Grad-CAM Guided Preprocessing and Convolutional Neural Network for Efficient Mammogram Images Classification

[1*]Belaggoune Mohammed, [2]Akou Anouar, [3]benblidia Nadjia
[1]National Higher-school of Computer Science, Algeria, Laboratoire des Méthodes de Conception de Systèmes (LMCS).
[2]The University of Science and Technology Houari-Boumediene, Algeria.
[3]Saad Dahlab University of Blida1, Algeria.
E-mail : m_belaggoune@esi.dz, anouar.onur@gmail.com, benblidia@univ-blida.dz
*Corresponding author

Keywords: deep learning, computer aided diagnosis system, grad-CAM, detection, classification, mammograms, breast cancer

Received: April 25, 2023

Breast cancer is now widely recognized as the second most deadly disease in women. Computer Aided Diagnosis system (CAD), and specifically deep learning (DL), have continued to provide a significant computational solution for early detection and diagnosis of this disease. Research efforts are advancing new approaches to improve the performance of deep learning-based models. We are developing a new system (CAD) to classify mammograms as normal or abnormal and then classify the abnormal ones as benign or malignant. The proposed system consists of three main steps. The pre-processing step consists mainly in cropping the images with the application of Grad-CAM (Class Activation Map) method which generates a heatmap that facilitates the detection of the area of interest to be cropped and minimizes the pre-processing steps usually done in previous works. A contrast enhancement is performed with the Contrast Limited Adaptive Histogram Equalization (CLAHE) method to differentiate the distinct elements of the image. We then proceed to data augmentation with two different ways combining the usual geometric transformations such as rotations, shifts and translations, before proceeding to the training of the different proposed models, namely: VGG16, Vgg19, Resnet50, Densenet121 and InceptionV3, thanks to which we achieved the highest accuracy of 99.13% for the benign and Malignant classification using the pre-trained network Resnet50, and an accuracy of 98.54% with the pre-trained network Vgg19.

Povzetek: Razvit je CAD (computer aided diagnosis) sistem z globokim učenjem in razširjanjem podatkov za zgodnje odkrivanje in diagnozo raka dojk. V primerjavi z obstoječimi metodami dosega visoko točnost pri klasifikaciji mamogramov.

1 Introduction

Today breast cancer is the major cause of women's death throughout the world. It spreads through cells and develops in the body in a regular way. Regarding women of all ages, according to GLOBOCAN [1], Algeria has 12,536 new cases of breast cancer out of 31,090 new cases of cancer in 2020, which is about 40.3%.

Several analyses have confirmed that early detection is the best way to improve the prognosis of life. Mammography plays an important role in diagnosis. This examination is the most effective in all aspects of early detection and monitoring for breast cancer. It is able to detect lesions that are difficult to spot. The screening of this dangerous cancer by mammography increases the number of daily cases to be analyzed by expert radiologists who will be faced with difficulties in the interpretation of mammographic images, which will create a risk of falling into error in some cases, also, the high cost when it comes to the need for a second reading of images by other radiologists.

Thus, the fight against cancer is far from over. Medicine is advancing on all fronts to improve patient care and defeat this disease of the century. Therefore, it is essential that several disciplines continue to contribute, including data mining and artificial intelligence. To provide robust and reliable medical diagnostic support, several computer-aided diagnosis (CAD) techniques have been proposed to speed up the process and assist in decision making [2] [3]. In the current CAD trend, machine learning techniques are at the forefront, and with the emergence of deep learning (a kind of machine learning), neural networks can be a powerful tool for distributed diagnostic applications.

Machine Learning has a primary place in the design of intelligent diagnostic systems, it consists of certain steps that are linked together, but the major disadvantage is that the steps are done separately, so if there is an error in one step, there is no guarantee that the propagation will not take place, and the result will be affected. Deep Learning is a branch of Machine Learning, but the contribution is that the steps are done in an automatic way which reduces the risk of error.

In our paper, a new computer-aided diagnosis (CAD) system was proposed and implemented to first classify mammograms into Normal and Abnormal, and then classify breast tumors in mammograms as malignant or benign.

In order to make classifications of different medical images, we have an interest in having a large dataset, well labeled to allow the learning network to properly set the weights of the nodes in order to minimize the error, where we are going to make a decision towards a health situation of a patient. The importance of the dataset is realized in the good quality of the instances as well as the number of them.

To do this, we use a very well-known and reliable database with respect to its labeling MIAS (Mammographic Image Analysis Society), but with a much reduced number, consequently we proceed in our approach to a data augmentation in order to satisfy our network thanks to several geometric transformations which will be illustrated in methodology section. This augmentation is indeed carried out after some primordial pre-processing for our approach, we mention the Grad-CAM which is the pillar of our method, which allowed us to reduce the number of pre-processing compared to previous works related to the problem, furthermore, it allows us to make a rather correct cropping avoiding us to proceed to the suppression of pectoral muscle and the various artifacts of the images and to locate the region of interest. A contrast enhancement is also applied with CLAHE method after the cropping for a better quality of the images. To achieve the classification of breast cancer, initially in normal/abnormal, then in malignant or benign using mammography images, We have specified some useful techniques to reduce the overfitting as well as to improve the speed of the model training, namely dropout and batch normalization during the training stage of the different models chosen mainly used and proved their efficiency, namely Vgg16, Vgg19, Resnet50, densenet121 and inceptionV3.

