Breast Mass Regions Classification from Mammograms using Convolutional Neural Networks and Transfer Learning.

Yuliana Jiménez-Gaona[a,b,d*], María José Rodríguez-Álvarez [b], Diana Carrión-Figueroa [c], Darwin Castillo-Malla[a,b,d], and Vasudevan lakshminarayanan [d,e]

[a]Departamento de Química y Ciencias Exactas, Universidad Técnica Particular de Loja, San Cayetano Alto s/n CP1101608, Loja, Ecuador; ydjimenez@utpl.edu.ec, https://orcid.org/0000-0001-7155-5546 dpcastillo@utpl.edu.ec, https://orcid.org/0000-0002-1800-1189

[b]Instituto de Instrumentación para la Imagen Molecular i3M, Universitat Politècnica de València (UPV)-Consejo Superior de Investigaciones Científicas (CSIC), E-46022 Valencia, Spain; mjrodri@i3m.upv.es, https://orcid.org/0000-0001-8333-8792

[c]Hospital Carlos Andrade Marín, IESS, Ayacucho N19-63 y Av. 18 de Septiembre, Quito, Ecuador, dcf1885@hotmail.com, https://orcid.org/0000-0002-7842-4714

[d]Theoretical and Experimental Epistemology Lab, School of Optometry and Vision Science, University of Waterloo, Waterloo, ON N2L3G1, Canada, vengulak@uwaterloo.ca, https://orcid.org/0000-0002-3473-1245

[e]Department of Systems Design Engineering, Physics, and Electrical and Computer Engineering, Unive rsity of Waterloo, Waterloo, ON N2L3G1, Canada

*corresponding author, ydjimenez@utpl.edu.ec , vengulak@uwaterloo.ca

Breast Mass Regions Classification from Mammograms using Convolutional Neural Networks and Transfer Learning.

Abstract: Breast cancer constitutes a prominent public health concern, with the early detection being of paramount importance for successful treatment outcomes. Digital mammography has solidified its position as a standard diagnostic tool, with the accuracy of detection closely tied to the quality of the images it produces. In this study, we introduce a novel approach aimed at enhancing the quality of digital mammography images through pre-processing techniques, with the goal of improving breast cancer detection accuracy. The primary objective of this research is to enhance image resolution, thus leading to more precise breast tissue segmentation and subsequent classification utilizing convolutional neural networks (CNNs).Our focus was directed towards three widely recognized public mammography databases: CBIS-DDSM, Mini-MIAS, and Inbreast for pre-processing data. This step is a pivotal part of the clinical decision-making process, intended to refine breast diagnosis. Our statistical findings revealed that the EDSR method (PSNR=39.05 dB /SSIM=0.90) consistently outperformed the visual quality of images when compared to SR-RDN (PSNR=32.68 dB/SSIM=0.82). Similarly, UNet demonstrated superior performance over SegNet, boasting an average Intersection over Union (IoU) of 0.862, an average Dice coefficient of 0.991, and an accuracy rate of 0.947 in Region of Interest (RoI) segmentation results. In conclusion, the ResNet model contributed to enhanced accuracy when compared to conventional machine learning algorithms, particularly when utilizing improved RoI data augmentation through affine transformation. However, it did not surpass state-of-the-art deep CNN-based classifiers, achieving an accuracy rate of 75%.

Key words: Breast cancer, classification, convolutional neural network, mammography images, segmentation, super resolution.

1. Introduction

The early detection of breast lesions remains a significant challenge in the field of medical research [1]. Various screening methods [2,3] and less invasive approaches to breast cancer detection [4-8], including x-ray radiographic techniques, have been developed to address this issue. However, digital mammography (DM) stands out as a superior diagnostic modality for the early detection of breast lesions [9]. DM offers precise control and data acquisition while minimizing radiation exposure to patients, making it a critical tool in the fight against breast cancer [10].

Deep learning-based computer-assisted diagnostic (CAD) systems have emerged as a promising technology for medical image processing [11-13], playing a significant role in aiding radiologists in both screening and acting as a second reader to improve diagnostic accuracy. One significant challenge with deep learning (DL) training models is their reliance on extensive datasets for training. However, the limited dataset sizes are often due to privacy and data protection concerns, among other reasons [14-16].

Another critical aspect is that the accuracy of lesion detection heavily depends on image quality. DM images frequently exhibit various types of noise, including Salt and Pepper, Gaussian, Speckle, and Poisson noise [17]. These issues often stem from factors like image transfer, blurring, compression, or general image degradation, which lead to the production of low-resolution (LR) images. As a result, image super-resolution becomes a pivotal element in computer vision.

Multiple super-resolution techniques have been proposed to enhance the quality of medical images, thus improving the accuracy of segmentation and classification processes. These processes are crucial in the accurate diagnosis of cancer. Convolutional neural networks (CNNs) have been adapted for enhancing image resolution, segmentation, and classification tasks [17-22]. These techniques have consistently demonstrated exceptional performance in image reconstruction, employing single image super-resolution (SISR) [23,24] and Multi-Image Super-Resolution (MISR) algorithms, with SISR being widely adopted due to its remarkable efficiency [25,26].

1.1 Related work-state of the art

1.1.1 Data Augmentation

Original image data augmentation based on basic transformations including: spatial translation, rotation, horizontal flipping, random cropping [11] and oversampling. Other augmentation methods are also available and include [27]: geometric transformations, colour space transformations, kernel filters, image blending, random erasure, feature space augmentation, adversarial training, GAN-based augmentation [28-30], neural style transfer, and meta-learning schemes. These include spatial translation, rotation, horizontal flipping, random cropping [11] and oversampling.

Other augmentation strategies extend to more advanced methods [27], including geometric transformations, colour space transformations, kernel filters, image blending, random erasure, feature space augmentation, adversarial training, GAN-based A noteworthy observation comes from Yu et al. [15], who have substantiated that deep convolutional neural networks (CNNs) can experience substantial enhancements in performance when trained on augmented data as opposed to non-augmented data.

However, it is essential to bear in mind the insights provided by Yadav et al. [31], who have explored the impacts of both simple and complex data augmentation techniques. Their findings suggest that highly intricate transformations may not consistently outperform simpler ones, and in some cases, overly complex augmentations may introduce additional noise into the feature set, potentially detrimental to the learning process. augmentation [28-30], neural style transfer, and meta-learning schemes.

Hence, a judicious balance between diversity and noise in training data is recommended when selecting data augmentation methods.

1.1.2 Single Image Super Resolution

The root causes of this degradation are multifaceted, originating from the real-world clinical settings where medical imaging data is acquired. Factors like equipment conditions, patient movements, and technical constraints influence to compromise image quality. Therefore, the necessity arises to bolster the resolution of these images before they undergo the rigors of segmentation and classification.

Image SR emerges as the pivotal technique, elevating images from a low-resolution (LR) state to high resolution (HR) [32,33]. This process takes center stage in ameliorating the screening process, especially when addressing challenges like macrocalcifications or dense breast tissue. The enhancement profoundly impacts the precision of subsequent classification and segmentation processes.

SISR algorithms are categorize into four distinct types: 1) prediction models, 2) edge-based methods, 3) image statistics and 4) example-based or patch-based [18,33].

Traditional SR techniques encompass nearest-neighbor interpolation, such as bilinear interpolation, bicubic interpolation, and learning-based methods. Despite their simplicity and efficiency, they often grapple with reduced accuracy [36].

To surmount these limitations, Convolutional Neural Networks (CNNs) have been harnessed to generate HR images, using techniques like Convolutional SR-CNN [37-40] and Generative Adversarial Networks (GAN) SR-GAN [41-44]. A case in point is the model introduced by Jiang et al. [41], known as "TSGAN," which combines texture loss and encourages local information matching with a gradient penalty. This model achieved commendable metrics with an average PSNR of 27.99 dB and an SSIM of 0.778, metrics used to assess image quality and signal reconstruction.

An array of CNN and GAN-based methods has been devised, encompassing Multi-scale deep super-resolution systems (MDSR), Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN)[45], Residual Dense Block (RDB) [46], Efficient sub-pixel convolutional neural network (ESPCN) [47], Very Deep Network for SR (VDSR)[48], SR-ResNet, Sparse Coding-based Network (SCN)[49], Deep Recursive Convolutional Network (DRCN)[50], and Deep Recursive Residual Network (DRRN) [51]. Nevertheless, these methods at times fail to exploit the full spectrum of information residing in each convolutional layer and the hierarchical features crucial for reconstruction, thereby limiting their architectural optimality [45].

Current research endeavors revolve around the adaptation of additional CNN techniques for super-resolution and segmentation [38,52-58]. For instance, Tong et al. [59] devised a dense skip connection to circumvent the vanishing gradient problem plaguing very deep networks. Abbass et al. [46] and Zhang et al. [37] introduced a Residual Dense Network (RDN) grounded in the DenseNet architecture, characterized by a higher growth rate and the adept utilization of all hierarchical features from the LR image, yielding a marked enhancement in overall performance.

Worth noting is the application of SRCNN to mammography images, showcasing its superiority over conventional interpolation methods when enhancing digital mammography images of dense breasts. Dong et al. [38,39] propounded an SRCNN image reconstruction technique based on end-to-end (E2E) mapping, eclipsing traditional methods like sparse coding, kernel regression, and random forest [60,61]. Evidently, conventional SR methods, reliant on mapping functions from dictionaries, pale in comparison to modern Dense Neural Networks (DNN) employing E2E mapping approaches [62].

