefore, applying deep transfer learning techniques to the healthcare sector can aid in the earlier detection and diagnosis of breast cancer by medical professionals. Despite numerous attempts to tackle these challenges by creating several Computer Aided Diagnosis (CAD) systems, using various machines and deep learning-based approaches, the existing pre-processing strategies and network architectures have proven inadequate for effective breast mass detection in mammogram scans. This deficiency directly impacts the accuracy of the proposed models. Advancements in technology and the integration of deep learning techniques have shown promising results in the field of medical imaging, particularly in mammography. However, the complexities involved in identifying breast masses demand robust and sophisticated methodologies. Pre-processing, which involves enhancing image quality, noise reduction, and normalization, is a critical step in preparing mammogram scans for accurate analysis. Yet, achieving optimal pre-processing methods that cater to the intricacies of breast tissue representation remains a challenge. Moreover, the design of the network architecture is of paramount importance in deep learning-based approaches. The network must be capable of extracting relevant features and patterns from mammogram images to distinguish between benign breast tissue and potential malignant masses accurately. Unfortunately, finding an optimal architecture that strikes the right balance between complexity and generalizability is a formidable task, and sub-optimal architectures can result in reduced accuracy. To achieve substantial progress in breast mass detection, researchers need to invest further effort in refining pre-processing techniques and developing innovative network architectures. Collaborative endeavors between medical experts and computer scientists can foster the creation of tailored approaches that cater specifically to the unique characteristics of mammogram images. By addressing these limitations head-on, we can enhance the accuracy and reliability of breast mass detection models, ultimately leading to improved early detection and patient outcomes.

Hence, aiming to address these issues, we focused to propose a reliable pre-processing strategy that is divided into three sub-sections, to distinguish the most suitable and robust sequence of filters for breast Masses classification in mammogram images. Furthermore, we aimed to validate the proposed pre-processing strategy through a classification phase based on deep learning approaches. In this paper, we propose a novel two-stage method to improve breast mass detection in mammography. We aimed to explore the gap in predicting breast masses. Following is a summary of overall contributions:

- To identify the most efficient sequence of filters for each type of breast mass, we propose a pre-processing strategy with three sub-strategies.
- We suggest a new novel CNN architecture for detecting breast masses. Where the feature extraction phase included a concatenation of two recent pre-trained CNNs in order to extract the maximum amount of features. Meanwhile, in the classification phase instead of using a traditional ML algorithm or fully connected layer, we used global pooling techniques.
- By reducing false positives and false negatives, we provide a method that will improve sensitivity and specificity. Hence, improve breast Mass detection's current accuracy.

The remainder of the suggested work is organized as follows: The second section provides an overview of the state of the art. Section "Proposed Methodology" describes the proposed approach in more detail. The section that follows deals with the results and discussion. Finally, Section " Declarations" includes the conclusion.

Related work

In the field of breast mass detection, there have been several works published. Yet, each group of them focused on a specific topic in this large domain. We can divide this work into three main strategies (Fig. 1). The first one mainly focuses on the pre-processing phase. Thus, it includes the usage of several filters on mammogram images, for image segmentation, ROI extraction, and so on. The second part concentrates on the feature extraction and classification or object detection phase. Thus, DCNNs and deep transfer learning techniques are applied. Lastly, the third part is a merge between the first two parts, where the work focuses on image pre-processing as the first step into the feature extraction and classification phase.

Pre-processing-based approaches

Starting with the strategy, where the authors focused on pre-processing the mammogram images in order to segment the

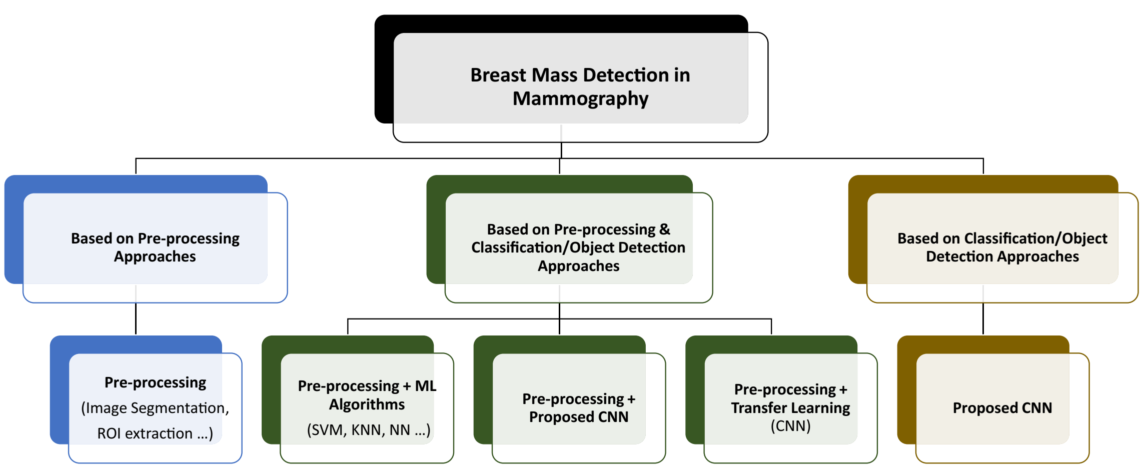


Fig. 1 Breast mass detection approaches

images, remove noise, and extract the ROI, which is the mass area in this case.

Divyashree and Kumar (2021), proposed a segmentation algorithm for background suppression, and pectoral muscle removal is performed using a gradient weight map followed by gray difference weight and fast marching method. Aiming for enhancing breast mass region, they used contrast limited adaptive histogram equalization (CLAHE) filter and de-correlation stretch. Mass detection achieved their highest accuracy of 97.64% and 94.66% for the MIAS and CBIS-DDSM datasets, respectively.

Famouri et al. (2021), proposed a work that studies the impact of noise on the training process of breast mass detection, and how to improve such a process. Using the CBIS-DDSM dataset, the authors demonstrate how noise propagates during training and decreases the R-CNN network's capacity to accurately discriminate lesions from the background due to an unsatisfactory match between the ground truth and the proposed bounding boxes for the network. The proposed methodology decreased the AUC by 9%.

Pre-processing and classification-based approaches

Secondly, we have a strategy that focuses on feature extraction and classification, Which can be divided into three categories:

- Image pre-processing followed by ML algorithm application.
- Image pre-processing followed by a proposed CNN architecture.
- Finally, image pre-processing followed by transfer learning techniques application, thus modifying a pre-trained CNN.

Loizidou et al. (2022), proposed an algorithm for the automatic detection and classification of benign and malignant breast masses. They collected a dataset of 196 images. Instead of proposing a CNN, they applied nine different classifiers: Linear Discriminant Analysis (LDA), k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Naive Bayes (NB), Random Forest (RF), AdaBoost (ADA), Bagging (BAG), Gradient Boosting (GB) and Voting. In addition to that, they tried different Neural Network (NN) architectures. They used cross-validation 7-fold, where the proposed NN reached the highest accuracy of 90.85%, 91.67% sensitivity, and 89.66% specificity. Meanwhile, Lbachir et al. (2021), proposed a CAD system for breast mass detection and diagnosis, that is divided into four main phases. Starting the three first phases could be cited under the name of pre-processing. First, the removed noise and enhance images. In the second step, the abnormalities were segmented using a proposed algorithm. Third, they reduced the false positives by texture and shape features and the bagged trees classifier. Finally, they used SVM to classify images into malignant or benign. The proposed system was tested with CBIS-DDSM and MIAS databases, reaching the highest accuracy of 94,2% and 0,95 AUC. On the other hand, Boumaraf et al. (2020) proposed an advanced computer-aided diagnosis (CAD) system for classifying mammographic masses into four BI-RADS categories. Their system used histogram equalization and region growing for enhancement and segmentation, as pre-processing phase. Extracting 130 BI-RADS features, they employed a genetic algorithm for selecting the most significant ones. Using a back-propagation neural network, their system achieved impressive classification accuracy of 84.5%. Their approach showed promise in supporting radiologists' decisions based on automatically assigned BI-RADS categories, improving breast mass diagnosis. Yet, the usage of transfer learning techniques for feature extraction could have boosted their system. Additionally, the authors applied a region of interest extraction to the DDSM dataset, where it does exist a ready dataset that is a crroped version of the dateset (CBIS-DDSM). Moreover, Sun et al. (2021), proposed a method for breast mass detection. Using the DDSM dataset, they started with the pre-processing phase via a mathematical morphology method and locates the suspected regions of a breast mass by the image template matching method. Next, using a proposed CNN (BD-CNN) they classified the extracted ROI (mass region) into breast mass and background categories. The proposed method reached the highest accuracy, sensitivity, specificity, and precision of 85.82%, 95.38%, 96%, and 50.81% respectively. Furthermore, Chugh et al. (2022), proposed a novel strategy for breast mass detection on mammogram images. Using the MIAS dataset, the pre-processing stage included a mix of morphological and multi-thresholding using Otsu's technique. Then, the inbuilt feature extraction was achieved using multiple color and texture feature approaches. The classification phase was divided into two primary strategies: first a model for dividing breast tissue into dense and non-dense categories. Second, a model for classifying breast regions into mass and non-mass. The highest obtained accuracy was 94%. However, Cao (2021), proposed a one-stage object detection architecture (BMassDNet). First, they started with a truncation normalization method and combine it with adaptive histogram equalization to enhance the contrast between the breast mass and the surrounding environment. Next, using data augmentation and transfer learning techniques, they managed to avoid overfitting issues reaching a sensitivity of 94.3% on the DDSM dataset, and 93% on the INBreast dataset.