In this paper, we will demonstrate our contribution after evocating some related important works as follows:

- We will describe the MIAS database which is an essential resource for the foundation of our results ;
- We will present the Grad-CAM method for image cropping which uses the gradient of the classification score with respect to the convolutional features determined by the network in order to understand which parts of the image are most important for classification.
- We will define the different data augmentations applied.
- We introduce a CNN training baseline, which produces competitive results to the state-of-the-art methods by itself.
- We present comprehensive experiment on MIAS dataset, the experiment results demonstrate that our method achieves superior performance over the state of art, in the form of tables, graphs, curves, comparisons and all the necessary analyses and discussions, based on the performance metrics

2 Related works

Breast cancer is currently the second leading cause of cancer death in women worldwide, and early detection of breast lesions is key to increasing the chance of survival and reducing the mortality rate [4][5]. Although digital mammography is a clinically accepted imaging method for detecting breast cancer in its early stages, radiologists face some challenges when analyzing mammograms for signs of cancer. Computer technologies and mathematical models can increase the detection rate and, therefore, effectively reduce the mortality rate from breast cancer [6][7]. The decision support system that relies on Deep Learning can provide radiologists with reliable assistance in detecting abnormalities in mammographic images [8][9]. Indeed, breast cancer detection using digital mammography improves when radiologists can use an artificial intelligence system as a diagnostic support tool, without requiring additional reading time [10][11]. Currently, computer-aided diagnosis is a very common and efficient method that analyses mammograms and helps radiologists in interpretation to detect the suspicious lesions as well as their type. Different techniques and methods have been studied for the classification of mammograms into normal abnormal, or benign and malignant classes.

- Ting F. F et al [12] presented a method based on Convolutional Neural Network for mammogram images classification as malignant, benign or normal. The method consists of three main steps:(a) feature wise pre-processing (FWP) for region of interest (ROI) extraction, (b) ROI classification based on convolutional neural network (CNN) of 28 convolutional layer (c) interactive detection-based lesion locator (IDBLL).
- Bakkouri.I et al [13] proposed a novel approach to distinguish normal from abnormal breast tissue using Pyramid-CNN for multiscale analysis. In order to improve image processing time, they extracted

representative region proposals from each mammogram using the Hessian operator determinant, their methodology was evaluated on a publicly available mammography dataset, such as the Breast Cancer Digital Repository (BCDR) database,

- Agnes. S. A et al [14] proposed a Multiscale All Convolutional Neural Network (MA-CNN) architecture for classifying mammogram images. The proposed model improves the accuracy by fusing the broader context of the information using multi-scale filters without negotiating the computational speed. The proposed system consists of two major phases: context feature extraction and classification.
- Tavakoli. N et al [15], introduced a novel approach to classify breast tissues as either normal or abnormal (indicating cancer) using deep learning techniques. The method primarily relies on a newly designed convolutional neural network (CNN) and a decision mechanism. Firstly, a preprocessing phase is performed, followed by feeding a block around each pixel into the trained CNN to determine its classification as normal or abnormal tissue. This process generates a binary map representing the suspicious areas. Subsequently, a thresholding technique is applied to a central block of the binary map as a decision mechanism to to assign labels
- Shen. L et al [16], introduced a novel computer-aided diagnosis system is for the early detection of breast cancer. The proposed method comprises five essential stages: noise reduction, suspicions breast tumor area extraction, mathematical morphology, feature extraction using a combination of discrete wavelet decomposition and GLCM, and classification based on a Deep Belief Network (DBN). To enhance the efficiency of the DBN, an optimized version of the sunflower optimization algorithm is employed.
- Li. H et al [17], proposed an enhanced DenseNet neural network model, called DenseNet II neural network model, to achieve efficient and precise classification of benign and malignant mammography images. The process begins with the mammography images normalization. Next, the DenseNet neural network model is improved by replacing the first convolutional layer with the Inception structure. The datasets utilized in the study were obtained from mammography images provided by the First Hospital of Shanxi Medical University, where mammary gland lesions were identified through full-field digital mammography (FFDM) examination.
- Benhassine. N et al [18], developed a novel Computer-Aided Diagnosis (CAD) system for classification of mammograms as either benign or malignant. The system encompasses three main steps. Firstly, the preprocessing stage involves filtering out noise, eliminating unwanted objects, and suppressing the pectoral muscle. To localize and extract the region of interest (ROI), the Seeded Region Growing (SRG) segmentation technique is applied within a triangular region containing the pectoral muscle. In the feature extraction step, the discrete wavelet transform (DWT) is applied to each ROI, and the most discriminative coefficients are selected using the discrimination power analysis (DPA) method. Finally, classification is performed using the support vector machine (SVM), artificial neural networks (ANN), random forest (RF), and Naive Bayes (NB) classifiers.
- Oyelade.O et al [19], proposed a novel convolutional neural network (CNN) model designed specifically for detecting architectural distortion. To enhance its performance, they employed a data augmentation technique and analyzed their impact on the performance of the proposed model. Furthermore, the model is adapted to handle images of both the right and left breasts, captured in MLO (Mediolateral Oblique) and CC (Cranio-Caudal) views. They also explored the performance of their model when using fixed size regions of interest (ROIs) and utilizing multi-size whole image inputs.
- Niu. J et al [20] proposed method using multi-scale image features for the classification of benign and malignant breast masses in mammograms. To accomplish this, they employed multi-scale residual networks and densely connected networks as backbone networks to extract and combine features from both global and local image patches. The two networks incorporated an attention module called Convolutional Block Attention Module (CBAM to help improving the network's ability to express and emphasize important features.
- Ayana. G et al [21] introduced a novel approach that involved incorporating a spatial pyramid pooling layer into an end-to-end multi-scale deep learning framework for classifying breast mass images. The model used the EfficientNetB2 as backbone. Table 1 includes a summary of the methods and results obtained in the aforementioned related work.

In this paper, we propose a scheme of our diagnostic aid system (DAS) for the classification of breast cancer in mammograms as normal or abnormal and then malignant or benign. The proposed system is divided into three main steps: Preprocessing, training model, and finally classification. In preprocessing we considered Grad CAM method which produces convincing qualitative results in cropping the regions of interest and saving time of preprocessing unlike methods evoked in related works.