Additionally, the deployment of the Enhanced Deep Residual Network (EDSR), predicated on multiple ResNet architectures, has garnered attention [32]. While experimentation involving different scaling factors and optimizers yielded no single superior optimization technique [62], the caveat is that an increased number of layers may engender a surge in parameters, potentially bottlenecking image detail. Lim et al. [45] introduced single and multi-scale SR networks premised on SRResNet with a deeper residual design [48]. Residual learning techniques have manifestly led to improved performance by eliminating superfluous modules compared to antecedent methods. Their results showcase the achievement of higher PSNR values, specifically 1.57 dB for SRResNet and 2.14 dB for EDSR, respectively.

1.1.2.1 Image Quality Assessment

When it comes to evaluating the quality of super-resolution images, a range of established metrics is commonly employed. These metrics include the Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Metric (SSIM), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Signal-to-Noise Ratio (SNR), Multi-Scale Structural Similarity (MS-SSIM), Task-Specific Similarity Assessment (TSSA), and Mean Opinion Score (MOS) [63].

However, within the domain of medical imaging super-resolution, the predominant metrics of choice are PSNR and SSIM. These metrics take center stage when quantifying the quality of the generated image in comparison to the original image [64]. Figure 1 illustrates the scale values for PSNR and SSIM.

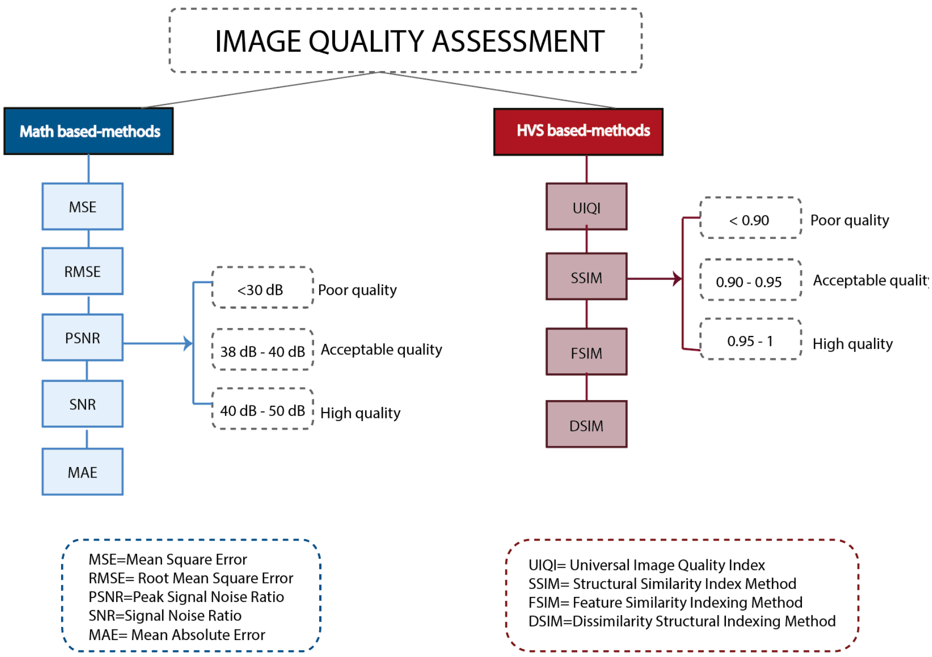


Fig. 1. Super resolution quality metrics, divided into (i) Math based methods and (ii) Human visual system-based methods [65-66].

For medical applications, the significance of PSNR and SSIM cannot be overstated. These metrics provide a valuable quantitative means of assessing the degree to which the enhanced images faithfully represent the original data. In the context of medical diagnostics, this fidelity is of utmost importance.

1.1.3 Segmentation

Segmentation is a critical task in medical image analysis, involving the separation of the region of interest (ROI) from the background in an image. Accurate tumour segmentation in medical images is particularly challenging due to the presence of various image artefacts and complexities. To address these challenges, researchers [67-70] have increasingly turned to deep learning (DL) methods, with a particular focus on CNNs. Various network architectures such as Visual geometric group (VGG-16) [71], ResNet [72], UNet [9,73-74], SegNet [75], ERFNet [76], have been applied to image segmentation.

1.1.4 Feature extraction and classification

Deep learning has established itself as a dominant approach for medical image classification. DL models have the capability to automatically extract relevant image features based on the shape of the ROI. Once these features are extracted and selected, they are utilized as input for a classifier, enabling the categorization of ROI samples into malignant or benign classes.

Several pre-trained networks have demonstrated high accuracy in medical image classification, including VGGNet, ResNet, DenseNet, and Inception [78-81]. Each of these CNN architectures has made significant contributions to the field of deep learning for computer vision tasks. The choice of architecture depends on the specific problem and involves trade-offs between factors such as model size, computational efficiency, and accuracy.

For example, DenseNet incorporates shorter skip connections between layers in a feed-forward architecture, resulting in enhanced accuracy, reduced susceptibility to overfitting, and efficient training through a cross-layer connection structure. VGGNet follows the classic CNN network structure, comprising a stack of convolutional, max-pooling, activation layers, and fully connected classification layers. ResNet offers flexibility with its fundamental shortcut connections that are task-dependent. In contrast, Inception networks employ convolution kernels of various sizes and pooling operations within a single layer [81].

Taking these factors into account, the proposed method aims to enhance the resolution of breast images, ultimately improving the accuracy of segmentation and classification. This approach leverages convolutional neural networks and transfer learning models.

Specifically, two novel CNN-based algorithms, EDSR (Enhanced Deep Residual Network) and RDN (Residual Dense Network), are introduced to address super-resolution challenges. For image segmentation, two primary pre-trained models, UNet and SegNet, are employed, while a CNN model (ResNet50) is used for image classification. The work is organized into four main sections: "Related Work," "Materials and Methods," "Results and Discussion," and "Conclusion and Future Work."

2.Methodology

2.1 Datasets

When evaluating computer-aided diagnosis (CAD) systems for breast cancer in mammography, several challenges arise. These include the absence of a standardized evaluation dataset and the necessity to adhere to ethical, regulatory, patient privacy, and data security considerations. Consequently, many CAD systems are assessed using private datasets or unspecified subsets of public databases. For this study, we harnessed three open-access datasets:

(i) CBIS-DDSM (Curated Breast Imaging Subset –Digital Database for Screening Mammography) [82]. This database consists of 2620 cases, each containing two different views: mediolateral oblique (MLO) and craniocaudal (CC). The images are stored in DICOM format and dimensions of approximately 3784 x 5912 pixels.

(ii) mini-MIAS (Mammographic Image Analysis Society) [83,84]. This dataset offers 322 MLO mammograms from 161 patients. The images have a resolution of 1024x1024 pixels and are categorized into 208 normal, 63 benign, and 51 malignant images.

(iii) Inbreast [85], this dataset comprises a total of 115 cases, with 90 of them having two views (MLO and CC). The image matrix size varies, featuring dimensions of 3328 x 4084 or 256 x 3328 pixels. Images in this dataset are stored in the DICOM format, see table 1.

The datasets were selected based on predefined inclusion criteria that encompass demographic characteristics (age ≥ 40 years and female gender) and clinical characteristics. The selected datasets adhere to specific criteria, including normal/cancer/benign cases with verified pathology information, breast density, and abnormality descriptions. Additionally, the datasets include images with two different views, MLO and CC, while excluding normal cases.

The research methodology for breast lesion classification using deep networks is structured into five key steps, as visualized in Figure 2: 1) Manual mask RoIs extraction, 2) RoIs cropping and data augmentation. 3) Super-resolution using EDSR and SR-RDSN algorithms. 4) RoIs segmentation using UNet and SegNet algorithms. 5) RoIs classification using ResNet-50 and finally, 6) Image quality evaluation using statistical metrics.

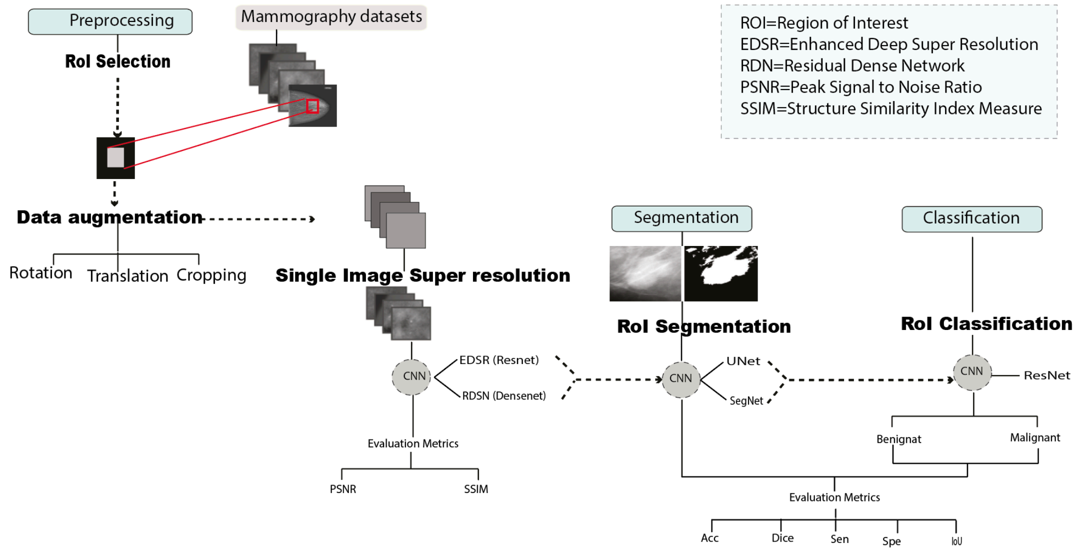


Fig. 2. The diagram describes the flowchart to breast lesion classification.