Classification-based approaches

Finally, for approaches that proposed a CNN architecture for breast mass detection without any previous pre-processing, Mahmood et al. (2021), proposed a CNN (ConvNet) for breast mass classification. Using the MIAS dataset along with the privately collected dataset, the ConvNet was trained and tested in order to reach its highest accuracy of 97%. Meanwhile in an extended version, Mahmood et al. (2021) introduced a ConvNet-based deep learning method for accurate mammography classification. The proposed ConvNet+SVM model achieved a remarkable accuracy of 97.8%, outperforming other methods such as VGGNet and ResNet. Their approach showed potential for supporting pathologists in predicting clinical outcomes using mammography images. However, the proposed approach did not include any filters for noise removal or image enhancement in the pre-processing phase.

Discussion of related work

To aid experts in breast mass identification, several studies have been carried out. starting with the first approach, where the purpose was to improve mammography images (Table 1). We can see that the suggested solutions take into account an evaluation strategy purely based on metrics, with no formal training procedure. However, employing a CNN model is the most effective way to assess how accurate a pre-processing method is. This was the main limit of the studies (Divyashree and Kumar 2021; Famouri et al. 2021). Moving on to the second strategy, where the work was based on a proposed CNN architecture. The pre-processing phase has a big effect on the results because having a great pre-processing strategy means providing clean data for the training phase. Hence, negligent this step and jumping directly into the classification phase is not a good method. Which was the main limit of Mahmood et al. (2021). Furthermore, the last strategy where the pre-processing phase is followed by classification is the most suitable strategy for breast mass detection for every problem of image classification. Even Though several methods do exist in this strategy according to the mentioned previous studies. However, when we have an image-included problem, CNNs are the most suitable for many reasons, mainly for the feature extraction aspect, which was one of the limits of Lbachir et al. (2021); Loizidou et al. (2022). And when we dive even further into using CNN nowadays, we should emphasize the presence of transfer learning, especially in the medical field, where we always have a lake of datasets, and as known CNNs always require a big amount of data. Thus, the usage of transfer learning techniques is currently the best solution for such an issue. Besides the dataset problem, using a pre-trained CNN model ensure a stronger and more effective architecture. That was mainly the limit for the studies (Sun et al. 2021; Chugh et al. 2022). In Cao (2021) the authors used pre-trained models to classify breast mass along with their proposed pre-processing strategy. In our study, we proposed our pre-processing strategy as well, and we proposed a model that is based on two recently built and accurate pre-trained CNNs, and our results were way better compared to the existing ones.

Table 1 Summary of the literature's review

Study	Dataset	PP[c]	CNN	TL[a]	DA[b]	Results (Accuracy)
Loizidou et al. (2022)	Collected	✓	×	×	×	90.85%
Divyashree and Kumar (2021)	CBIS-DDSM	✓	×	×	×	94.66%
Divyashree and Kumar (2021)	MIAS	✓	×	×	×	97.64%
Mahmood et al. (2021)	MIAS	×	✓	×	✓	97.05%
Sun et al. (2021)	DDSM	✓	✓	×	×	85.825
Lbachir et al. (2021)	CBS-DDSM	✓	×	×	×	90.44%
Lbachir et al. (2021)	MIAS	✓	×	×	×	94.2%
Cao (2021)	DDSM	✓	✓	✓	✓	94.3% Sensitivity
Cao (2021)	INBreast	✓	✓	✓	✓	93% Sensitivity
Famouri et al. (2021)	CBIS-DDSM	✓	×	×	×	AUC (up with 9%)
Chugh et al. (2022)	MIAS	✓	✓	×	×	93.24%
Boumaraf et al. (2020)	DDSM	✓	×	×	×	84.5%
Mahmood et al. (2021)	MIAS/INbreast	×	✓	✓	✓	97.8%
Our Approach	CBIS-DSSM	✓	✓	✓	✓	98.13%

[a]Transfer Learning techniques

[b]Data Augmentation techniques

[c]Filters included in the Pre-Processing phase

Proposed methodology

In this study we aimed to propose a new approach for breast cancer mass detection. Our approach (Fig. 2) is divided on two main parts: First, a pre-processing strategy. Second, a CNN model.

Dataset

We should underline the dearth of publicly accessible mammographic datasets first and foremost. The Curated Breast Imaging Subset (CBIS) is an updated and standardized version of the Digital Database for Screening Mammography (DDSM) (Heath et al. 1998), supplying data that is simple to get and enhanced ROI segmentation. The entire mammography image is not necessary for some CAD procedures; instead, just anomalies (the area of the image in the ROI) must be analysed. In our case, we do not need the whole mammogram image to indicate the type of the existing mass tumor either benign or malignant. Instead, we need only the mass area. Hence, CBIS-DDSM https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629 was the best choice for our study. It is publicly available in The Cancer Imaging Archive (TCIA) (Clark et al. 2013). The dataset comprising a total of 6775 participants has been split into two subsets for training and testing purposes. Specifically, 20% of the data, which corresponds to 1355 participants, has been allocated to the testing set, while the remaining 80% of the data, accounting for 5420 participants, constitutes the training set. It includes 891 mass cases and 753 calcification cases, giving it a size that may be used to analyze mammography decision support systems (Lee et al. 2017).

Pre-processing

The pre-processing phase plays a crucial role in developing an image-based detection/classification model. It is essential to present well-prepared and clear data as input during the training process to achieve a more effectively trained model. In the medical domain, image pre-processing involves three primary tasks: Region Of Interest (ROI) extraction, noise removal, and image enhancement. Each task serves a specific purpose and contributes to refining the data for subsequent analysis. Various filters designed for these tasks have been utilized in the existing literature, enabling researchers to enhance the quality of medical images and extract relevant information effectively. In this study, we used the publicly available CBIS-DDSM dataset. Several filters were used in this stage for different reasons:

- Non Local Means Denoising is a filter for noise removal, the catted ROI images were filtered with this Algorithm (Figs. 3, 4, 5 and 6).
- ROI Extraction: A. Hekal et al in Hekal et al. (2021) proved that the grey level of a breast tumor image (from CBIS-DDSM dataset) starts from the value 193. Based on that we can extract the real ROI of a tumor from an image. Thus, we created a binary image that takes only