3 Approach

In this section, we will explain the different steps of our approach illustrated in Figure 1. At the beginning a global scheme summarizing the procedure to follow is presented, starting with the collection of the dataset, passing by the different steps of pre-processing, then the data augmentation, training of the model with the different pre-trained networks, and finally the prediction of the class, Normal or Abnormal, then Malignant or Benign.

3.1 Dataset

The MIAS (Mammographic Image Analysis Society) dataset is a well-known and widely used dataset in tasks related to mammographic image analysis such as breast cancer detection, classification, and segmentation. MIAS consists of 322 digitized grayscale and 1024x1024 images acquired from 161 patients, with each patient contributing two images representing the left and right breast [22].

The dataset includes different types of abnormalities, such as masses, microcalcifications, and architectural distortions. Additionally, the final diagnosis for each case is also provided in the dataset. Some of the example images are presented in Figure 2.

3.2 Preprocessing

Digital mammograms are medical images that are difficult to interpret due to their complexity. Image preprocessing techniques play a crucial role in various aspects related to mammograms. These techniques are essential for unifying the orientation of the mammogram, eliminating artifacts, irrelevant areas and enhancing the overall image quality. By performing preprocessing steps before any image analysis task, it becomes possible to limit the influence of background characteristics and irrelevant areas in mammography for making any decision.

- **Application of Grad-cam and ROI cropping**

Grad-CAM consists of searching which parts of the image led a convolutional neural network to its final decision.

This method consists in producing heat maps representing the activation classes on the images received as input.

An activation class is associated with a specific output class. These classes will indicate the importance of each pixel in relation to the class concerned by increasing or decreasing the intensity of the pixel [23].

Grad-CAM uses the gradient information flowing into the last convolutional layer of the CNN to understand the importance of each neuron for making decisions. In order to obtain the class discriminative localization map for a particular class c, the method first computes the gradient of the score y^c (before softmax) with respect to the feature maps A^k: [24]

$$g_c(A^k) = \frac{\partial y^c}{\partial A^k}, \tag{1}$$

In which k is the channel index. Then, it averages the gradients as the neural importance weight α c k in each channel:

$$\alpha_k^c = \frac{1}{Z}\sum_i \sum_j \frac{\partial y^c}{\partial A_{i,j}^k}, \tag{2}$$

Table 1: Summary of related work

Authors	Target task	Approach	Preprocessing	Dataset	Performance metrics
Ting F. F et al [12]	Classification of mammogram images as malignant, benign or normal	Proposal ROIs - ROIs classification - interactive detection based lesion locator (IDBLL).	Decomposing mammogram images into patches	MIAS (Mammographic Image Analysis Society)	Accuracy : 90.50% Specificity : 90.71 % Sensitivity : 89.47% AUC: 0.901 ± 0.0314
Bakkouri. I et al [13]	Classification of breast tissues as normal or abnormal	Representative region proposals extraction using the determinant of the Hessian matrix – classification using multi-scale CNNs	-	Breast Cancer Digital Repository (BCDR)	Accuracy : 96.84% Sensitivity : 92.12% Specificity : 98.02% Precision : 92.15% F1-score : 92.12% AUC : 96.76%
Agnes. S. A et al [14]	Classification of mammogram images as malignant, benign or normal	Multiscale All Convolutional Neural Network (MA-CNN)	Artifacts and pectoral muscle removal	Mini-MIAS	Accuracy : 96.47% Sensitivity : 96% AUC : 0.99

Tavakoli. N et al [15]	Classification of breast tissues as normal or abnormal	Pixels level classification of the inputted suspicious region by a CNN, and Assigning a single label of 'normal' or 'abnormal' to each ROI by a decision mechanism	Breast region extraction, pectoral muscle suppression, mask creation, and contrast enhancement	MIAS (Mammographic Image Analysis Society)	Accuracy: 95% Sensitivity: 93.33% Specificity: 95.31 % AUC: 94.68%
Tavakoli. N et al [15]	Classification of breast tissues as normal or abnormal	Pixels level classification of the inputted suspicious region by a CNN, and Assigning a single label of 'normal' or 'abnormal' to each ROI by a decision mechanism	Breast region extraction, pectoral muscle suppression, mask creation, and contrast enhancement	MIAS (Mammographic Image Analysis Society)	Accuracy: 95% Sensitivity: 93.33% Specificity: 95.31 % AUC: 94.68%
Shen. L et al [16]	Classification of mammogram images into normal and abnormal	Combination of discrete wavelet decomposition and GLCM for feature extraction + classification based on Deep Belief Network (DBN) optimized by an improved version of the Sunflower optimization algorithm	Noise reduction, suspicions breast tumor area extraction, mathematical morphology	MIAS	Accuracy: 91.5% Specificity: 72.4% Sensitivity: 94.1%
Li. H et al [17]	Classification of mammogram images benign and malignant	DenseNet-II (the first convolutional layer was replaced with the Inception structure)	Image normalization	Dataset from the First Hospital of Shanxi Medical University	Accuracy :94.55% Specificity: 95.36% Sensitivity: 95.60 %
Benhassine. N et al [18]	Classification of mammogram images benign and malignant	Discrete wavelet transform (DWT) for feature extraction, and the most discriminative coefficients are selected using the discrimination power analysis (DPA). Classification using SVM, ANN, RF, and Naive Bayes.	Noise filtering, pectoral muscle suppression, region of interest (ROI) extraction	MIAS	Accuracy: 99.41 % Specificity: 99.37% Sensitivity: 99.50%
Oyelade.O et al [19]	Architectural distortion detection	CNN model from scratch	Noise suppression, contrast enhancement, e background area, labels, artefacts, and pectoral muscle removal, ROIs extraction	Combination of (MIAS, INbreast, DDSM) datasets	Accuracy: 93.75%
Niu. J et al [20]	Classification of benign and malignant breast masses in mammograms	Feature extraction and combination from both global and local image patches using CBAM ResNet and DenseNet	Noise suppression, artefacts, and image patches extraction	DDSM INbreast	Accuracy: 0.9626% ± 0.0110 Sensitivity: 0.9719% ± 0.0126 Specificity: 95.31 % AUC: 0.9576% ± 0.0064