This comprehensive methodology encompasses data sources, image processing steps, and deep learning techniques employed for breast lesion classification and image quality enhancement. It provides a systematic approach to address the complexities of breast cancer diagnosis in mammography.

2.1.1 Manual Selection of RoI masks

The initial phase of our methodology involved the meticulous manual selection of a total of 784 binary Regions of Interest (RoIs) and their corresponding binary masks. This selection process is visually demonstrated in Figures 3a and 3b. The manual selection was carried out with the assistance of ImageJ software, which is accessible at [https://imagej.net/ij/docs/intro.html](see table 1).

Table 1. The distribution of benign and malignant cases per dataset.

Database	Benignant	Malignant	Total
CBIS-DDSM [82]	305	318	623
Mini-MIAS [83,84]	17	25	42
Inbreast]	70	49	119
Total	392	392	784

Our dataset encompasses 392 benign and 392 malignant images. To streamline the computational performance and facilitate the subsequent training of network models, all images were consistently resized to a dimension of 128 x 128 pixels.

This methodical data preparation is of paramount importance, as it forms the foundation for the subsequent phases in our methodology. It ensures that the network models can effectively learn and extract features from the RoIs, thereby contributing to the overall efficiency of the breast lesion classification system.

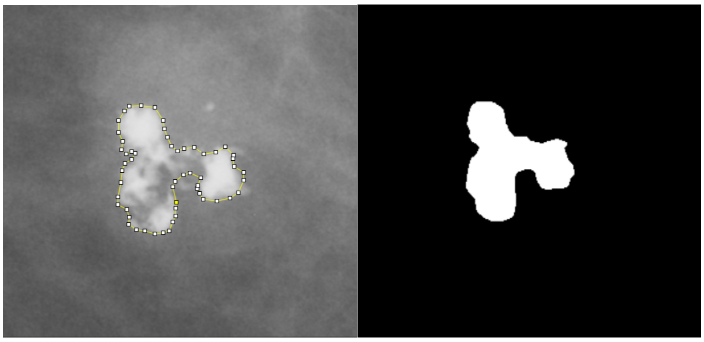


Fig. 3. (a)Manually RoI selection and (b) Binary mask.

Figures 3a and 3b provide a visual representation of the diligently selected RoIs and their corresponding binary masks, underscoring the crucial role of this data curation process in our methodology.

2.2 Data augmentation

Traditional data augmentation techniques are widely adopted in medical image analysis to mitigate the issue of limited data for training deep learning models. These techniques involve applying various transformations to the existing data, enhancing the model's generalization capabilities, and reducing the risk of overfitting.

In our study, the original dataset images underwent expansion through the application of basic geometric transformation operations, including: Blurring (Blur_1.5); Flipped (fliph, flipv); Translation (trans_20_20), rotation (rot_90, rot_180) and scaling [26] to generate new RoIs images from the selected databases. This data augmentation process yielded an additional 4704 RoI images. When combined with the original set of 784 RoIs, the dataset's size was extended to a total of 5486 RoIs.

2.3 Cross-validation

Cross-validation is a statistical technique employed to effectively partition the augmented dataset into subsets for model evaluation. In our case, the dataset was divided into three subsets: a training set, a validation set, and a testing set. The proportions for this division were determined as follows: Training Set: 4380 images (80%), Testing Set: 546 images (10%) and Validation Set: 560 images (10%).

This random yet systematic division of the dataset, as outlined in Table 2, ensures robust evaluation of the CNN models. Cross-validation is a key component in validating the performance and reliability of the breast lesion classification system.

Table 2. Data split into three sets: training, validation, and test.

Datasets	**Benignant**	**Malignant**	**Total**
Training	2190	2190	4380

Validation	280	280	560
Testing	273	273	546
Total	2743	2743	5486

Table 2 offers a concise summary of the dataset division based on the cross-validation technique, underscoring its integral role in our methodology.

2.4 Image Quality Assessment Using PSNR and SSIM

To assess the quality of our processed images, we employed well-known metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Metric (SSIM). These metrics require a reference image (ground truth) for comparison. In our case, we calculated these indexes by comparing the images generated by our model with high-resolution images (ground truth) from the CBIS-DDSM breast images. Importantly, the CBIS-DDSM dataset used for this evaluation was distinct from the data used for training the model, ensuring that the assessment was conducted on unseen data. The results of PSNR and SSIM are presented in Table 4.

2.5 Image Super Resolution Using Transfer Learning

In our study, we employed transfer learning to tackle the challenge of image Super-Resolution. We utilized two distinct models for this task: EDSR (Enhanced Deep Super-Resolution) based on the ResNet architecture and SR-RDN (Super-Resolution Residual Dense Network) based on the DenseNet architecture.

The training process involved the careful selection and optimization of hyperparameters, which play a vital role in enhancing the accuracy of the Convolutional Neural Networks

(CNNs) [86]. The optimization was conducted through a systematic exploration of different hyperparameter combinations, with Python's GridSearchCV (ParameterGrid) using scikit-learn. This approach enabled us to identify the combination of hyperparameters that minimized the margin of error, resulting in the most effective models.

Key hyperparameters details are described below:

EDSR hyperparameters: Scaling factor: 0.1, number of epochs: 20, Loss function:L1, number of blocks:50, optimization algorithm: ADAM/SGD/RMsProp, ResBlocks: 32, Number of filters:256, Upsampling factor x3.

RDN hyperparameters: Convolutional layer size: 3 × 3, Kernel size for local and global feature fusion: 1 × 1. Kernel size for convolutional layer: 3 × 3 with zero-padding, Local and global feature fusion layers: 64 filters with a Kernel size of 1x1, Upsampling factor: x4, Number of epochs: 30, Number of blocks: 16, Number of layers: 8, Batch size: 50, Patch size: 10, Activation function: ReLU.

2.6 Segmentation

2.6.1 UNet architecture

The UNet architecture is a fundamental component of our segmentation process. This model involves applying a series of convolutional operations to the input image, effectively compressing information, and detecting essential features. Subsequently, a new image is generated using the learned features acquired during the contraction process.

The hyperparameter details of the UNet architecture are outlined in Table 3, and you can also refer to Figure 4 for a visual representation of this architecture.

Tabla 3. Hyperparameters for U-net and SegNet training architecture.

Hiperparameter	Unet/ Segnet
Number of epochs	40
Batch size	4
Steps	125
Optimizer	Adadelta
Learning rate	0.001
Loss function	Binary-crossentropy
Activation function	ReLu, Sigmoid

Table 3 presents a comprehensive overview of the hyperparameters associated with the UNet model, while Figure 4 provides a graphical representation of the architecture's structure. This UNet architecture plays a crucial role in the segmentation of breast lesions, contributing to the accuracy of our analysis.

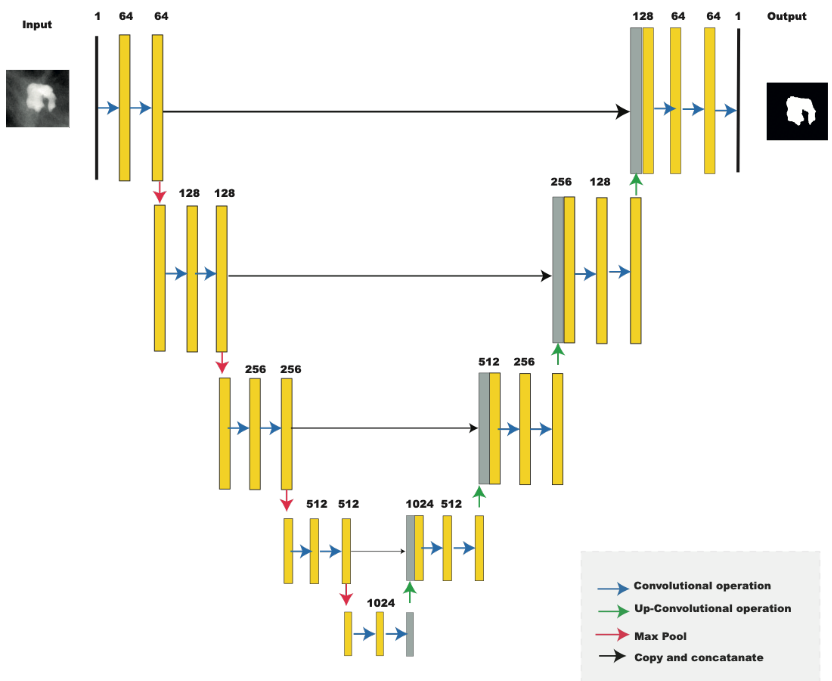


Fig. 4. The U-net architecture consists of an expansive path on the right side (upsampling) and a contracting path on the left side (downsampling). The contracting path follows the typical CNN architecture, which each yellow box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The grey boxes represent duplicated feature maps. The arrows denote the different operations.

2.5.2 SegNet architecture

SegNet consists of an encoder-decoder network (see Figure 5) followed by a pixel-wise classification layer. It lacks fully connected layers, relying solely on convolutional layers. The decoder upsamples its input using pooling indices, while the encoder produces a feature map, and subsequently performs convolution to densify the feature map. The final decoder output feature maps are then fed to a Softmax classifier for pixel-wise classification.