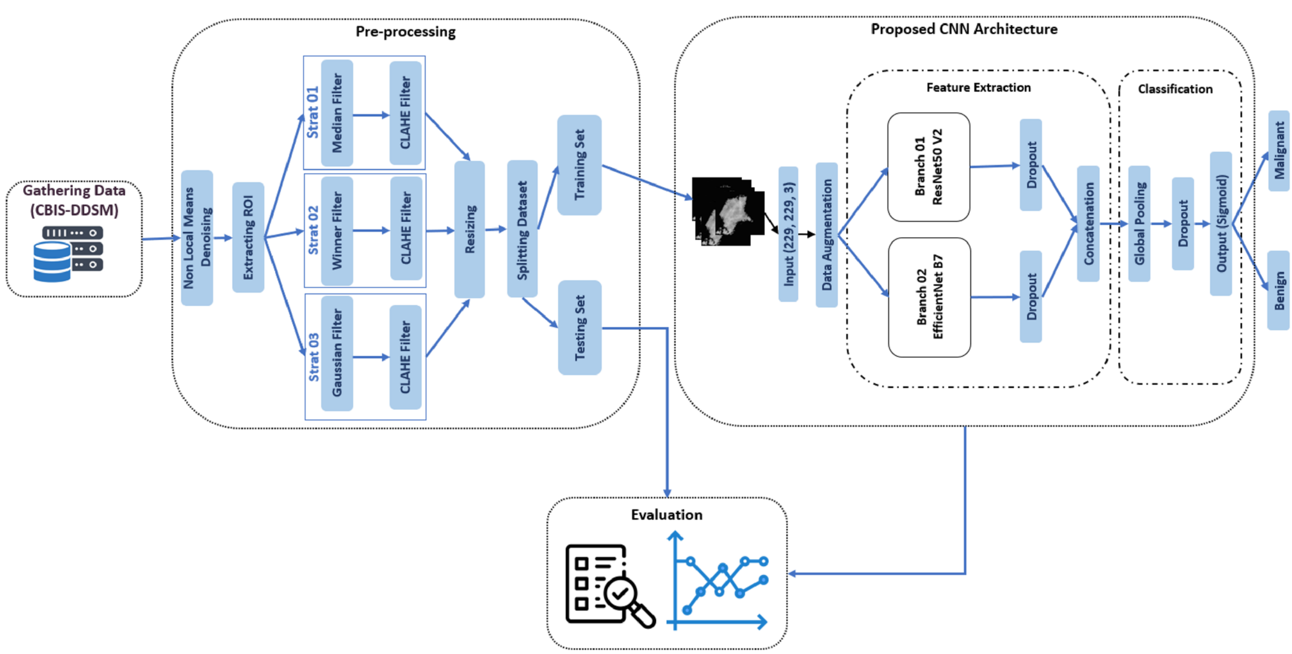


Fig. 2 Proposed approach

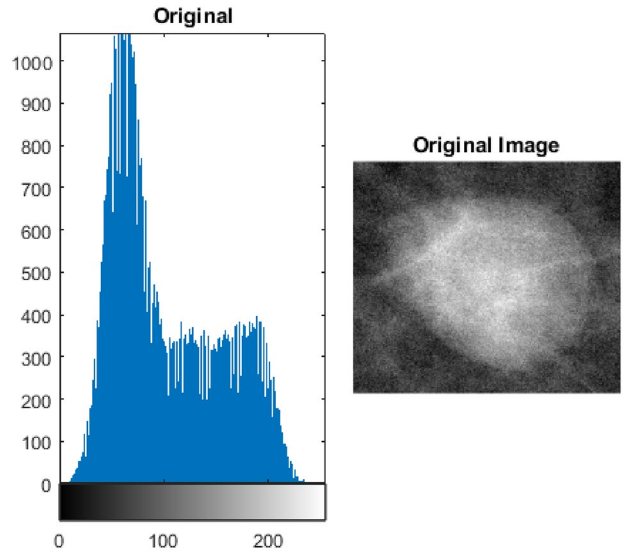


Fig. 3 Original image histogram

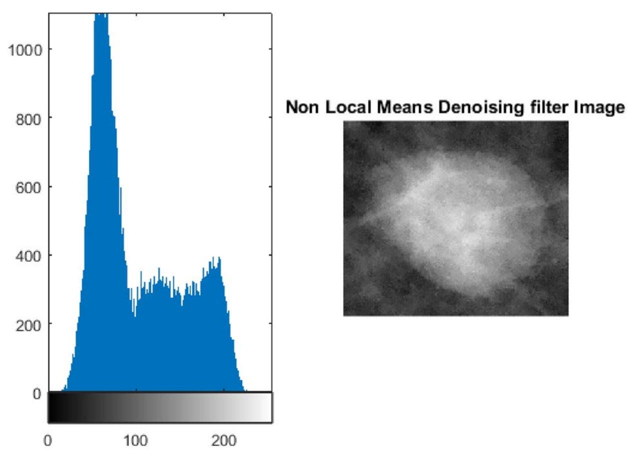


Fig. 4 Non local means denoising filter image histogram

the pixels values that are higher then 193. After that, the resulting image will be divided by the original image so we get only the tumors area as shown in Fig. 5.

- Noise Removing Filters: previous studies were satisfied by using just one noise removal filter, but as nown every noise removing filter are detected for one task. Hence, using a second round of noise removing could provide better images for training. Several filters exists for such a task, for mammogram images the mainly used filters are: Median, Wiener, and Gaussian. In order to determine the best one for our study we did three trials, each one with a filter.

  - Median: a spatial non-linear filter, is regarded as the greatest filter in the median filtering category (Al-Najdawi et al. 2015). Due to its superior noise reduction, reduced blurring, and preservation of sharp edges, it is more durable than traditional linear smoothing filters (Al-Najdawi et al. 2015).
  - Wiener: is a type of filter that improves image quality by reducing noise levels by comparing them to noise-free signal estimations (Babu and Jerome 2022). It is a filter with a minimal mean square error that will aid in restoring the performance of images that have been affected by blurring and additive noise (Babu and Jerome 2022).
  - Gaussian: is a rotationally symmetric low-pass filter that achieves smoothing by giving the central pixel more weight and the neighbors less weight, resulting in a bell-shaped curve (Bhateja et al. 2019). The Gaussian standard deviation determines how much smoothing is applied (Bhateja et al. 2019).

- After a noise removing step, it is needed to enhance the resulting images especially for mass tumor containing images (Boudouh and Bouakkaz 2022). The Contrast Limited Adaptive Histogram Equalization (CLAHE) filter is considered as a histogram equalization modification that addresses the problem of local detailing loss

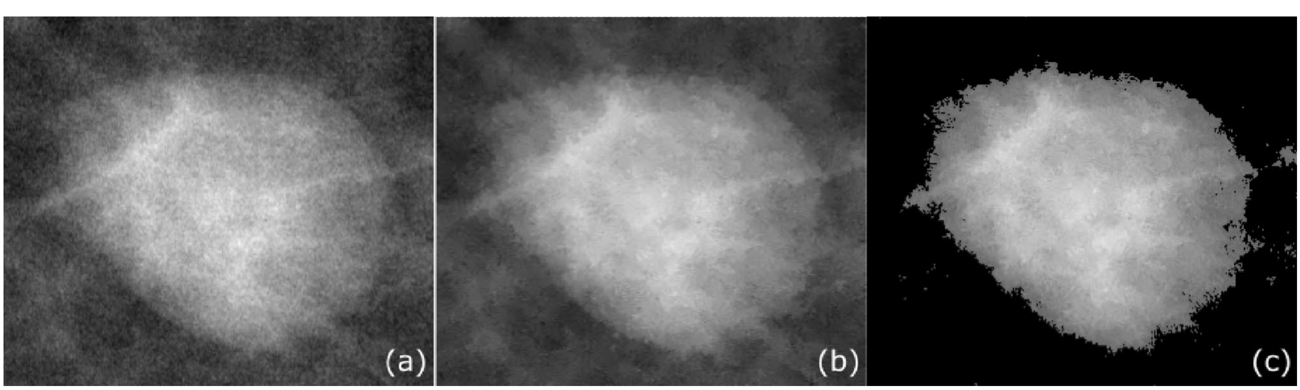


Fig. 5 **a** Original mammogram Image(CBIS-DDSM), **b** Non Local Means Denoising filter, **c** Extracted ROI