Ayana. G et al [21]	Classification of benign and malignant breast masses in mammograms	Incorporating a spatial pyramid pooling layer into EfficientNetB2	Contrast enhancement	DDSM INbreast	Accuracy: 99.08% on INbreast and 99.99% on DDSM

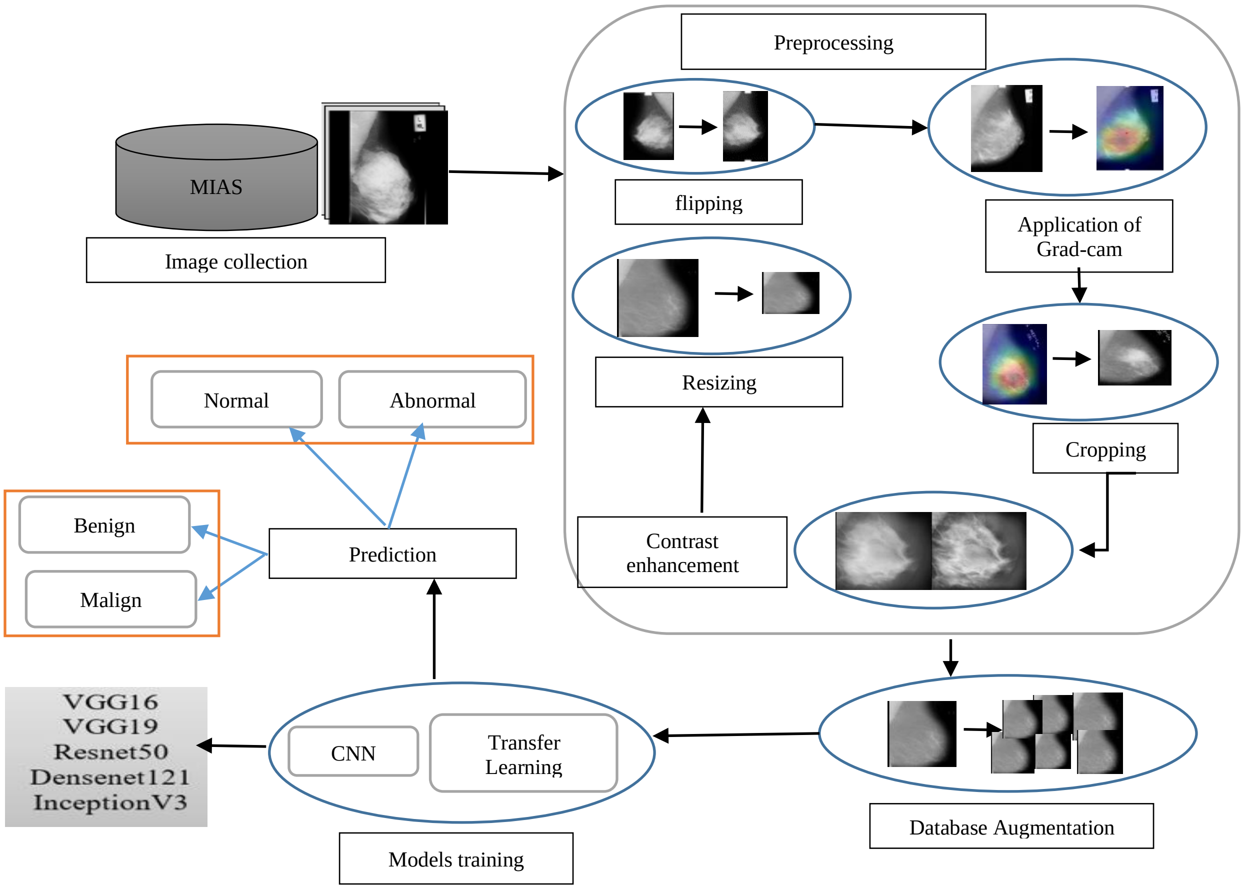

Figure1: Block diagram of the proposed

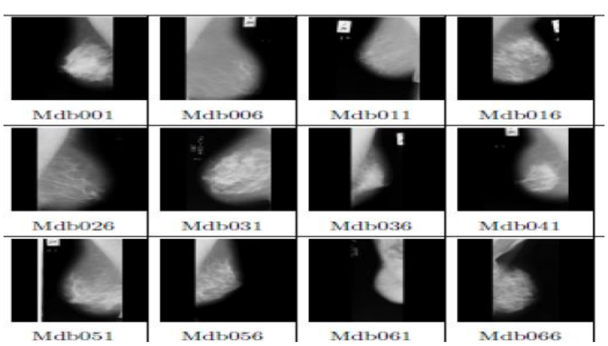

Figure 2: Example of histopathology images from MIAS dataset.

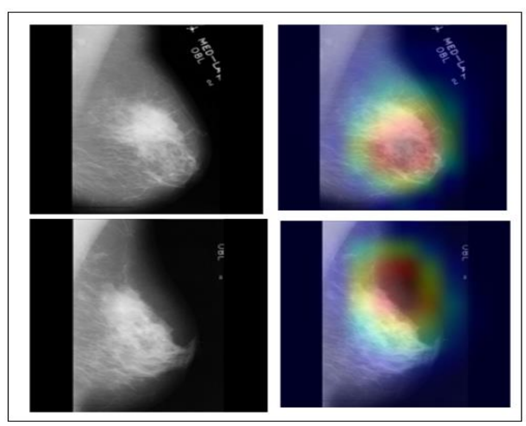

Figure 3: application of grad CAM on two Benign Class breast images.