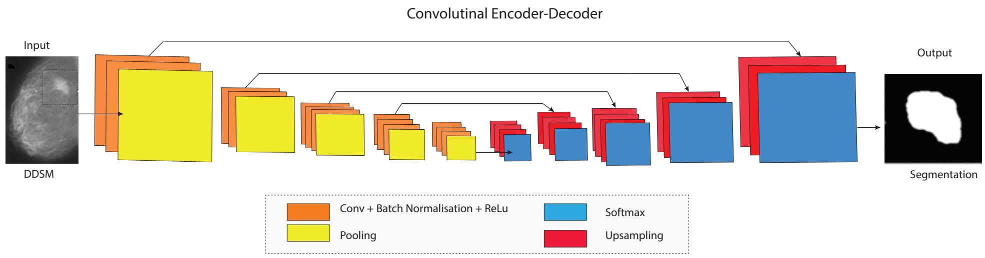


Fig. 5. SegNet architecture flowchart, adapted from Badrinarayanan, V. [75]

We used the same hyperparameters values for training U-net and SegNet networks (see Table 3).

2.6 Classification

The Deep Residual Network (ResNet) is one of the pre-trained models used in transfer learning, particularly in computer vision. It has been introduced for automatic feature extraction and classification, to address the problem of vanishing gradients providing good performance with less training time and fewer data samples compared to training a deep network from scratch.

In this research, we chose to train the ResNet-50 model, which is widely used in medical image classification. The model skips one or more layers and manages the gradient vanishing problem, in addition to its ease optimization. Model accuracy can be improved by increasing its depth. Therefore, two or three layers of the ResNet-50 model are directly connected to each layer (not to the adjacent layer), employing the ReLu non-linear activation function (Figure 6). The hyperparameters details are described below:

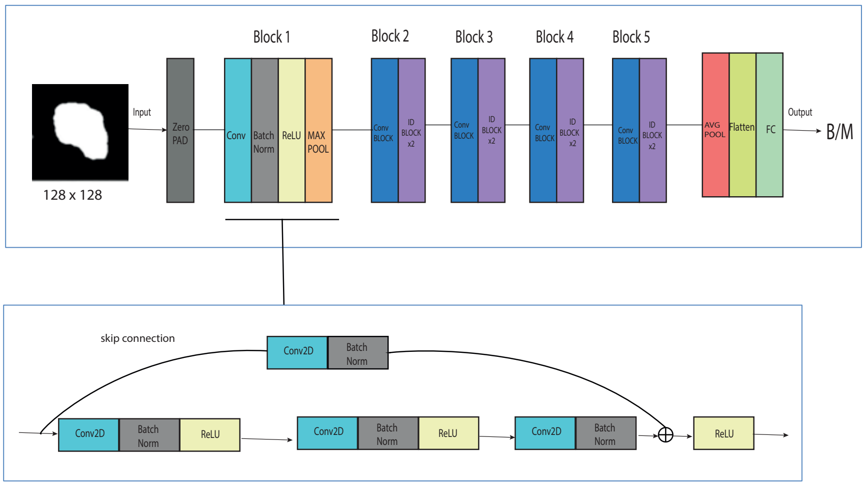


Fig. 6. The ResNet-50 architecture consists of 5 blocks, each containing 3 convolutional identity blocks and 3 convolutional blocks with skip connections.

Resnet hyperparameters: We trained the model with data augmentation and without data augmentation using standard benchmark data sets with the next hyperparameter values: Adadelta optimizer ε=1e^{-07}, δ= 0.95, a learning rate = 0.001, batch size = 2, number of epochs = 60 with 2190 steps, categorical-cross entropy loss function over 2,190 iterations. We experimented with two activation functions: *ReLU* was used during the training of convolutional layers, while the *Sigmoid* function was used for binary class prediction.

All CNN models were training using a cloud service based on Jupyter Notebooks on Google Colab Pro GPU (model V100) and python libraries such as Keras, Matplolib and TensorFlow.

2.7 Statistical Measures

Likewise, the most frequently used statistical metrics for assessing image restoration and HR image quality are PSNR and SSIM index. A higher PSNR value indicates higher image quality, while a small value implies high numerical differences between images [87]. Typical PSNR values range from 30 dB to 50 dB, and SSIM values between -1 and 1, where 1 indicates perfect similarity, 0 indicates no similarity, and -1 indicates perfect anti-correlation (see Figure 1). For segmentation performance evaluation, we use Dice (F1 score) and Intersection over union (IoU or Jaccard index). Finally, for classification performance evaluation, we consider accuracy (Acc), precision (Prec), sensitivity (Sen), specificity (Spec) and Area Under the Curve (AUC) [11].

3. Results

This section discusses the most important results, assessment metrics, and graphs obtained from the network training.

3.1 Image Quality comparison

Table 4 presents the average PSNR and SSIM values after the image quality evaluation. These metrics offer valuable insights into the quality of the processed images when compared to high-resolution ground truth images from the CBIS-DDSM dataset. The EDSR model showed a significant improvement compared to the other model.

Tabla 4. The most relevant PSNR/SSIM values for x3 and x4 factor scaling.

ID	SR-RDN		ID	EDSR	
	PSNR (dB)	SSIM		PSNR (dB)	SSIM
DDSM_043__blur1	40.76	0.97	**DDSM_043__blur1**	46.41	0.97
DDSM_0607	40.46	0.95	**DDSM_0381__blur1**	46.23	0.97
DDSM_0120	40.33	0.95	**DDSM_0275__blur1**	45.95	0.97
DDSM_0468	40.22	0.94	**DDSM_0228__blur1**	45.76	0.97
DDSM_0466	40.20	0.96	**DDSM_0120**	45.63	0.91
DDSM_0168__blur1	40.18	0.94	**DDSM_0467**	45.57	0.97
DDSM_0188	40.14	0.94	**DDSM_0466**	45.51	0.97
DDSM_0275__blur1	40.14	0.97	**DDSM_0212__blur1**	45.38	0.97
DDSM_0369	40.05	0.95	**DDSM_079__blur1**	45.35	0.95
DDSM_0538	39.98	0.90	**DDSM_0368**	45.2	0.92
DDSM_0212__blur1	35.87	0.98	**DDSM_021__blur1**	39.5	1
DDSM_0228__blur1	37.00	0.97	**DDSM_0295__blur1**	38.17	1
DDSM_0275__blur1	40.14	0.97	**DDSM_0106__blur1**	41.14	0.97
DDSM_043__blur1	40.76	0.97	**DDSM_0212__blur1**	45.38	0.97
DDSM_0189__blur1	35.12	0.96	**DDSM_0228__blur1**	45.76	0.97
DDSM_0256__blur1	37.91	0.96	**DDSM_0275__blur1**	45.95	0.97
DDSM_0381__blur1	39.11	0.96	**DDSM_0381__blur1**	46.23	0.97
DDSM_0466	40.20	0.96	**DDSM_043__blur1**	46.41	0.97
DDSM_0467	39.95	0.96	**DDSM_0466**	45.51	0.97
DDSM_0549	38.83	0.96	**DDSM_0467**	45.57	0.97

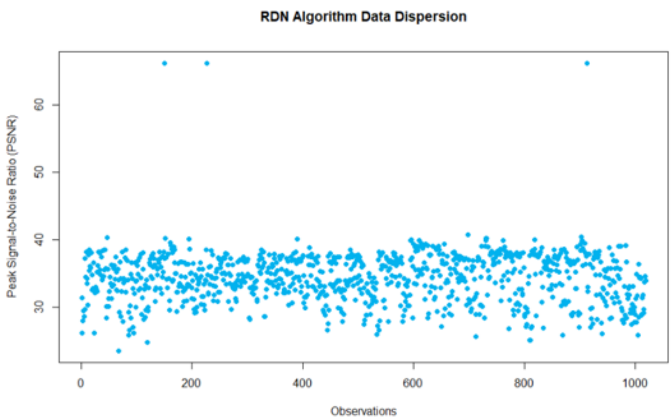
3.2 Data dispersion

Figures 7a-d present the statistical results, where a and b are the dispersion data obtained from SR-RDN. The blue data in Figure 7a represent the PSNR metric, ranging from 30 to 40 dB, while the red points in Figure 7b represent the SSIM metric, ranging from -1 to

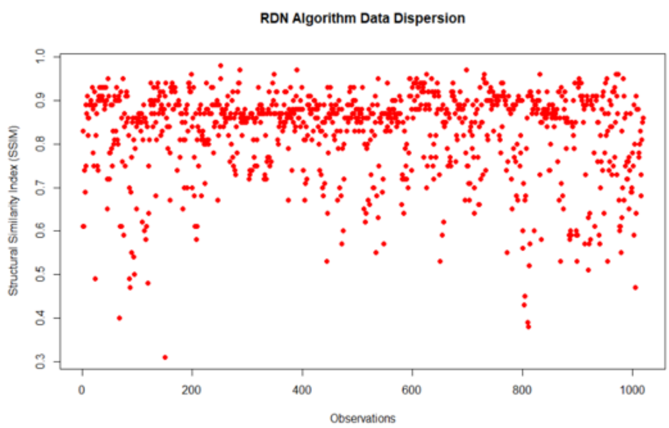
1. PSNR is used to measure the quality of the restored image when it is affected by noise and blur. Similarly, SSIM is defined as a function of luminance comparison.

The linear dependence factor is computed using the correlation coefficient in SSIM index and can be find broad applications in mammographic diagnosis and cancer detection fields [63].