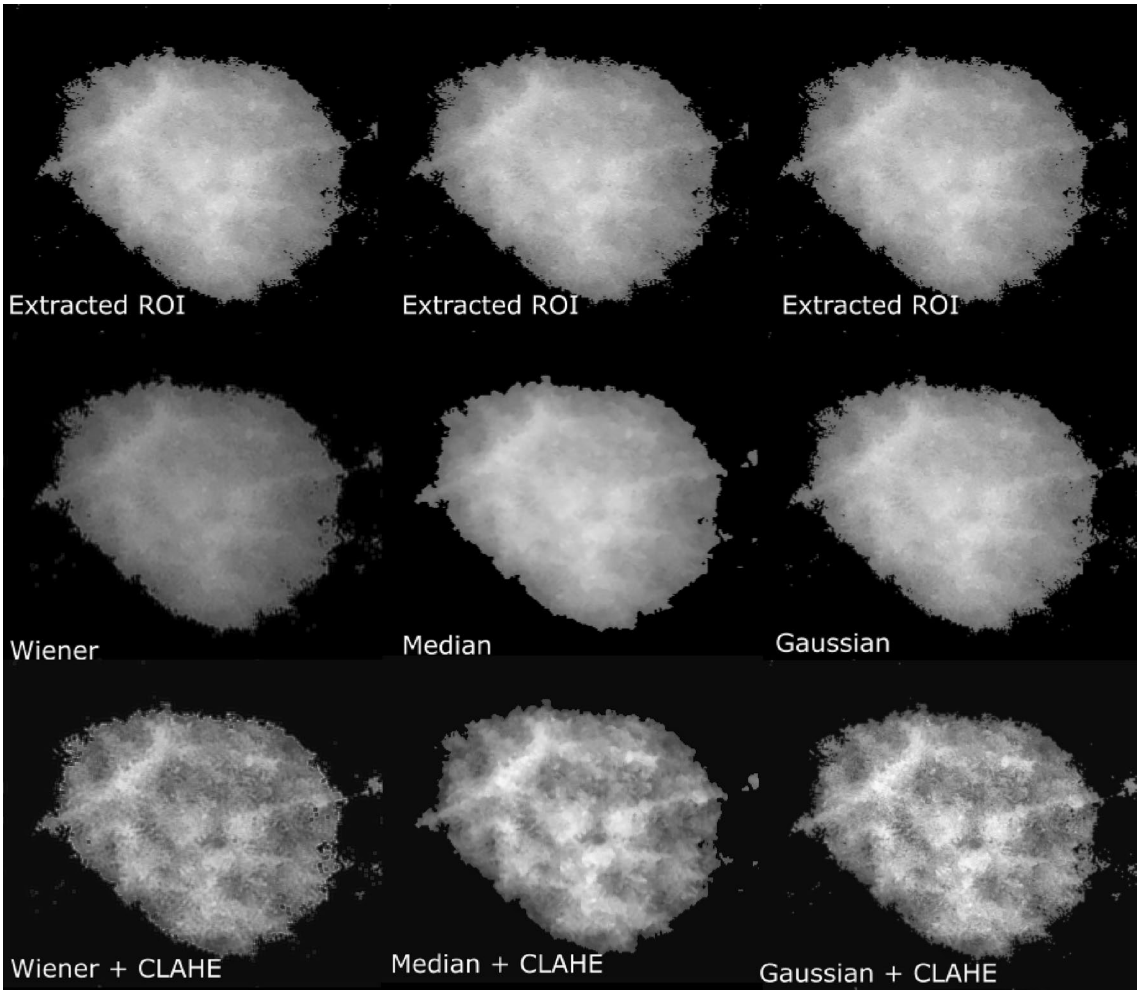


Fig. 6 Noise removing and enhancing filters

by controlling the overamplification of noise (Babu and Jerome 2022). The purpose of contrast enhancement is to enhance the mass pixel imagery's low contrast (Boudouh and Bouakkaz 2022), which made it the best filter in our case.

Proposed CNN Architecture

In our approach, we proposed a hybrid CNN architecture to classify breast Mass mammography scans into benign and malignant (Fig. 7). First, we had an input layer with (229, 229, 3) shape, after that we added a sequential data augmentation layer (Table 2). Following, the feature extraction phase, we applied transfer learning techniques. Instead of using the traditional fine-tuning technique, we combined two recent pre-trained CNNs.

Thus, our model was divided into two branches: The First one for ResNet50 V2, where we called the model with ImageNet weights, without an input nor an output layer, without including on top any fully connected layer, and all the layers were activated to trained. Based on our earlier study on breast tumor detection (Boudouh and Bouakkaz 2022), we determined that activating pre-trained CNN model layers to train will provide better results than not doing so. The same steps were taken for the second model which was EfficientNet B7. Each model was followed by a dropout with a value in the range (of 0.2, to 0.5). Then, in order to join the models, a concatenation layer was added, which receives a list of identically shaped tensors as input and outputs a single tensor that concatenates all of the inputs. Hence, instead of having features extracted from one model, we will have features extracted from two different pre-trained models. As shown in Fig 9, ResNet50 V2 produced 2048 feature maps, meanwhile, EfficientNet B7 had 2560. After the concatenation, we had 4608 feature maps.

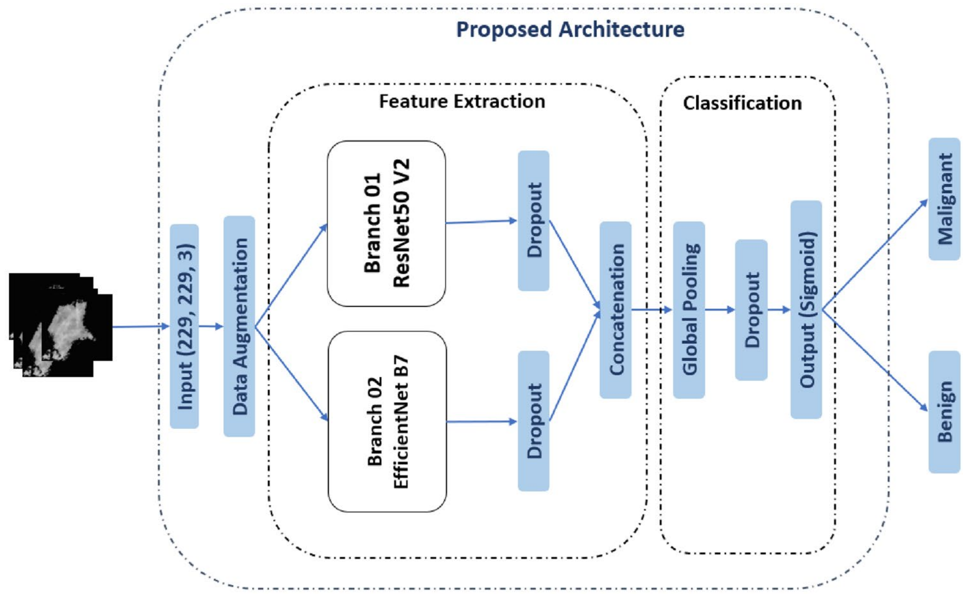


Fig. 7 Proposed CNN architecture

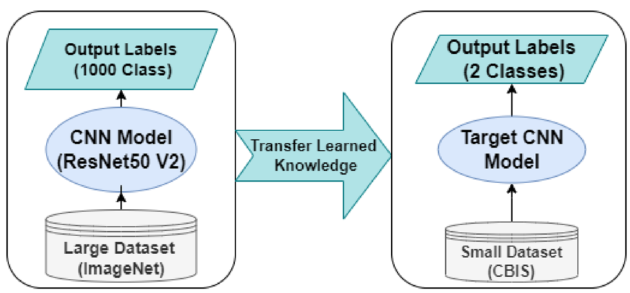


Fig. 8 Concept of transfer learning

Table 2 Data augmentation parameters

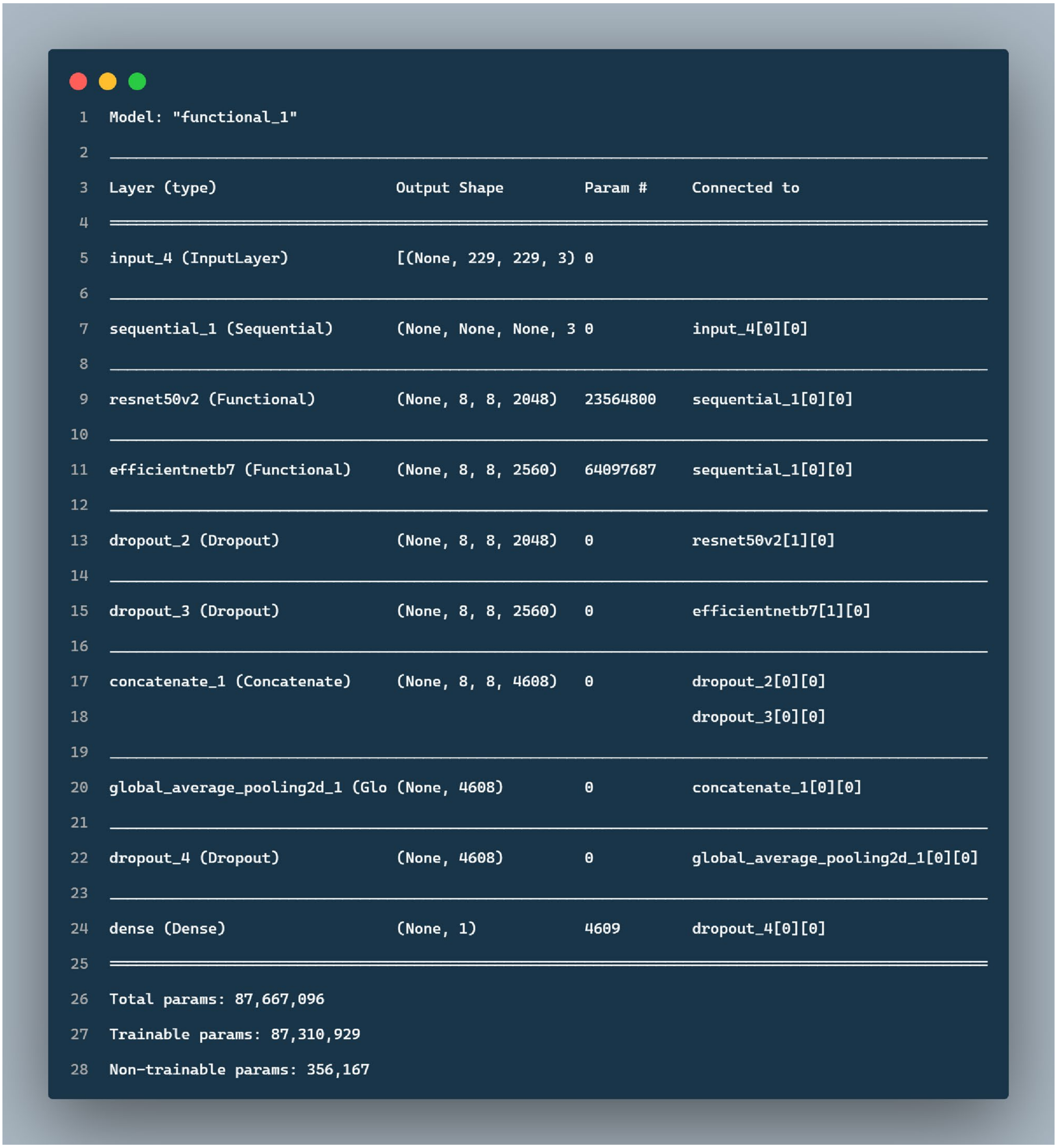
Random rotation	Range (0.125, 0.150)
Random width	Range (0.1, 0.15)
Random height	Range (0.1, 0.15)
Random zoom	Range (−0.1, 0.01)
Random translation width/height factor	Range (0.01, 0.1)
Random flip	Horizontal