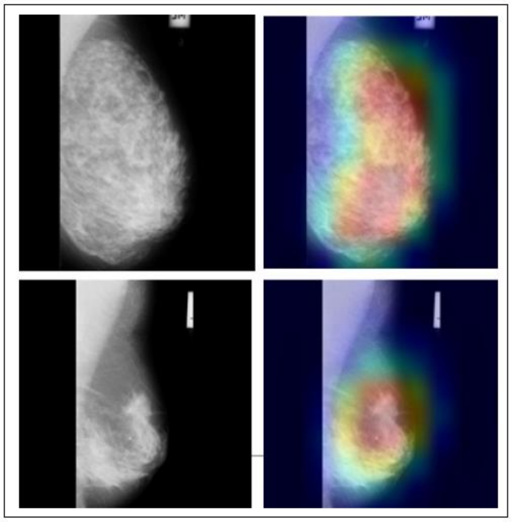

Figure 4: application of grad CAM on two malign class breast images.

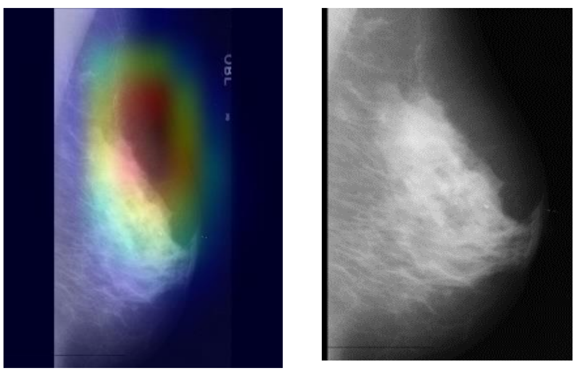

Figure 5: cropping of the original image from the grad-CAM images.

In which (i, j) is the spatial index and Z is the spatial resolution of the feature map. We call this weight as a grad weight. Finally, Grad-CAM is a weighted sum of feature maps, followed by a ReLU operator:

$$H^{c}_{\text{Grad-CAM}} = ReLU\left(\sum_{k} \alpha^{c}_{k} A^{k}\right). \tag{3}$$

First, we used Vgg16 pre-trained model (model) that maps the input image to the activations of the last convolutional layer as well as the output predictions, then, we compute the gradient of the top predicted class for our input image with respect to the activations of the last convolutional layer.

We return the gradient of the output neuron (top predicted or chosen) with regard to the output feature map of the last convolutional layer. Then we return a vector where each entry is the mean intensity of the gradient over a specific feature map channel, we multiply each channel in the feature map array by "how important this channel is" with regard to the top predicted class then sum all the channels to obtain the heatmap class activation. For visualization purpose, we will also normalize the heatmap between 0 & 1. [25]. Figures 3&4 shows the application of grad CAM.

Thanks to the thermal images generated by Grad-CAM, we can now proceed to the cropping of the parts that interest us. The cropping is applied for two types of images on which we will perform our tests later. The first type of image, with the same coordinates resulting from the thermal images, we crop the original images, and the second type of image is to crop directly the thermal image with the same coordinates of the heatmap. In the Grad-CAM crop, most part of the image that we are not interested in is eliminated, in most of the images passed to the Grad CAM, we see the removal of artifacts as well as a good part of the pectoral muscle, and the separation of the background. An example is shown in figure 5.

- **Contrast enhancement**

Contrast is the range or difference in luminance that distinguishes an object from other objects in the same image. Contrast enhancement is an image processing technique used to increase the contrast and visibility of the images in question. It improves the visual clarity and allows to differentiate the distinct elements of the image. Mammogram images have low contrast so the Contrast Limited Adaptive Histogram Equalization (CLAHE) method was applied to improve the contrast of the images [26]. Figure 6 shows an example image.

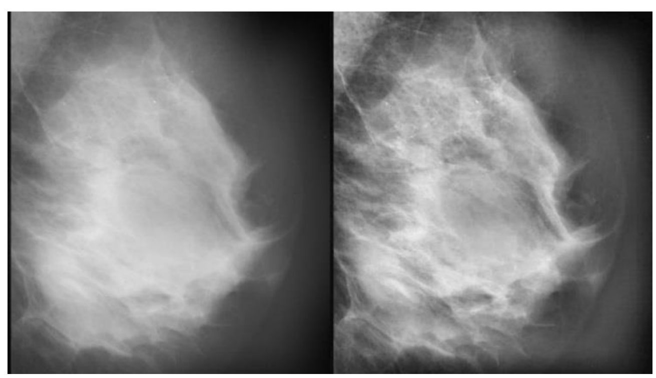

Figure 6: representative image of the contrast improvement of the cropped images.

- **Data augmentation**

The process of data augmentation involves applying a set of predefined operations to each image simple, generating new augmented dataset that are similar but slightly different from the original dataset. In this work, we based on geometric transformations which includes:

Translation: which is nothing more than moving images relative to the center, the translations used are: Shift right, Shift left, and Shift Medium.

Rotation: another strategy is to rotate images by a random number of degrees. In this way, the augmented images can better resemble the instances that were scanned at a small angle, we rotated from 0 to 360 degrees with a step size of 8 degrees and without including the 0 and 360 that represent the same image.

3.3 Model training

In order to perform the classification, we proceeded to transfer learning, we chose five pre-trained models, in order to be able to compare the results namely: VGG16, VGG19, Resnet50, Densenet121 and InceptionV3.

The aforementioned pre-trained models were adapted to handle the binary classification of the MIAS dataset.