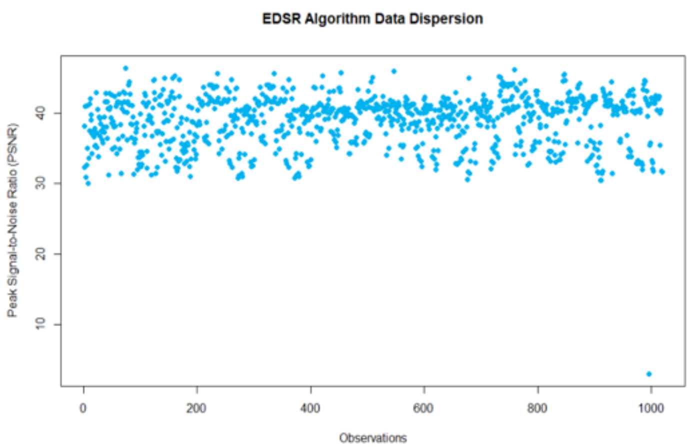
In figure 7c, the blue data show more signal with a higher quality rate and better-quality image using EDSR algorithm. Figure 7d presents better SSIM statistical results using EDSR in comparison to 7b using SR-RDN algorithm. It indicates better luminance (ranging from 0.90 to 0.95), contrast, and structural information in restructured EDSR images.



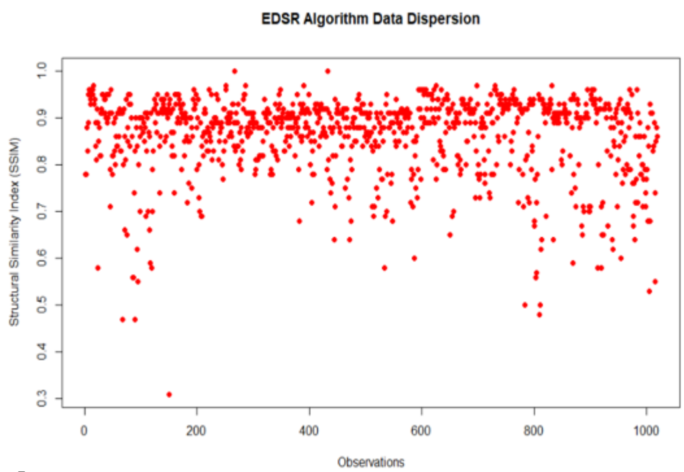
a.



b.



c.



d.

Fig. 7. Dispersion values in both super-resolution algorithms SRRDN and EDSR. (a) PSNR vs. observations in RDN algorithm. (b) SSIM vs. observations in RDN algorithm. (c) PSNR vs. observations in EDSR algorithm. (d) SSIM vs. observations in EDSR algorithm.

3.3 SegNet and Unet comparison

The experiments (see table 6) demonstrate that UNet achieved better RoI segmentation performance (IoU=0.862, Dice=0.991 and Acc=0.947) than the SegNet model, across all mammogram datasets used in this study.

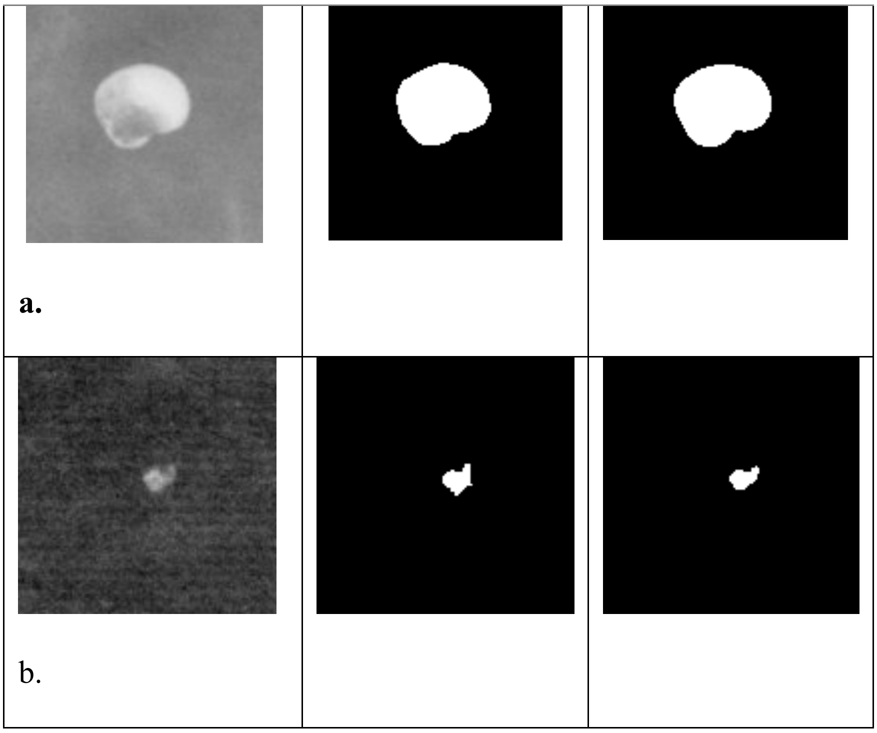
Tabla 6. Average values compared from segmentation models results.

Method	**Segmentation metrics**					
	Acc	**Prec**	**Sen**	**Spec**	**Dice**	**IoU**
Unet	0.947	0.930	0.925	0.956	0.991	0.862
SegNet	0.889	0.848	0.836	0.950	0.810	0.709

Table 7 displays the RoI input image with their manual segmentation and automatic segmentation using U-net and SegNet models.

Table 7. Comparison between manual segmentation and automatic RoI segmentation (Unet/Segnet) from original RoI images: a) DDSM_0504), b) DDSM_0526.

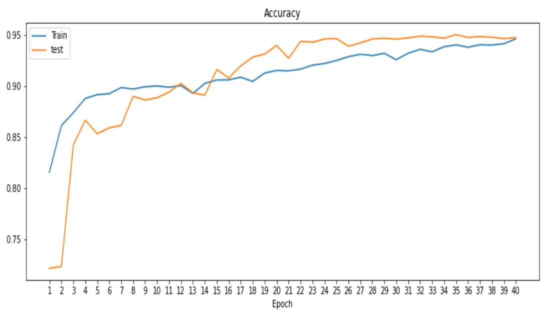
Original RoI	**Manual Segmentation**	**Unet/Segnet**



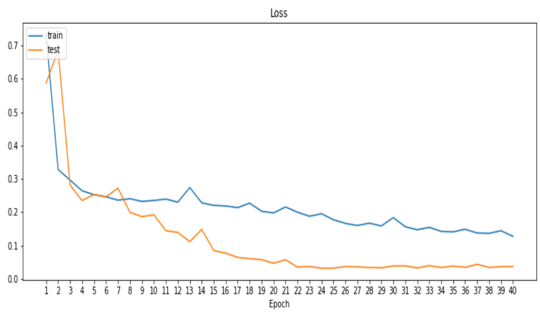
a.

b.

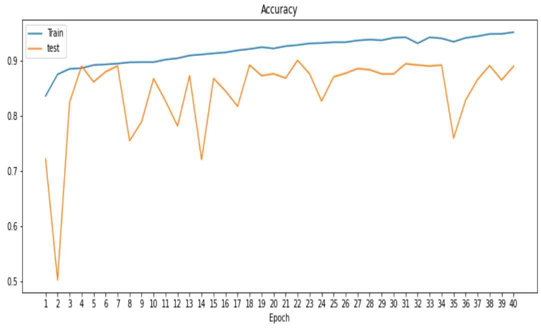
Therefore, it is important to monitor their evolution and performance of the models during training and validation. Figure 8 presents plots of the indices obtained for each epoch during the testing of different models. Figures 8a and 8b display the loss value and accuracy by each epoch in the Unet, while Figures 8c and 8d show the loss value and accuracy by each epoch in Segnet. The results indicate that the Unet model is more stable, with a consistent learning rate, and does not exhibit overfitting.

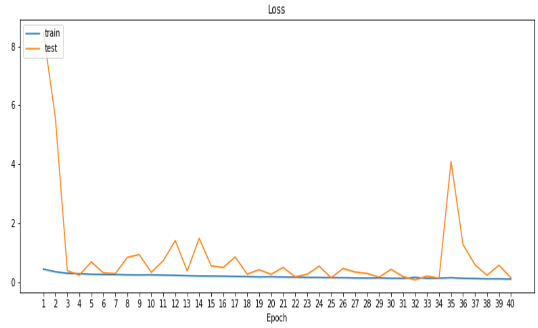


a.



b.





c. d.

Fig. 8. Accuracy and loss values in a. Unet training dataset, b. Unet validation dataset c. Segnet training dataset and d. Segnet validation dataset.

3.4 Classification using ResNet-50

The results of the classification experiments are described in Table 8, where the values show a high accuracy using the hyperparameters with data augmentation.

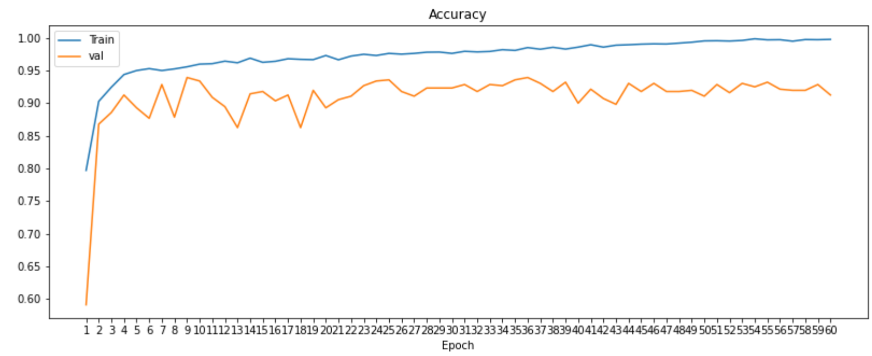
Tabla 8. ResNet-50 evaluation metrics on two datasets.