Deep learning uses the phrase "transfer learning" to describe how one type of knowledge can affect another type of learning or how completing one activity can affect the completion of another (Zhou et al. 2020). Larger datasets commonly outperform smaller ones when used with CNNs. Transfer learning can be useful in CNN applications with a small dataset (Agnes et al. 2019). It involves employing a new model with a small dataset in conjunction with a trained model (such as ResNet50 V2) with a large dataset (like ImageNet) (Fig. 8). In this study, two well-known, very accurate, and recent pre-trained CNN models were chosen. Each of ResNet50 V2 and EfficientNet B7 were modified, merged, then used as a branch for breast mass detection.

- ResNet50 V2: Basic baselines for the deep Residual Network (ResNet) are mostly inspired by the concept of VGG networks (He et al. 2016). The convolutional layers consist mainly of 3*x*3 filters and follow two straightforward design concepts: First, for a given output feature map size, each layer has the same amount of filters. Second, the number of filters must be increased to maintain the time complexity per layer if the size of the feature map is cut in half (He et al. 2016). Each two-layer block in the 34-layer net is swapped out for a three-layer bottleneck block to create the 50-layer ResNet (ResNet50V2) (He et al. 2016).
- EfficientNet B7: employs a unique scaling method that uses a compound coefficient to consistently scale all dimensions, including width, depth, and resolution (Tan and Le 2019). The main concern in the data science community is that traditional ConvNets would not scale "efficiently." Compound scaling is a method that EfficientNets uses to make the scaling difficulty easier. EfficientNet B7 was the version we employed for this study.

Therefore, a combination of EfficientNet B7 and ResNet50V2 will provide a maximum amount of features, thus a reliable feature extractor network.

Moving to the classification phase, we used global pooling techniques rather than traditional ML algorithms or fully connected layers (Dense), which have been used in previous studies. Global Average/Max Pooling (GAP/GMP) is a pooling technique that replaces fully-connected layers in



```
Model: "functional_1"
__________________________________________________________________________________________________
Layer (type)                    Output Shape         Param #     Connected to
==================================================================================================
input_4 (InputLayer)            [(None, 229, 229, 3) 0
__________________________________________________________________________________________________
sequential_1 (Sequential)       (None, None, None, 3 0           input_4[0][0]
__________________________________________________________________________________________________
resnet50v2 (Functional)         (None, 8, 8, 2048)   23564800    sequential_1[0][0]
__________________________________________________________________________________________________
efficientnetb7 (Functional)     (None, 8, 8, 2560)   64097687    sequential_1[0][0]
__________________________________________________________________________________________________
dropout_2 (Dropout)             (None, 8, 8, 2048)   0           resnet50v2[1][0]
__________________________________________________________________________________________________
dropout_3 (Dropout)             (None, 8, 8, 2560)   0           efficientnetb7[1][0]
__________________________________________________________________________________________________
concatenate_1 (Concatenate)     (None, 8, 8, 4608)   0           dropout_2[0][0]
                                                                 dropout_3[0][0]
__________________________________________________________________________________________________
global_average_pooling2d_1 (Glo (None, 4608)         0           concatenate_1[0][0]
__________________________________________________________________________________________________
dropout_4 (Dropout)             (None, 4608)         0           global_average_pooling2d_1[0][0]
__________________________________________________________________________________________________
dense (Dense)                   (None, 1)            4609        dropout_4[0][0]
==================================================================================================
Total params: 87,667,096
Trainable params: 87,310,929
Non-trainable params: 356,167
```

Fig. 9 Summary of the proposed model

CNNs. Its primary concept is that one feature map should be created over the final convolutional layer for each matched class in the classification process (Lin et al. 2014). Instead of building fully-connected layers on top of the feature maps, we take the average (or maximum) of each feature map and pass the resultant vector straight into the output layer. In our case, the concatenation layer has a (8, 8, 4608) feature map. Calculating the average (or maximum) for each of the 4608 feature maps is the goal of GAP. The 4608 feature points produced by the feature maps may then be concatenated to create a 1×4608 feature vector, which we can then input into our sigmoid output for the classification. The GAP can be described using Eq. 1, for which x is the k^{th} feature map with a size of $n \times m$ in the final convolutional layer, Y_{GAP}^{k} is the output of GAP (Boudouh and Bouakkaz 2023).

The global pooling layer was added as average and max separately for each trial in order to determine the most suitable one for our study. Following it, a dropout layer was added with a value in the range (of 0.2, to 0.5). Finlay, a Sigmoid output layer is in the proposed architecture, due to the binary classification type (Fig 10).

$$Y_{GAP}^{k} = \frac{1}{n \times m} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} x_{ij}^{k} \tag{1}$$

Multiple experiments were performed to determine the main criteria that could influence the models' performance. Table 3 shows the hyper parameters.

Experimental setup

We used Matlab 2017a for the pre-processing stage. For the training and testing, we used TensorFlow (TensorFlow 2023) version 2.9.1 with python 3.9.13 and Keras version 2.9.0. As for the used machine, had an Intel® Core™ i7-10870 H 2.20GHz processor and 16GB of RAM, Nvidia RTX 3060 GPU with a performance capacity of 16 GB.

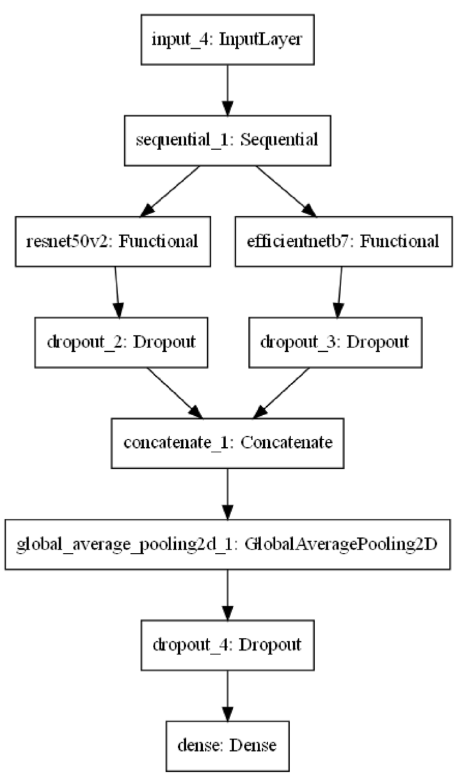


Fig. 10 Architecture of the proposed model

Table 3 Hyper parameters

Global pooling layers	Max/Average
From logits	True
Dropout layer	Range (0.2, 0.5)
Optimizer	Keras SGD
Learning rate	0.0001/0.001
Momentum	0.9
Nesterov	True
Metrics	Binary accuracy/ Binary crossentropy (Loss)
Number of epochs	100

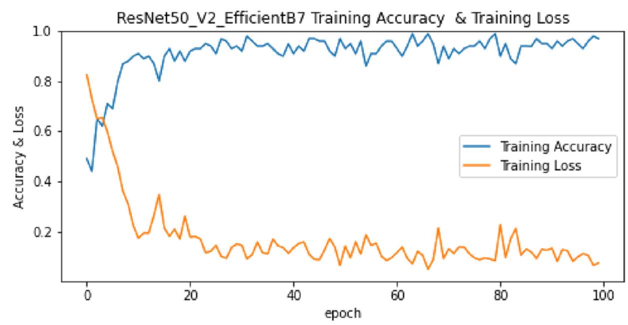


Fig. 11 Training accuracy and loss of the gaussian strategy

Results and discussion

In this section, we will state the achieved results, starting with the training results and then passing on to the evaluation results. After that, we will discuss and compare our obtained outcomes with state-of-the-art approaches.