In order to reduce overffiting, the dropout technique was used. A probability of 0.5 was used to select neurons that do not participate in training for Vgg16, Vgg19 and inceptinV3, and a probability of 0.2 was used for Resnet50 and Densenet121; thus, a fully connected network is reduced in training at each step [27].

We used also Batch normalization between layers to mainly improve the training speed with our model.

In normal abnormal classification, we used a single node for the last layer with the Sigmoid activation function. The value for the sigmoid function is between 0 and 1 only, the result can be predicted easily to 1 if the value is greater than 0.5 and 0 otherwise.

4 Results and discussions

4.1 Normal, abnormal classification

We consider in this part two different experiment with and without contrast enhancement with the same data augmentation using only rotations, with the number of epochs of 50 for the pretrained model VGG16. For the data augmentation, we went from 322 images to 14490 of which 11592 are used in learning, and 2898 are used for test. Clearly the contrast enhancement has a remarkable improvement of accuracy which passed from 89.26% to 93.68% the results are illustrated in the table 2.

Table 2: summary results of using Vgg16 with and without contrast enhancement.

Without contrast enhancement Data augmentation: rotation	**Precision %**	87.785
	Recall %	88.735
	F1-score %	88.22
	Accuracy %	89.26
With contrast enhancement Data augmentation: rotation	**Precision %**	92.93
	Recall %	93.32
	F1-score %	93.12
	Accuracy %	93.68

For the rest of experiments, we keep the contrast enhancement and we first consider the data augmentation of rotations only then the super augmentation where we consider rotations and translations. We summarize the different in results in the table3.

Table 3: summary results of different pretrained models using contrast enhancement.

Model	**Precision %**	**Recall %**	**F1-score %**	**Accuracy %**
Vgg16 Nb epoch = 50	92.93	93.32	93.12	93.68
Vgg19 Nb epoch = 60	94.885	94.955	93.555	95.34
Resnet50 Nb epoch = 60	93.555	95.905	94.555	95.10
Densenet121 Nb epoch = 60	91.525	90.295	90.825	91.40
InceptionV3 Nb epoch = 60	89.71	90.27	89.975	90.82

For the super augmentation where we consider rotations and translations. We went from 322 images to 57960 of which 46367 are for learning and 11593 for testing. The results are illustrated in the table 4.

Throughout the different tests performed for the normal and abnormal classification of mammography images, we found that for different models used, the Accuracy progresses and significantly improved when we have more images, in our case data augmentation method are translation and rotations where the number is four times greater than the case where the data augmentation used was only the rotations of images. In addition, some models have better prediction than others, indeed the best performances were obtained for the pre-trained network vgg19 with Accuracy of 98.54% then the resnet50 with Accuracy of 97.85% then the densenet121 with Accuracy of 97.79%, finally the VGG16 with Accuracy of 96.41%.

Figure 7 depicts also, for more illustration, a graphical comparison of accuracy according to the pre-trained models for the two data augmentation cases.

Table 4: summary results of different pretrained models with data super augmentation.

Model	Precision %	Recall %	F1-score %	Accuracy %
Vgg16 Nb epoch = 50	96.19	96.035	96.11	96.41
Vgg19 Nb epoch = 60	98.48	98.36	98.42	98.54
Resnet50 Nb epoch = 60	97.45	97.88	97.66	97.85
Densenet121 Nb epoch = 60	97.75	97.475	97.61	97.79
InceptionV3 Nb epoch = 60	92.93	93.32	93.12	93.68

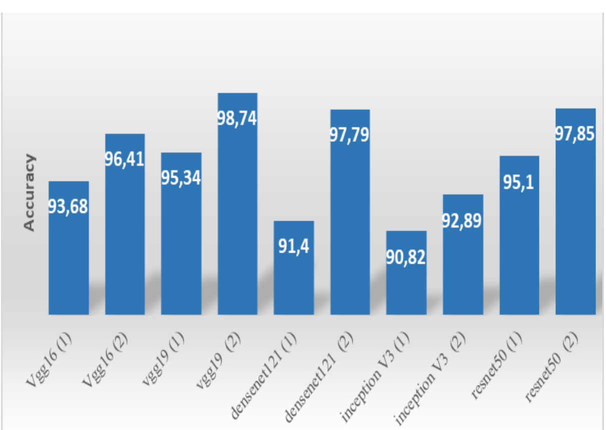

Figure 7: Graphical comparison of accuracy for the two data augmentation cases.

4.2 Benign, malign classification

Along this part, the Learning rate is estimated to 10-5, the Batch size is 64. The number of epochs varies between 40 and 80 and the considered optimizer is 'Adam'.

For the data augmentation using only rotations we went from 115 images to 3195 of which 2556 are used for training, and 639 are used for testing. For the super augmentation we used, rotations plus translations. We went from 115 images to 20700 of which 16560 are for training and 4140 for testing.

In this part, we found that in the different models used, the accuracy progressed and significantly improved in datasets where we have more examples, in our case the data augmentation methods performed is translation and rotations and the number is four times greater than the dataset where the data augmentation considered only the rotations. In addition, some models have a better prediction than others, indeed the best performance was obtained for the pre-trained network Resnet50 with an accuracy of 99.13% then the VGG19 with an accuracy of 98.93% then the VGG16 with an accuracy of 98.81%, finally the densenet121 with an accuracy of 98.52%.

We summarize the results in tables 5&6 for different models, and different data augmentation methods.

To further illustrate, Figure 8 shows the graphical representation of the accuracy according to the pre-trained models for both augmented and super-augmented datasets.

Table 5: summary results of different pretrained models using data augmentation with only rotation.

Model	Precision %	Recall %	F1-score %	Accuracy %
Vgg16 Nb epoch = 60	95.3	95.3	95.3	95.30
Vgg19 Nb epoch = 50	96.88	96.85	96.86	96.87
Resnet50 Nb epoch = 80	90.32	91.3	89.94	89.98
Densenet121 Nb epoch = 60	95.59	95.635	95.61	95.61
InceptionV3 Nb epoch = 40	96.63	96.85	96.70	96.71

Table 6: Summary results of different pretrained models using data augmentation with rotations and translations.