ResNet-50 model	**Average classification metrics**					
	Image number	**Acc**	**Prec**	**Sen**	**Spec**	**F1 score**
Data augmentation	5486	0.75	0.68	0.55	0.82	0.64
Without Data augmentation	784	0.68	0.65	0.77	0.59	0.71

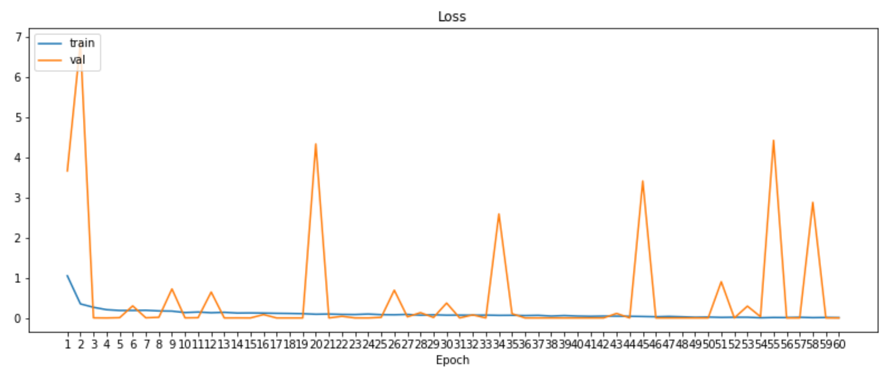
The ResNet model improved the accuracy to 75%, surpassing other traditional machine learning algorithms when using enhanced RoI data augmentation images by affine transformation.

The ResNet model utilizes a loss function to indicate it proximity to making correct predictions. The risk during the training process is the potential for the model to overfit to the training set, meaning it might learn an overly specific function that performs well on the training data but fails to generalize to unseen images.

Figures 9a and 9b illustrate that the model learns effectively from the training and validation data sets.



a.



b.

Fig. 9. (a) Line plots of model accuracy on the training (blue) and validation (orange) datasets. (b) Loss values on the training (blue) and validation (orange) datasets for each epoch.

This indicates that the model is likely to perform well on new images. However, it's important to note that while these factors can contribute to a model's ability to perform well on new images, there are no guarantees. The real-world performance of a machine learning model often requires ongoing monitoring, refinement, and adaptation to changing data distributions and conditions.

4.Discussion

DM images are often acquired at lower resolutions to minimise radiation exposure while maintaining adequate diagnostic quality. However, low-resolution images can compromise the ability to detect subtle features or abnormalities, such as microcalcifications in breast cancer lesions. The use of CNNs for SR image enhancement and segmentation can indeed be valuable and significantly improve the visibility of fine details, making it easier for radiologists to identify and classify breast cancer lesions, especially in low-resource settings where DM images may have poor resolution.

After enhancing the image resolution, the next step is to locate and delineate regions of interest, such as potential breast cancer lesions, within the mammogram. Once the image is enhanced and the lesions are segmented, CNNs can be used for the actual classification of breast cancer lesions into benign and malignant.

In our work, to test the proposed method, three public mammography databases were selected: CBIS-DDSM, Mini-MIAS and Inbreast. In the SR task the EDSR results provide high enhancement in image quality, with the PSNR and SSIM index (39.05 dB and 0.90) exceeding those of SR-RDN (32.68 dB and 0.82).

The average EDSR index values indicated that successfully reconstructed of detailed textures and edges in the RoIs and exhibited better quality output, in comparison with other results in the literature [45-47]. Lim et al. [45] proposed the EDSR CNN-based algorithm to improve super resolution in natural image (DIV2K), by removing the batch normalization layer, accelerating the training process, and achieving better performance (PSNR of 32.62 dB and SSIM of 0.8984) compared to other methods such as ESPCN (30.90 dB/-), VDSR (31.35 dB/0.8838), DRCN (31.53 dB/0.8838), SRResNet (32.05 dB/0.9019), RDN (32.61 dB/0.9003), MDSR (32.60 dB/0.8982), and DBPN (32.47 dB/0.898).

Additionally, another custom CNN (Unet) was used to perform image segmentation on the high-resolution RoIs generated by EDSR, achieving an average Intersection over Union (IoU) of 0.862, an average Dice similarity coefficient of 0.991, and an accuracy of 0.947 in segmentation results, surpassing the results of the SegNet model.

Similar research has proposed UNet as a segmentation network. Almajalid et al. [73] used UNet for breast tumour segmentation using ultrasound images and achieved a F1 score of 0.994 with the training set and 0.8252 with the testing set. Likewise, Zhou et al. [88] improved UNet by using skip connections and achieved an average IoU gain of 3.9 over the standard U-Net.

Our results align with Vianna et al. [88], who compared U-Net and SegNet for the breast lesions segmentation in ultrasonography, were U-Net demonstrated better Dice results (86.35%) compared to Segnet-Dice of 81.1%.

By analysing the segmented regions, ResNet50 provides a classification or likelihood score for the presence of breast lesions. Table 8 shows enhanced model results for breast lesion image classification using transfer learning with data augmentation through affine transformation, in comparison with training model without data augmentation.

However, the average classification results show that synthetic data cannot fully substitute real images for training CNN classifiers. The absence of real images in the training set can lead to overfitting and lower model accuracy (75%) compared to other state-of-the-art deep CNN-based classifiers, such as Dense Convolutional Network (Densenet) [80,90-91], VGGNet [92], and Inception [93]. Chen et al. [94] introduced data augmentation and ResNet transfer learning for the automatic extraction of features and classification of mammography images, achieving good performance metrics (Acc= 93.15%, Spe=92.17%, Sen=93.83%, AUC=0.95, and loss=0.15). Similarly, Wu et al. [79] presented a deep CNN-ResNet method for breast cancer classification, achieving an AUC of 0.895 in predicting the presence of breast cancer.

These results may be attributed to the theory explained by Lan et al. [95] regarding generated images by traditional augmentation methods. Such images tend to share a similar distribution with the original ones and may not be suitable for processing medical images. Guan et al. [96] demonstrated that RoIs generated by GANs are more similar to real RoIs than affine-transformed RoIs in terms of mean, standard deviation, skewness, and entropy.

However, we acknowledge that breast tumour classification using DM has limitations when using traditional data augmentation, as the model did not significantly improve accuracy (75%) compared to other state-of-the-art deep CNN-based classifiers. This limitation could potentially be overcome by incorporating GAN models for the generation of synthetic data.

One major limitation of our work is the limited number of SR studies based on CNN models using 2D breast images such as mammography and ultrasound, as most of the literature primarily focuses on urban and natural images. In our study, we developed a deep CNN approach for mammography SR, segmentation, and classification of RoIs, resulting in good indices and quality values.

In summary, employing CNNs for SR image enhancement, segmentation, and breast cancer lesion classification can significantly enhance diagnostic capabilities in DM, particularly in settings where resource constraints may limit traditional diagnostic approaches. This approach has the potential to improve early detection and enhance patient outcomes in breast cancer diagnosis.

Conclusions

This article has presented a novel Computer-Aided Diagnosis (CAD) system framework based on deep learning for breast mammography super-resolution, segmentation, and classification, utilizing the concept of transfer learning. We implemented data augmentation with affine transformations for RoIs to enhance the performance of CNN networks. The synthetic RoIs data served as input to two different SR-CNN algorithms, EDSR and SR-RDN, resulting in improved image quality with enhanced resolution and precision for the subsequent segmentation and classification tasks.

We found that EDSR outperformed SR-RDN in the super-resolution task, as evidenced by higher PSNR and SSIM indices. Additionally, the U-Net model was selected as the preferred RoI segmentation technique due to its more reliable results, as demonstrated by Dice, Intersection over Union (IoU), and accuracy metrics.

However, while the ResNet-50 architecture improved accuracy over traditional machine learning algorithms when using generated images with affine transformations, it could not achieve the same accuracy (75%) as other state-of-the-art deep classifiers. This limitation can be attributed to the inability of traditional data augmentation to accurately simulate the real distribution of medical images, as opposed to generative models.

In summary, the importance of SR in lesion segmentation depends on the specific context and characteristics of the medical images in question. Clinical validation and evaluation should guide the selection of image enhancement methods to ensure they enhance diagnostic accuracy without introducing unintended effects.

Future research may involve comparative studies to assess the impact of SR on segmentation accuracy. Additionally, data augmentation techniques based on Generative Adversarial Network (GAN) models will be explored for the generation of synthetic mammography data. This synthetic data could be used as a training dataset for alternative breast mass classifiers based on convolutional networks (e.g., DenseNet, NasNet, VGGNet) with the aim of improving breast lesion classification accuracy and reducing overfitting.

Funding. This project has received co-financed from the Spanish Government Grant PID2019-107790RB-C22, "Software development for a continuous PET crystal systems applied to breast cancer".

Author Contributions: Conceptualization, Y.J.G., V.L.; methodology, Y.J.G,D.C, A.V, D.C.M; formal analysis, Y.J.G., M.J.R.-Á., and V.L.; investigation, Y.J.G.; resources, Y.J.G.; writing—original draft preparation, Y.J.G.; writing—review and editing, Y.J.G., M.J.R.-Á., and V.L.; visualization, Y.J.G.; supervision, M.J.R.-Á. and V.L.; project administration, M.J.R.-Á. and V.L.; funding acquisition M.J.R.-Á. All authors have read and agreed to the published version of the manuscript.

Disclosure statment. "The authors report there are no competing conflicts of interest to declare."

Data availability. The data that support the findings of this study are openly available in Code 1 https://opticapublishing.figshare.com/s/d3fb2f7112f41c58fa07 (Ref. [96]).