As mentioned before several trials with different hyper-parameters and with different pre-processing strategies were applied in our study. Yet, all the trials had 100 training epochs and every trial took five hours. In the training phase, accuracy and loss were chosen as the metrics. Figs 11, 12, 13 are presenting the graphs of training process accuracy and loss for every best-obtained results trial, of the three proposed pre-processing strategies.

The graphs illustrates the proposed hybrid CNN models' accuracy (ascending) and loss (descending) across the training set. For each strategy the training graphs are close together, confirming that the proposed hybrid model trained adequately. Furthermore, to confirm that Table 4 explains better the best trials obtained training accuracy. Both the strategies of Median and Gaussian had the highest training accuracy of 100%, yet the Median filter strategy achieved a better loss outcome than the others.

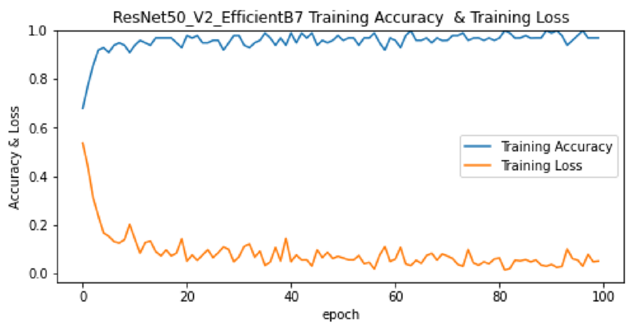


Fig. 12 Training accuracy and loss of the median strategy

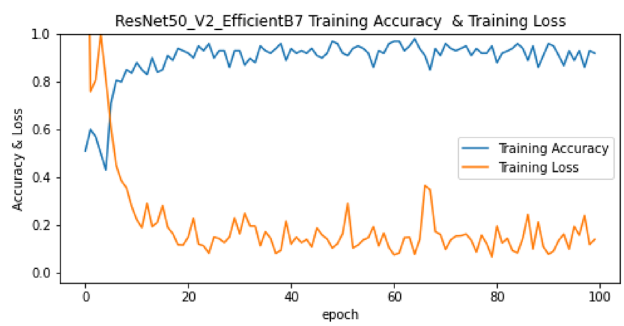


Fig. 13 Training accuracy and loss of the wiener strategy

Several evaluation metrics exist in the machine and deep learning field. Among those metrics we used: confusion matrix (Figs 14, 15, 16), accuracy (2), specificity (4) also known as True Negative Rate (TNR), sensitivity (3)also known as True Positive Rate (TPR) or recall, precision (5), F1-Score (6), and the Area Under the Curve(AUC) (Fig 17).

$$Accuracy = \frac{TP + TN}{TN + FP + FN + TP}\% \tag{2}$$

$$Sensitivity = \frac{TP}{FN + TP}\% \tag{3}$$

$$Specificity = \frac{TN}{TN + FP}\% \tag{4}$$

$$Precision = \frac{TP}{TP + FP}\% \tag{5}$$

$$F1Score = \frac{2 \times Specificity \times Precision}{Specificity + Precision}\% \tag{6}$$

- TP: True Positive is the total number of truly identified malignant breast masses.

Table 4 Final training results

Strategy	Accuracy	Loss
Gaussian	100%	0.01292
Median	100%	0.00302
Wiener	85%	0.17848

- TN: True Negative is the total number of truly identified benign breast masses.
- FP: False Positive is the total number of images predicted by the model to have malignant breast masses and have benign barest masses in fact.
- FN: False Negative is the total number of images predicted by the model to have benign breast masses and have barest malignant masses in fact.

In this study, a testing set was used to evaluate each model's performance. Several hyper-parameters were changed in each trial, yet the main criteria that did influence the outcomes were: The choice of the global pooling layer (either GAP or GMP), the learning rate value, and the last dropout layer value.

The best-obtained trials are shown in Table 5 where we can see that the learning rate value, the type of global pooling layer that was employed after concatenating the two pretrained models, and the value of the final dropout layer were the key outcomes influencers.

In the strategy where the Gaussian filter was applied for noise removal, we can see that using a final average pooling layer was more beneficial than maximum pooling. Yet, following the GAP with a high dropout (such as 0.35) did not achieve great results as opposed to a reduced value (like 0.2). Furthermore, for all the trials reducing the final dropout layer value outcomes great results, specifically the 0.2 value. On the other hand, the most successful tests employing the wiener filter revealed that GMP yields superior outcomes when using that filter. The GAP results, however, were far superior to the Wiener filter trials for both the Gaussian and Median filters. As for the learning rate, for all the trials it was clear that reducing it will achieve better results.

Moving to the obtained evaluation results, the pre-processing strategy that included the Gaussian filter achieved the best testing accuracy of 96%, where the training accuracy was 100%. Yet, the TPR was 98.26%, much higher than the TNR, which was 92.41%. Furthermore, when comparing all the best trials of this filter, always the obtained TPR was superior to the TNR. Knowing that only the hyper-parameters were changed, and the models were basically the same, especially the extracted feature maps, the main reason behind the outcomes should be the selected noise removal filter itself. Hence, the Gaussian filter is not well suited for cases that include malignant masses compared to

Table 5 Best obtained trials results

Filter	GP[a]	LR[b]	Dropout	TrainAcc[c]	Loss	TestAcc[c]	TNR	TPR	AUC	Precision	F1-Score
Gaussian	A[d]	0.001	0.35	100%	0.088	95.46%	89.65%	99.10%	0.94	93.82%	91.69%
Gaussian	A[d]	0.0001	0.2	100%	0.012	96%	92.41 %	98.26 %	0.95	95.35%	93.86%
Gaussian	M[e]	0.0001	0.5	69.99%	0.46	89.06%	84.13%	92.17%	0.88	90.21%	87.06%
Median	M[e]	0.001	0.35	100%	0.065	97.33%	100%	95.65%	0.98%	100%	100%
Median	A[d]	0.0001	0.2	100%	0.003	98.13%	100%	96.95%	0.98%	100%	100%
Median	M[e]	0.001	0.5	60%	0.559	54.13%	94.92%	13.95%	0.54%	61.53%	74.66%
Wiener	M[e]	0.0001	0.2	85%	0.178	94.13%	86.2%	99.13%	0.93	91.93%	88.97%
Wiener	A[d]	0.001	0.35	94.99%	0.125	90.4%	89.65%	90.86%	0.9	97.66%	93.48%
Wiener	M[e]	0.001	0.5	44.99%	0.698	61.33%	100%	0%	0.5	0%	0%
Median[f]	A[d]	0.0001	0.2	20.00%	0.717	95.46%	99.13%	89.65%	0.94%	98.48%	98.80%
Median[g]	A[d]	0.0001	0.2	85.00%	0.228	86.66%	83.91%	91.03%	0.87%	78.10%	80.90%

[a]Global Pooling Layers Learning rate, [b]Accuracy, [c]Average, [d]Max, [e]ResNet50V2, [f]EfficientNetB7