Model	Precision %	Recall %	F1-score %	Accuracy %
Vgg16 Nb epoch = 60	98.86	98.75	98.80	98.81
Vgg19 Nb epoch = 40	98.95	98.9	98.92	98.93
Resnet50 Nb epoch = 60	99.06	99.19	99.12	99.13
Densenet121 Nb epoch = 40	98.49	98.53	98.51	98.52
InceptionV3 Nb epoch = 40	95.24	96.09	95.55	95.62

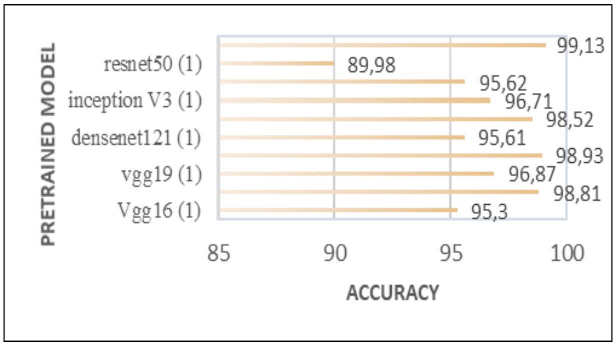

Figure 8: Graphical representation of accuracy in % according to the pre-trained models for both augmented and super augmented data sets.

4.3 Comparison with existing methods

To assess the performance of our method, we conducted a series of experiments and evaluations using MIAS dataset. We compare our method against some existing works that use the same dataset and deep learning techniques for mammogram images classification to demonstrate the advantages of our approach. The key aspects considered in this comparison is preprocessing steps and accuracy rate. As reported in table7, our method achieved the highest results with an accuracy of 99.13%. These results prove the performance of the proposed approach.

As shown in Table 7, Ting. F. F et al [12] achieved an accuracy of 90.50% for mammogram images classification as malignant, benign or normal using feature wise pre-processing (FWP) for region of interest (ROI) extraction and CNN for classification.

Regarding Agnes. S. A et al [14] they achieved an accuracy of 96.47%% using multiscale All Convolutional Neural Network (MA-CNN) architecture. For mammogram images preprocessing they removed noises and artifacts using the median filter and global thresholding followed by standard morphological operations respectively. The pectoral muscles are removed using single seeded region growing algorithm.

Tavakoli. N et al [15] obtained an accuracy of 95% for classification of breast tissues as normal or abnormal. The method primarily relies a newly designed convolutional neural network (CNN). Preprocessing phase includes four main steps: breast region extraction employing Otsu's thresholding, pectoral muscle suppression, mask creation, and contrast enhancement using CLAHE method.

Shen Lizhen et al [16] reached an accuracy of 91.5% for breast cancer mammogram images classification into normal and abnormal based on noise reduction, suspicions breast tumor area extraction, mathematical morphology for preprocessing and using a combination of discrete wavelet decomposition and GLCM for feature extraction, and the classification was based on a Deep Belief Network (DBN) using the sunflower optimization algorithm.

Table7: Comparison with existing methods.

Authors	Performance
Ting. F. F et al [12]	Accuracy: 90.50%
Agnes. S. A et al [14]	Accuracy: 96.47%
Tavakoli. N et al [15]	Accuracy: 95%
Shen Lizhen et al [16]	Accuracy: 91.5%
Our method	Accuracy: 99.13%

5 Conclusion

In this paper, we proposed a new system (CAD) to classify mammograms as normal or abnormal and then classify the abnormal ones as benign or malignant. We have initially performed preprocessing necessary to ensure optimal performance of our system. A crop applied after performing the Grad-CAM that provided us heat maps designating the possible region of interests, from which the crop is based, which saved us time compared to the preprocessing phase, comparing to the previeus works cited in the state of the art.

As the data are very small and unbalanced, we proceeded to a data augmentation based on translations and rotations in order to improve the classification performances. Several pretrained models are exploited and compared thereafter namely: the Vgg16, Vgg19, Resnet50, Densenet121 and inceptioinV3, while performing Dropout and batchnormalization for better performances. For Normal and Abnormal classification, we obtained, an accuracy of 96.41%., 98.54%, 97.58% and 97.79% for Vgg16, Vgg19, Resnet50 and Densenet121 respectively. As for the Begnine and Malignant classification, we obtained an accuracy of 99.13%, 98.93%, 98.81%, and 98.52% for Resnet50, VGG19, VGG16 and densenet121 respectively and for the InceptionV3 model the accuracy was was 90% for the Normal and abnormal classification, and around 96% for the Begnine and Malignant classification.