Abbreviations:

AUC	Area under curve
CAD	Computer aided system
CC	Cranio caudal
CNN	Convolutional neural network
DNN	Deep neural network
DDSM	Digital Database for Screening Mammography
DM	Digital mammography
DL	Deep learning
EDSR	Enhanced Deep Residual Network
E2E	End to End
ESRGAN	Enhanced Super-Resolution Generative Adversarial Networks
ESPCN	Efficient sub-pixel convolutional neural network
GAN	Generative adversarial network
HR	High resolution
IoU	Intersection over Union
LR	Low resolution
MDSR	Multi-scale deep super-resolution
MLO	Mediolateral Oblique
PSNR	Peak signal to Noise Ratio
RoI	Region of interest
RDN	Residua Dense Network
RDB	Residual Dense Block
RNN	Recurrent Neural Network
ReLU	Rectified Linear Unit
SR-GAN	Super-Resolution Using a Generative Adversarial Network
SSIM	Structural Similarity Index Metric
SISR	Single image super resolution
SegNet	Segmentation Network
TP	True positive
TN	True negative
FP	False positive
FN	False negative
VGG	Visual geometric group
VDSR	Very Deep Network for SR

References

[1]. American Cancer Society. 2018. Global Cancer: Facts & Figures, 4th edition URL:http://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/global-cancer-facts-and-figures/global-cancer-facts-and-figures-4th-edition.Pdf

[2]. Oeffinger, K. C., Fontham, E. T., Etzioni, R., et al.(2015). Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. Jama, 314(15), 1599-1614.

[3]. Mainiero, M. B., Lourenco, A., Mahoney, M. C., Newell, M. S., et al. (2016). ACR appropriateness criteria breast cancer screening. Journal of the American College of Radiology, 13(11), R45-R49.

[4]. Mann, R. M., Kuhl, C. K., Kinkel, K., & Boetes, C. (2008). Breast MRI: guidelines from the European society of breast imaging. European radiology, 18(7), 1307-1318.

[5]. Jalalian, A., Mashohor, S. B., Mahmud, H. R., et al. (2013). Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. Clinical imaging, 37(3), 420-426.

[6]. Sarno, A., Mettivier, G., & Russo, P. (2015). Dedicated breast computed tomography: basic aspects. Medical physics, 42(6Part1), 2786-2804.

[7]. Ponraj, D. N., Jenifer, M. E., Poongodi, P., & Manoharan, J. S. (2011). A survey on the preprocessing techniques of mammogram for the detection of breast cancer. Journal of Emerging Trends in Computing and Information Sciences, 2(12), 656-664.

[8]. Kalles, V., Zografos, G. C., Provatopoulou, X., Koulocheri, D., & Gounaris, A. (2013). The current status of positron emission mammography in breast cancer diagnosis. Breast Cancer, 20(2), 123-130.

[9]. Abdallah, Y. M., Elgak, S., Zain, H., Rafiq, M., Ebaid, E. A., & Elnaema, A. A. (2018). Breast cancer detection using image enhancement and segmentation algorithms.

[10]. Hendrick, R. E. (2010). Radiation doses and cancer risks from breast imaging studies. Radiology, 257(1), 246-253.

[11]. Jiménez-Gaona, Y., Rodríguez-Álvarez, M. J., & Lakshminarayanan, V. (2020). Deep-Learning-Based Computer-Aided Systems for Breast Cancer Imaging: A Critical Review. Applied Sciences, 10(22), 8298.

[12]. Chiao, J. Y., Chen, K. Y., Liao, K. Y. K., Hsieh, P. H., Zhang, G., & Huang, T. C. (2019). Detection and classification the breast tumors using mask R-CNN on sonograms. Medicine, 98(19).

[13]. Chougrad, H., Zouaki, H., & Alheyane, O. (2018). Deep convolutional neural networks for breast cancer screening. Computer methods and programs in biomedicine, 157, 19-30.

[14]. Hussain, Z., Gimenez, F., Yi, D., & Rubin, D. (2018). Differential data augmentation techniques for medical imaging classification tasks. [American Medical Informatics Association.]. AMIA ... Annual Symposium Proceedings - AMIA Symposium. AMIA Symposium, 2017, 979–984.

[15]. Yu, X., Wu, X., Luo, C., & Ren, P. (2017). Deep learning in remote sensing scene classification: A data augmentation enhanced convolutional neural network framework. GIScience & Remote Sensing, 54(5), 741–758. https://doi.org/10.1080/15481603.2017.1323377

[16]. Munir, K., Elahi, H., Ayub, A., Frezza, F., & Rizzi, A. (2019). Cancer diagnosis using deep learning: a bibliographic review. Cancers, 11(9), 1235.

[17]. Salem, M. A. M. (2018). Mammogram-Based cancer detection using deep convolutional neural networks. In 2018 13th International Conference on Computer Engineering and Systems (ICCES) (pp. 694-699). IEEE.

[18]. Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence, 38(2), 295–307.

[19]. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).

[20]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

21. Li H, Meng X, Wang T, Tang Y, Yin Y. Breast masses in mammography classification with local contour features. Biomed Eng Online. 2017;16(1):44. doi:10.1186/s12938-017-0332-0

[22]. Xu M, Qi S, Yue Y, et al. Segmentation of lung parenchyma in CT images using CNN trained with the clustering algorithm generated dataset. Biomed Eng Online. 2019;18(1):2. doi:10.1186/s12938-018-0619-9

[23]. Robinson, M. D., Chiu, S. J., To th, C. A., Izatt, J. A., Lo, J. Y., & Farsiu, S. (2017). New applications of super-resolution in medical imaging. Book Chapet in Super-resolution imaging. Peyman Milanfar (Editor) (pp. 383-412). CRC Press.

[24]. Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J. H., & Liao, Q. (2019). Deep learning for single image super-resolution: A brief review. IEEE Transactions on Multimedia, 21(12), 3106-3121.

[25]. Wan, J., Yin, H., Chong, A. X., & Liu, Z. H. (2020). Progressive residual networks for image super-resolution. Applied Intelligence, 1-13.

[26]. Yang, C. Y., Ma, C., & Yang, M. H. (2014). Single-image super-resolution: A benchmark. In European Conference on Computer Vision (pp. 372-386). Springer, Cham.

[27]. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.

[28]. Motamed, S., Rogalla, P., & Khalvati, F. (2021). Data Augmentation Using Generative Adversarial Networks (GANs) For GAN-Based Detection Of Pneumonia And COVID-19 In Chest X-Ray Images.

[29]. Antoniou, A., Storkey, A., & Edwards, H. (2017). Data augmentation generative adversarial networks. arXiv preprint arXiv:1711.04340.

[30]. Tran, N. T., Tran, V. H., Nguyen, N. B., Nguyen, T. K., & Cheung, N. M. (2021). On data augmentation for GAN training. IEEE Transactions on Image Processing, 30, 1882-1897.

[31]. Yadav, S. S., & Jadhav, S. M. (2019). Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data, 6(1), 1-18

[32]. Gulati, T., Sengupta, S., & Lakshminarayanan, V. (2020). Application of an enhanced deep super-resolution network in retinal image analysis. In Ophthalmic Technologies XXX (Vol. 11218, p. 112181K). International Society for Optics and Photonics.

[33]. Freeman, W. T., Jones, T. R., & Pasztor, E. C. (2002). Example-based super-resolution. IEEE Computer graphics and Applications, 22(2), 56-65.

[34]. Chang, H., Yeung, D. Y., & Xiong, Y. (2004). Super-resolution through neighbor embedding. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. (Vol. 1, pp. I-I). IEEE.

[35]. Gribbon, K. T., & Bailey, D. G. (2004). A novel approach to real-time bilinear interpolation. In Proceedings. DELTA 2004. Second IEEE International Workshop on Electronic Design, Test and Applications (pp. 126-131). IEEE.

[36]. Wang, Z., Chen, J., & Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. IEEE transactions on pattern analysis and machine intelligence.4

[37]. Zhang, Yulun et al. (2018). “Residual Dense Network for Image Super-Resolution.” : 0–10.Jiang, Y., Chen, L., Zhang, H., & Xiao, X. (2019). Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. PloS one, 14(3), e0214587.

[38]. Dong, C., Loy, C.C., He, K. and Tang, X. (2014). Learning a Deep Convolutional Network for Image Super-Resolution. Proceedings of European Conference on Computer Vision, Zurich, 6-12 September, 184-199.

[39]. Dong, C., Loy, C.C., He, K. and Tang, X. (2016). Image Super-Resolution using Deep Convolutional Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, 295-307. https://doi.org/10.1109/TPAMI.2015.2439281

[40]. Umehara, K., Ota, J., & Ishida, T. (2017). Super-resolution imaging of mammograms based on the super-resolution convolutional neural network. Open Journal of Medical Imaging, 7(4), 180-195.

[41]. Jiang, Y., & Li, J. (2020). Generative Adversarial Network for Image Super-Resolution Combining Texture Loss. Applied Sciences, 10(5), 1729.

[42]. Tschuchnig, M. E., Oostingh, G. J., & Gadermayr, M. (2020). Generative Adversarial Networks in Digital Pathology: A Survey on Trends and Future Potential. arXiv preprint arXiv:2004.14936.

[43]. Ledig, C., Theis, L., Huszár, F., Caballero, J.,et al, (2017). Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4681-4690).

[44]. Sood, R., Topiwala, B., Choutagunta, K., Sood, R., & Rusu, M. (2018). An application of generative adversarial networks for super resolution medical imaging. In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 326-331). IEEE.

[45]. Lim, B., Son, S., Kim, H., Nah, S., & Mu Lee, K. (2017). Enhanced deep residual networks for single image super-resolution. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 136-144).

[46]. Abbass, M. Y. (2020). Residual Dense Convolutional Neural Network for Image Super-Resolution. Optik, 165341.