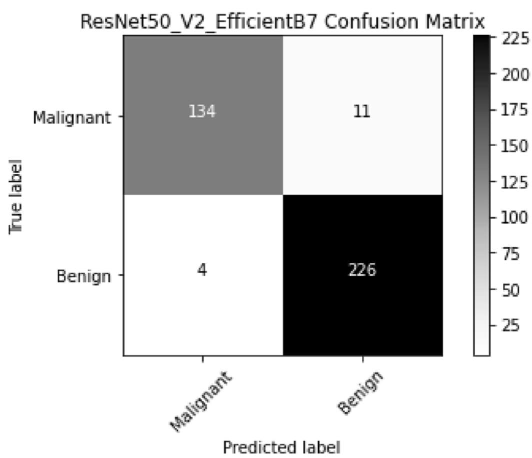


Fig. 14 Gaussian Strategy Confusion matrix

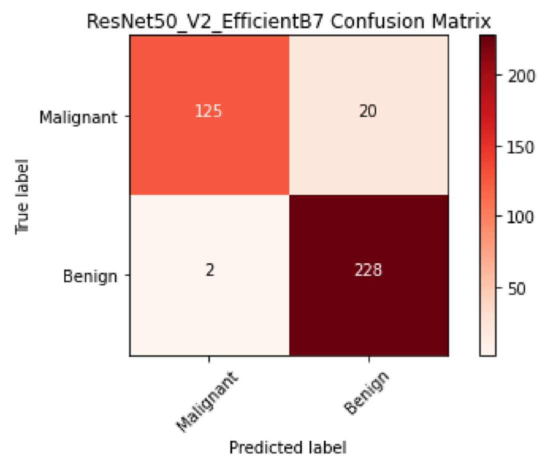


Fig. 16 Wiener Strategy Confusion matrix

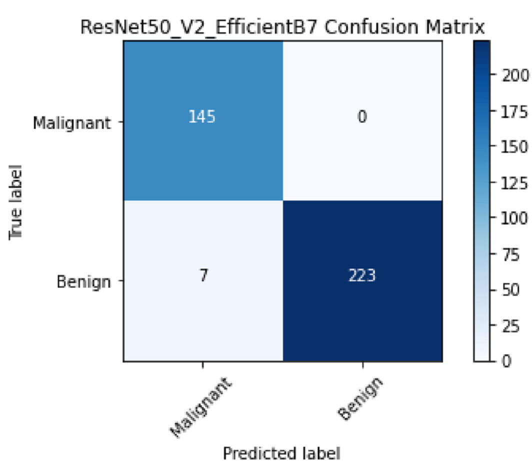


Fig. 15 Median Strategy Confusion matrix

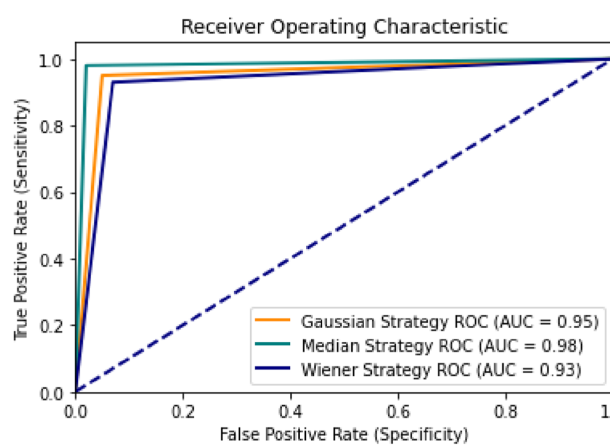


Fig. 17 ROC Curves

Table 6 Comparison between the proposed approach with existing ones

Study	Dataset	Accuracy	TNR	TPR	AUC	Precision	F1-Score
Loizidou et al. (2022)	Collected	90.85%	91.67%	89.66%	0.91	/	/
Divyashree and Kumar (2021)	CBIS-DDSM	94.66%	91.21%	90.27%	/	/	/
Divyashree and Kumar (2021)	MIAS	97.64%	98.53%	92.15%	/	/	/
Mahmood et al. (2021)	MIAS	97.05%	/	/	/	/	/
Sun et al. (2021)	DDSM	85.82%	95.38%	96%	/	50.81%	66.31%
Lbachir et al. (2021)	CBS-DDSM	90.44%	90, 85%	/	0, 90	/	/
Lbachir et al. (2021)	MIAS	94.2%	93, 15%	/	0, 95	/	/
Cao (2021)	DDSM	/	94.3%	/	/	/	/
Cao (2021)	INBreast	/	93%	/	/	/	/
Famouri et al. (2021)	CBIS-DDSM	/	/	/	up with 9%	/	/
Chugh et al. (2022)	MIAS	93.24%	84.74%	93.03%	/	87.23%	/
Boumaraf et al. (2020)	DDSM	84.5%	84.4%	79.3%	/	/	/
Mahmood et al. (2021)	MIAS/INbreast	97.8%	97.7%	/	0.91	97.8%	97.6%
Our Approach (Gaussian)	CBIS-DDSM	96%	92.41%	98.26%	0.95	95.35%	93.86%
(Median)	CBIS-DDSM	98.13%	100%	96.95%	0.98%	100%	100%
(Wiener)	CBIS-DDSM	94.13%	86.2%	99.13%	0.93	91.93%	88.97%

benign masse, since the TNR is always lower than TPR, thus the TP is always poor.

Moreover, the same equation was achieved by the Wiener filter, where we can see the big difference between the TNR and TPR, especially in the best-obtained trial where the TPR reached the maximum and best result among all the trials with 99.13%, yet the TNR was 86.20%. Thus, the Wiener filter failed to suit and reduce noise for malignant mass cases, depending on the TP number achieved. In contrast, even in comparison to other filters, the TN attained its maximum using the Wiener filter.

However, using the Median filter, yielded the total number of TP right with 100% TNR in the two best cases. Followed by a lower TPR. This leads to conclude that the Median filter was the most suitable filter for malignant mass cases, reaching the total number of true positives. Lastly, comparing the test accuracy, we can declare that the Median filter achieved the best results with the highest accuracy of 98.13%.

Meanwhile, when we tested the best pre-processing strategy with just one feature extractor, we can notice (Table 5) that the usage of the ResNet50V2 as a feature extractor yielded better results than EfficientNetB7. However, when observing the training results, the trial of ResNet50V2 had a low accuracy than the testing results, which conclude that there was an underfitting problem. Thus, even though the test accuracy of the ResNet50V2 trial was better, but EfficentNetB7 had a better training process where no underfitting occurred. Furthermore, The results of the hybrid feature extractor yielded even better results, proving that the combination idea was more suitable for breast Masses in mammography. Hence, the number of extracted features is a main criterion that influences the performance of the classifier of breast Masses types in mammogram images, in which the increase of obtained feature maps outcome better results. That was proved in our previous study for breast tumor detection in mammography.

Moreover, the ROC curve was closer to the upper left corner, and the AUC (Area Under Curve) was 0.98, as shown in the ROC curve (Fig. 17) for the best trial of the Median filter, which was expected given that only seven benign images were mistaken and classified as malignant Masses from 230 test images (Fig. 15), followed by the Gaussian filter with o.95 AUC which was mistaken on both classes (Fig. 14). On the other hand, the Wiener filter trial had the lowest AUC value of 0.93 misclassifying samples from both classes as well (Fig. 16). Furthermore, for all the best trials, the ROC curve was close to the upper left corner but not as much as in the best case of the Median filter 0.99 AUC.

As mentioned before, in the state of the art of breast mass classification into benign and malignant, there have been several proposed solutions. However, when compared to previously proposed solutions (Table 6), our approach achieved better results. Several reasons are behind that. Starting with the chosen dataset, the MIAS dataset (Divyashree and Kumar 2021; Mahmood et al. 2021; Lbachir et al. 2021; Chugh et al. 2022) is considered a small dataset for training and testing. Additionally, with the existence of more accurate extracted ROI datasets such as CBIS-DDSM, there is no need to use full mammography as an input when the problem is tumor type classification because the proposed model is not required to distinguish between normal tissue and tumor, it must be able to distinguish between malignant and benign tissue. Thus, the ROI is just the tumor's area, not the whole organ.

Second, the pre-processing technique is essential in any machine or deep learning tasks, particularly when images are involved. Providing clear and noiseless images, produce great results always. However, with the large existing filters and strategies, specifically for mammogram images, creating an effective pre-processing phase has become challenging. Proposing a new strategy, divided into three different trials, where several filters were included was not proposed before as a first step into breast mass detection in the literature. Instead, studies proposed strategies that included two filters in a row as a maximum (Loizidou et al. 2022; Divyashree and Kumar 2021; Lbachir et al. 2021; Cao 2021; Famouri et al. 2021; Chugh et al. 2022; Boumaraf et al. 2020), and some of them were not even followed by a classification stage in order to prove prove their efficiency (Loizidou et al. 2022; Divyashree and Kumar 2021; Lbachir et al. 2021; Famouri et al. 2021; Mahmood et al. 2021).

On the other hand, when compared to the study that employed transfer learning techniques for breast mass detection (Cao 2021), our results were way superior. Even though (Cao 2021) used ResNet (101), which we used as well in our hybrid model with a different version (ResNet50 V2). That was due to the new hybridization of the transfer learning algorithms approach that we proposed. Instead of modifying one pre-trained CNN model, we concatenated two recent and efficient CNNs with a major modification, to create an accurate breast mass detection model.