References

[1] Globocan (2018) site web:
[2] https://gco.iarc.fr/today/data/factsheets/populations/12-algeria-fact-sheets.pdf.
[3] Al-Jammali, K. (2023). Prediction of Heart Diseases Using Data Mining Algorithms. Informatica, 47(5).
[4] https://doi.org/10.31449/inf.v47i5.4467
[5] Ullah, Z., Odeh, A., Khattak, I., & Hasan, M. A. (2023). Enhancement of Pre-Trained Deep Learning Models to Improve Brain Tumor Classification. Informatica. Vol. 47, N° 6 165-172.
[6] DOI:10.31449/inf.v47i6.4645
[7] Puliti, D., Duffy, S. W., Miccinesi, G., De Koning, H., Lynge, E., Zappa, M., & Paci, E. (2012). Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. Journal of medical screening, 19(1_suppl), 42-56.
[8] https://doi.org/10.1258/jms.2012.012082.
[9] Elmore, J. G., Armstrong, K., Lehman, C. D., & Fletcher, S. W. (2005). Screening for breast cancer. Jama, 293(10), 1245-1256. doi:10.1001/jama.293.10.1245.
[10] Beura, S., Majhi, B., & Dash, R. (2015). Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer. Neurocomputing, 154, 1-14. https://doi.org/10.1016/j.neucom.2014.12.032.
[11] Buciu, I., & Gacsadi, A. (2011). Directional features for automatic tumor classification of mammogram images. Biomedical Signal Processing and Control, 6(4), 370-378.
[12] https://doi.org/10.1016/j.bspc.2010.10.003.
[13] Verma, B. (2008). Novel network architecture and learning algorithm for the classification of mass abnormalities in digitized mammograms. Artificial Intelligence in Medicine, 42(1), 67-79.
[14] https://doi.org/10.1016/j.artmed.2007.09.003.
[15] Rodríguez-Ruiz, A., Krupinski, E., Mordang, J. J., Schilling, K., Heywang-Köbrunner, S. H., Sechopoulos, I., & Mann, R. M. (2019). Detection of breast cancer with mammography: effect of an artificial intelligence support system. Radiology, 290(2), 305-314.
[16] https://doi.org/10.1148/radiol.2018181371.
[17] Sánchez-Ferrero, G. V., & Arribas, J. I. (2007). A statistical-genetic algorithm to select the most significant features in mammograms. In Computer Analysis of Images and Patterns: 12th International Conference, CAIP 2007, Vienna, Austria, August 27-29, 2007. Proceedings 12 (pp. 189-196). Springer Berlin Heidelberg.
[18] https://doi.org/10.1007/978-3-540-74272-2_24.
[19] Brunelle, F., & Brunelle, P. (2019). Intelligence artificielle et imagerie médicale: Définition, état des lieux et perspectives. Bulletin de l'Académie Nationale de Médecine, 203(8-9), 683-687.
[20] https://doi.org/10.1016/j.banm.2019.06.016.
[21] Ting, F. F., Tan, Y. J., & Sim, K. S. (2019). Convolutional neural network improvement for breast cancer classification. Expert Systems with Applications, 120, 103-115.
[22] https://doi.org/10.1016/j.eswa.2018.11.008.
[23] Bakkouri, I., & Afdel, K. (2019). Multi-scale CNN based on region proposals for efficient breast abnormality recognition. Multimedia Tools and Applications, 78, 12939-12960.
[24] https://doi.org/10.1007/s11042-018-6267-z.
[25] Agnes, S. A., Anitha, J., Pandian, S. I. A., & Peter, J. D. (2020). Classification of mammogram images using multiscale all convolutional neural network (MA-CNN). Journal of medical systems, 44, 1-9. https://doi.org/10.1007/s10916-019-1494-z.
[26] Tavakoli, N., Karimi, M., Norouzi, A., Karimi, N., Samavi, S., & Soroushmehr, S. R. (2019). Detection of abnormalities in mammograms using deep features. Journal of Ambient Intelligence and Humanized Computing, 1-13. https://doi.org/10.1007/s12652-019-01639-x.
[27] Shen, L., He, M., Shen, N., Yousefi, N., Wang, C., & Liu, G. (2020). Optimal breast tumor diagnosis using discrete wavelet transform and deep belief network based on improved sunflower optimization method. Biomedical Signal Processing and Control, 60, 101953. https://doi.org/10.1016/j.bspc.2020.101953.
[28] Li, H., Zhuang, S., Li, D. A., Zhao, J., & Ma, Y. (2019). Benign and malignant classification of mammogram images based on deep learning. Biomedical Signal Processing and Control, 51, 347-354.
[29] https://doi.org/10.1016/j.bspc.2019.02.017.
[30] Benhassine, N. E., Boukaache, A., & Boudjehem,

D. (2020). A new cad system for breast cancer classification using discrimination power analysis Of Wavelet's coefficients and support vector machine. Journal of Mechanics in Medicine and Biology, 20(06), 2050036. https://doi.org/10.1142/S0219519420500360.

[31] Oyelade, O. N., & Ezugwu, A. E. (2021). A deep learning model using data augmentation for detection of architectural distortion in whole and patches of images. Biomedical Signal Processing and Control, 65, 102366. https://doi.org/10.1016/j.bspc.2020.102366.

[32] Niu, J., Li, H., Zhang, C., & Li, D. (2021). Multi-scale attention-based convolutional neural network for classification of breast masses in mammograms. Medical Physics, 48(7), 3878-3892.

[33] https://doi.org/10.1002/mp.14942.

[34] Ayana, G., Park, J., & Choe, S. W. (2022). Spatial Pyramid Pooling based end-to-end deep learning for mammogram breast masses classification. Proceedings of the Korean Society of Communications and Communications Society 893-895.

[35] J Suckling et al (1994): The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Med- ica. International Congress Series 1069 pp375-378. http://peipa.essex.ac.uk/info/mias.html.

[36] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE international conference on computer vision (pp. 618-626).

[37] Chen, L., Chen, J., Hajimirsadeghi, H., & Mori, G. (2020). Adapting grad-cam for embedding networks. In proceedings of the IEEE/CVF winter conference on applications of computer vision (pp. 2794-2803).

[38] Chattopadhay, A., Sarkar, A., Howlader, P., & Balasubramanian, V. N. (2018, March). Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. In 2018 IEEE winter conference on applications of computer vision (WACV) (pp. 839-847). IEEE.

[39] Q.-C. Tian. (2018). Color Correction and Contrast Enhancement for Natural Images and Videos. PhD thesis, PSL Research University.

[40] Priya, T. Sathya. (2021). Resnet based feature extraction with decision tree classifier for classificaton of mammogram images. Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12.2: 1147-1153.https://turcomat.org/index.php/turkbilmat/article/view/1136.