[47]. Wang, X., Yu, K., Wu, S., et al.(2018). Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 0-0).

[48]. Kim, J., Kwon Lee, J., & Mu Lee, K. (2016). Accurate image super-resolution using very deep convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1646-1654).

[49]. Shi, W., Caballero, J., Huszár, F., et al.(2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1874-1883).

[50]. Wang, Z., Liu, D., Yang, J., Han, W., & Huang, T. (2015). Deep networks for image super-resolution with sparse prior. In Proceedings of the IEEE international conference on computer vision (pp. 370-378).

[51]. Kim, J., Kwon Lee, J., & Mu Lee, K. (2016). Deeply-recursive convolutional network for image super-resolution. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1637-1645)

[52]. Tai, Y., Yang, J., & Liu, X. (2017). Image super-resolution via deep recursive residual network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3147-3155).

[53]. Al-Masni, M. A., Al-Antari, M. A., Park, J. M., Gi, G., et al (2018). Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. Computer methods and programs in biomedicine, 157, 85-94.

[54]. Geras, K. J., Wolfson, S., Shen, Y., Wu, N., Kim, S., Kim, E. et al (2017). High-resolution breast cancer screening with multi-view deep convolutional neural networks. arXiv preprint arXiv:1703.07047.

[55]. Lee, J. G., Jun, S., Cho, Y. W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: general overview. Korean journal of radiology, 18(4), 570-584.

[56]. S. Sengupta, A. Singh, H. A. Leopold, T. Gulati, V. Lakshminarayanan (2020). Ophthalmic diagnosis using deep learning with fundus images – A critical review, Artificial Intelligence in Medicine, Volume 102,2020,101758,ISSN 0933-3657, https://doi.org/10.1016/j.artmed.2019.101758

[57]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[58]. Singh, V. K., Rashwan, H. A., Romani, S., et al,. (2020). Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network. Expert Systems with Applications, 139, 112855.

[59]. Tong, T., Li, G., Liu, X., & Gao, Q. (2017). Image super-resolution using dense skip connections. In Proceedings of the IEEE International Conference on Computer Vision (pp. 4799-4807).

[60]. Yang, J., Wright, J., Huang, T. S., & Ma, Y. (2010). Image super-resolution via sparse representation. IEEE transactions on image processing, 19(11), 2861-2873.

[61]. Schulter, S., Leistner, C., & Bischof, H. (2015). Fast and accurate image upscaling with super-resolution forests. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3791-3799).

[62]. Kaji, S., & Kida, S. (2019). Overview of image-to-image translation by use of deep neural networks: denoising, super-resolution, modality conversion, and reconstruction in medical imaging. Radiological physics and technology, 12(3), 235-248.

[63]. Rajkumar, S., & Malathi, G. (2016). A comparative analysis on image quality assessment for real time satellite images. Indian J. Sci. Technol, 9(34).

[64]. Juhong, A., Li, B., Yao, C. Y., Yang, C. W., Agnew, D. W., Lei, Y. L., ... & Qiu, Z. (2023). Super-resolution and segmentation deep learning for breast cancer histopathology image analysis. Biomedical Optics Express, 14(1), 18-36.

[65].Sheba, K. U., & Raj, S. G. (2016). Objective quality assessment of image enhancement methods in digital mammography–a comparative study. Signal & Image processing: An International Journal, 7(4), 1-13.

[66]. Gao, X., Lu, W., Tao, D., & Li, X. (2010). Image quality assessment and human visual system. In Visual Communications and Image Processing 2010 (Vol. 7744, p. 77440Z). International Society for Optics and Photonics.

[67]. Abdelhafiz, D., Yang, C., Ammar, R., & Nabavi, S. (2019). Deep convolutional neural networks for mammography: advances, challenges and applications. BMC bioinformatics, 20(11), 281.

[68].Soulami, K. B., Kaabouch, N., Saidi, M. N., & Tamtaoui, A. (2021). Breast cancer: One-stage automated detection, segmentation, and classification of digital mammograms using UNet model based-semantic segmentation. Biomedical Signal Processing and Control, 66, 102481.

[69]. Valvano, G., Santini, G., Martini, N., Ripoli, A., Iacconi, C., Chiappino, D., & Della Latta, D. (2019). Convolutional neural networks for the segmentation of microcalcification in mammography imaging. Journal of Healthcare Engineering, 2019.

[70].Abdelhafiz, D., Bi, J., Ammar, R., Yang, C., & Nabavi, S. (2020). Convolutional neural network for automated mass segmentation in mammography. BMC bioinformatics, 21(1), 1-19.

[71].Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[72].He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[73].Almajalid, R., Shan, J., Du, Y., & Zhang, M. (2018). Development of a deep-learning-based method for breast ultrasound image segmentation. In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 1103-1108). IEEE.

[74].Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.

[75].Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence, 39(12), 2481-2495.

[76].Romera, E., Alvarez, J. M., Bergasa, L. M., & Arroyo, R. (2017). Erfnet: Efficient residual factorized convnet for real-time semantic segmentation. IEEE Transactions on Intelligent Transportation Systems, 19(1), 263-272.

[77]. Lazo, J. F., Moccia, S., Frontoni, E., & De Momi, E. (2020). Comparison of different CNNs for breast tumor classification from ultrasound images. arXiv preprint arXiv:2012.14517.

[78]. Alkhaleefah, M., Chittem, P. K., Achhannagari, V. P., Ma, S. C., & Chang, Y. L. (2020). The influence of image augmentation on breast lesion classification using transfer learning. In 2020 International Conference on Artificial Intelligence and Signal Processing (AISP) (pp. 1-5). IEEE.

[79]. Wu, N., Phang, J., Park, J., Shen, Y., Huang, Z., Zorin, M., ... & Geras, K. J. (2019). Deep neural networks improve radiologists' performance in breast cancer screening. IEEE transactions on medical imaging, 39(4), 1184-1194.

[80]. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[81]. Ferreira, C. A., Melo, T., Sousa, P., Meyer, M. I., Shakibapour, E., Costa, P., & Campilho, A. (2018). Classification of breast cancer histology images through transfer learning using a pre-trained inception resnet v2. In International Conference Image Analysis and Recognition (pp. 763-770).

[82]. Lee, R., Gimenez, F., Hoogi, A. et al. A curated mammography data set for use in computer-aided detection and diagnosis research. Sci Data 4, 170177 (2017). https://doi.org/10.1038/sdata.2017.177

[83]. Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., Ricketts, I., et al. (2015). Mammographic Image Analysis Society (MIAS) database v1.21 [Dataset]. https://www.repository.cam.ac.uk/handle/1810/250394

[84].Li, S., Hatanaka, Y., Fujita, H., Hara, T., & Endo, T. (1999). Automated Detection of Mammographic Masses in MIAS Database. Medical Imaging Technology, 17(4), 427-428.

[85]. Moreira, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J., & Cardoso, J. S. (2012). Inbreast: toward a full-field digital mammographic database. Academic radiology, 19(2), 236-248.

[86].Yu, T., & Zhu, H. (2020). Hyper-parameter optimization: A review of algorithms and applications. arXiv preprint arXiv:2003.05689.

[87]. Sara, U., Akter, M., & Uddin, M. S. (2019). Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. Journal of Computer and Communications, 7(3), 8-18.

[88] Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2019). Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE transactions on medical imaging, 39(6), 1856-1867.

[89].Vianna, P., Farias, R., & de Albuquerque Pereira, W. C. (2021). U-Net and SegNet performances on lesion segmentation of breast ultrasonography images. Research on Biomedical Engineering, 37(2), 171-179.

[90].Jiménez Gaona Y., Rodriguez-Alvarez M.J., Espino-Morato H., Castillo Malla D., Lakshminarayanan V. (2021) DenseNet for Breast Tumor Classification in Mammographic Images. In: Rojas I., Castillo-Secilla D., Herrera L.J., Pomares H. (eds) Bioengineering and Biomedical Signal and Image Processing. BIOMESIP 2021. Lecture Notes in Computer Science, vol 12940. Springer, Cham. https://doi.org/10.1007/978-3-030-88163-4_16

[91]. Rybiałek, A., & Jeleń, Ł. (2020). Application of DenseNets for Classification of Breast Cancer Mammograms. In International Conference on Computer Information Systems and Industrial Management (pp. 266-277). Springer, Cham.

[92].Jahangeer, G. S. B., & Rajkumar, T. D. (2021). Early detection of breast cancer using hybrid of series network and VGG-16. Multimedia Tools and Applications, 80(5), 7853-7886.

[93].Khan, S., Islam, N., Jan, Z., Din, I. U., & Rodrigues, J. J. C. (2019). A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognition Letters, 125, 1-6.

[94].Chen, Y., Zhang, Q., Wu, Y., Liu, B., Wang, M., & Lin, Y. (2018). Fine-tuning ResNet for breast cancer classification from mammography. In The International Conference on Healthcare Science and Engineering (pp. 83-96). Springer, Singapore.

[95].Lan, L., You, L., Zhang, Z., Fan, Z., Zhao, W., Zeng, N., ... & Zhou, X. (2020). Generative adversarial networks and its applications in biomedical informatics. Frontiers in Public Health, 8, 164.

[96]. Guan, S., & Loew, M. (2019). Breast cancer detection using synthetic mammograms from generative adversarial networks in convolutional neural networks. Journal of Medical Imaging, 6(3), 031411.

[96].Y. Jimenez-Gaona, "Breast Mass Regions Classification from Mammograms using Convolutional Neural Networks and Transfer Learning", figshare (2023), https://opticapublishing.figshare.com/s/d3fb2f7112f41c58fa07