Moreover, the strategy that includes the median filters reached the best and maximum accuracy compared to other strategies including ours. Which made it the best strategy for breast mass detection in mammography using this dataset. However, we concluded also that for cases that included benign mass tumors the Wiener filter was the most effective, reaching the highest number of true negatives of 228 (Fig. 16). On the other hand, for malignant cases, the Median filter was the most suitable with the highest number of true positives of 145 (Fig. 15). Yet, the low performance of the Gaussian and Wiener filters on the detection of malignant cases can be due to the shortage of images in this class, hence it is caused by unbalanced data.

Conclusion

In this study, we proposed a new approach to breast mass classification. Divided into two parts, our approach starts with a pre-processing phase that included several filters. Three strategies were proposed in this phase. Secondly, we suggested a new novel CNN architecture for detecting breast Masses. Where the feature extraction phase included a concatenation of two recent pre-trained CNNs (ResNet50 V2 and EffiecentNet B7) instead of traditional feature extraction fine-tuning, guaranteeing a maximum amount of extracted features. Meanwhile, in the classification phase rather than using a traditional ML algorithm or fully connected layer, we used global pooling techniques. As a result, by creating such a model and training it using our pre-processed dataset, we established that the pre-processing strategy that included the Median filter as noise removal achieved the best outcomes of 98.13% accuracy, followed by the Gaussian filter trial with 96% accuracy. Meanwhile, the Wiener filter trial achieved the lowest accuracy of 94.13%. Additionally, we determined as well that the usage of the strategy that included the Median filter will provide well-clean malignant Mass mammography scans, in which the TPR was 100%. On the other hand, the usage of the Wiener filter will suit better benign Mass mammograms with a maximum TNR of 99.13%. Additionally, the attained results, where we were able to improve each of the current accuracy, TNR, TPR, precision, F1-Score, and AUC, further show that the suggested CNN architecture is well suited for such a situation.

However, this work contains some limitations. Even though we used a large dataset compared to the existing mammography databases, using a second or even third dataset might be better. In addition to that, despite the selected pre-trained CNNs being recent and accurate, still concatenating other pre-trained CNN may also achieve better results.

Thus, as future perspectives:

- Since the hybridization method achieved great results, it can be used with other models in other problems (such as breast calcification classification).
- We can try to use the Median filter for malignant mass cases, and the Wiener filter for benign mass-containing mammograms, as a new strategy that will increase the TNR and TPR, thus the accuracy.
- Since the proposed pre-processing strategies were successful, they could be modified and used on other organ tumors (such as brain tumors).

Author contributions SSB wrote the main manuscript text. All authors reviewed the manuscript.

Funding This study was not funded. The authors have no relevant financial or non-financial interests to disclose.

Availability of data and materials The dataset analysed during the current study are available in: CBIS-DDSM https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629.

Declarations

Conflict of interest Authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Agnes SA, Anitha J, Pandian SIA, Peter JD (2019) Classification of mammogram images using multiscale all convolutional neural network (MA-CNN). J Med Syst 44:1–9

Al-Najdawi N, Biltawi M, Tedmori S (2015) Mammogram image visual enhancement, mass segmentation and classification. Appl Soft Comput 35:175–185

Babu A, Jerome SA (2022) Mammogram image grade gauging of denoising filters; enhancement methods. In: 2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), pp. 1–6. https://doi.org/10.1109/iceeict53079.2022.9768548

Bhateja V, Misra M, Urooj S (2019) Non-linear filters for mammogram enhancement

Boudouh SS, Bouakkaz M (2022) Breast cancer: Breast tumor detection using deep transfer learning techniques in mammogram images. In: 2022 International Conference on Computer Science and Software Engineering (CSASE), pp. 289–294. https://doi.org/10.1109/csase51777.2022.9759702

Boudouh SS, Bouakkaz M (2023) Breast cancer: toward an accurate breast tumor detection model in mammography using transfer learning techniques. Multimedia Tools Applic

Boumaraf S, Liu X, Ferkous C, Ma X (2020) A new computer-aided diagnosis system with modified genetic feature selection for bi-rads classification of breast masses in mammograms. BioMed Res Int. https://doi.org/10.1155/2020/7695207

Cao H (2021) Breast mass detection in digital mammography based on anchor-free architecture. Comput Methods Prog Biomed 205:106033

Chugh S, Goyal S, Pandey A, Joshi S (2022) Morphological and otsu's technique based mammography mass detection and deep neural network classifier based prediction. Traitement du Signal 39(4):1283–1294

Clark KW, Vendt BA, Smith KE, Freymann JB, Kirby JS, Koppel P, Moore SM, Phillips SR, Maffitt DR, Pringle M, Tarbox L, Prior FW (2013) The cancer imaging archive (tcia): Maintaining and operating a public information repository. J Digit Imag 26:1045–1057

DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Sauer AG, Jemal A, Siegel RL (2019) Breast cancer statistics. CA Cancer J Clin. https://doi.org/10.3322/caac.21583

Divyashree BV, Kumar GH (2021) Breast cancer mass detection in mammograms using gray difference weight and mser detector. SN Comput Sci 2:63

Famouri S, Morra L, Mangia L, Lamberti F (2021) Breast mass detection with faster r-cnn: On the feasibility of learning from noisy annotations. IEEE Access 9:66163–66175

He K, Zhang X, Ren S, Sun J (2016) Identity mappings in deep residual networks. CoRR https://arxiv.org/abs/1603.05027

Heath M, Bowyer K, Kopans D, Moore R, Jr, PK (1998) The digital database for screening mammography, 10

Hekal AA, Elnakib A, Moustafa HE-DS (2021) Automated early breast cancer detection and classification system. Signal Image Video Proc 15:1497–1505

Jafari SH, Saadatpour Z, Salmaninejad A, Momeni F, Mokhtari M, Nahand JS, Rahmati M, Mirzaei H, Kianmehr M (2018) Breast cancer diagnosis: imaging techniques and biochemical markers. J Cell Physiol 233:5200–5213

Khan A, Sohail A, Zahoora U, Qureshi AS (2020) A survey of the recent architectures of deep convolutional neural networks. Artific Intell Rev 53:5455–516

Lbachir IA, Daoudi I, Tallal S (2021) Automatic computer-aided diagnosis system for mass detection and classification in mammography. Multim Tools Appl 80:9493–9525

Lee RS, Gimenez F, Hoogi A, Miyake KK, Gorovoy M, Rubin D (2017) A curated mammography data set for use in computer-aided detection and diagnosis research. Scient Data 4:1–9

Lin M, Chen Q, Yan S (2014) Network in network. CoRR https://arxiv.org/abs/1312.4400

Loizidou K, Skouroumouni G, Nikolaou C, Pitris C (2020) An automated breast micro-calcification detection and classification technique using temporal subtraction of mammograms. IEEE Access 8:52785–52795

Loizidou K, Skouroumouni G, Savvidou G, Constantinidou A, Nikolaou C, Pitris C (2022) Benign and malignant breast mass detection and classification in digital mammography: The effect of subtracting temporally consecutive mammograms. 2022 IEEE-EMBS Int Conference Biomed Health Inform (BHI) 1–4

Mahmood T, Li J, Pei Y, Akhtar F (2021) An automated in-depth feature learning algorithm for breast abnormality prognosis and robust characterization from mammography images using deep transfer learning. Biology 10:859

Mahmood T, Li J, Pei Y, Akhtar F, Jia Y, Khand ZH (2021) Breast mass detection and classification using deep convolutional neural networks for radiologist diagnosis assistance. 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), 1918–1923

Mahmood T, Li J, Pei Y, Akhtar F, ur Rehman M, Wasti SH, (2022) Breast lesions classifications of mammographic images using a deep convolutional neural network-based approach. PLoS One 17:e0263126

Society medical and editorial content team, T.A.C.: About Breast Cancer. Technical report (2019). https://www.cancer.org/content/dam/CRC/PDF/Public/8577.00.pdf Accessed 2022-02-18

Sun L, Sun H, Wang J, Wu S, Zhao Y, Xu Y (2021) Breast mass detection in mammography based on image template matching and cnn. Sensors (Basel, Switzerland) 21:2855

Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71:209–249

Tan M, Le QV (2019) Efficientnet: Rethinking model scaling for convolutional neural networks. ArXiv https://arxiv.org/abs/1905.11946

TensorFlow. https://www.tensorflow.org/ Accessed 2023-01-10

Waks AG, Winer EP (2019) Breast cancer treatment: a review. JAMA 321:288–300

Wellings EH, Vassiliades LA, Abdalla R (2016) Breast cancer screening for high-risk patients of different ages and risk - which modality is most effective? Cureus 8(12):e945

Zhou J, Yang X, Zhang L, Shao S, Bian G (2020) Multisignal vgg19 network with transposed convolution for rotating machinery fault diagnosis based on deep transfer learning. Shock Vibrat 2020:1–12