IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 6 December 2022, accepted 24 December 2022, date of publication 13 January 2023, date of current version 20 January 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3236612

RESEARCH ARTICLE

A New Fat-Removal-Based Preprocessing Pipeline for MLO View in Digital Mammograms

JUANITA HERNÁNDEZ LÓPEZ[1], JUAN HUMBERTO SOSSA AZUELA[1], (Senior Member, IEEE), ALBERTO SALVADOR NUÑEZ VARELA[2], CÉSAR AUGUSTO RAMÍREZ GAMEZ[2], VIRGINIA CANSECO GONZÁLEZ[3], FRANCISCO EDUARDO MARTÍNEZ PÉREZ[2], SANDRA EDITH NAVA MUÑOZ[2], HÉCTOR GERARDO PÉREZ GONZÁLEZ[2], AND JOSÉ IGNACIO NÚÑEZ VARELA[2]

[1]Centro de Investigación en Computación del Instituto Politécnico Nacional, Mexico City 07738, Mexico
[2]Facultad de Ingeniería, Universidad Autónoma de San Luis Potosí, San Luis Potosi 78290, Mexico
[3]Hospital Central Dr. Ignacio Morones Prieto, San Luis Potosi 78290, Mexico

Corresponding author: Juanita Hernández López (juanita.hernandez@cinvestav.mx)

This work was supported in part by Consejo Nacional de Ciencia y Tecnología (CONACYT)–Fondo Institucional para el Desarrollo Científico, Tecnológico y de Innovación "FORDECYT-PRONACES" in "Sistema automático para la interpretación inmediata de mamografías para la determinación del riesgo de cáncer" under Project 6005.

ABSTRACT Specific anatomical structures from the female body, such as the axillary slope, armpit, pectoral muscle, or abdominal tissue, can be present in mammograms and might affect the proper mammogram analysis, especially in female populations with overweight issues, as is the case in Mexico. This work aims to determine if better results can be obtained in an automatic mammogram analysis by removing the abdominal and axillary fatty tissues as a preprocessing step. The experimentation is carried out by applying a pectoral muscle segmentation technique in a Mexican mammogram dataset and comparing the results by removing the fatty tissues from the same mammograms. Furthermore, the same experimentation is performed using the Portuguese public datasets, INbreast and BCDR. The conducted experiment will allow us to determine the differences in results across different populations. The fat removal method is based on the breast contour-edge from which points of interest are detected to cut the fat tissue. Our results suggest that the proposed method is suitable as a preprocessing technique, obtaining a 94.18% of acceptance, according to the qualitative analysis performed, and showed that removing the fatty tissue yields better results if the mammogram contains significant fatty tissue such as in the mammograms of Mexican patients.

INDEX TERMS Axillary and abdominal fat removal method, breast cancer, mammography analysis, preprocessing technique.

I. INTRODUCTION

Breast cancer is a disease that mainly affects middle-aged-adult women [1], aged between 15 and 54 [2]. In Mexico and worldwide, the high mortality and incidence rates due to malignant neoplasms are increasing [3]. According to the World Health Organization (WHO), out of 1,350,000 cases of breast cancer, there are 460,000 deaths each year worldwide [4]. Also, the WHO forecasts that by 2025, the number of breast cancer cases will be 19.3 million [5]. An alternative for breast cancer early detection is a mammography study [5]. Mammography is a non-invasive imaging method that has been shown to reduce the mortality rate by 30-70% due to its high sensitivity in detecting breast cancer in early stages [6].

Mammographic imaging can be performed conventionally or digitally. The patient experience is the same. However, digital mammography does have some advantages over film-based mammography. Some of the main advantages of digital mammograms is the ability to store the mammograms for a long time without losing their quality, they are easy to replicate, and the patients are exposed to less radiation [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Yongjie Li.

Screening mammography is usually done on the left and right breasts to take two standard mammographic views, craniocaudal (CC) and mediolateral oblique (MLO) views [5]. The MLO is the most used view for mammographic analysis because it describes most of the breast tissue [8]. However, this view includes undesirable artifacts such as pectoral muscle that are not needed for breast analysis [9]. For this reason, the preprocessing step is almost always included before training methods, like machine learning and deep learning [10], [11], in the general pipeline of a computer-aided diagnosis system (CAD) [1], [12].

Several preprocessing techniques are commonly applied in digital mammograms to improve their quality and benefit the following steps in constructing a CAD system (e.g., classification). Many preprocessing techniques enhance mammography quality, such as denoising, contrast enhancement, etc. The use of a specific technique depends on the particular preprocessing task to be solved.

Typical MLO mammogram preprocessing techniques involve enhancing and delimiting breast tissue area and removing pectoral muscles [13], [14]. Several methods in the literature are focused on the segmentation process for breast and pectoral muscle delimitation [15]. However, these methods do not consider that in the real world, the mammographic acquisition may include unwanted anatomical components such as the axillary slope, a section of the arm, and fat in the armpit and abdominal tissues. Although these components are natural anatomical structures in the female body, they are not necessary for classification tasks [16], and these components may affect the proper characterization of breast tissue and the pectoral muscle removal to a lesser or greater extent.

Assume the case shown in Fig. 1 where a mammogram shows abdominal and axillary fat such that these two areas are greater or equal to the proportion of breast tissue. For algorithms that analyze the local texture of the pattern or methods that consider ratios between areas such as fat, glandular and dense tissues, such fat components may trigger a false positive regarding density labeling. Although such cases can be isolated and could depend on the patient and the population analyzed, we need to consider these cases because obesity issue in the Mexican population is a reality [17]. In this sense, we have identified a new preprocessing task to be addressed. The new preprocessing task consists of removing the abdominal and axillary fat included in mammograms to obtain a cleaner breast tissue area, which may be helpful for classification tasks (e.g., density analysis [2]). It is worth mentioning that other researchers have previously also detected this need [16], but they have not focused on removing the abdominal and axillary fat included in mammograms. So, we did not find similar work to our research process. To the best of our knowledge, no study to date has used for removal abdominal and axillary fat included in digital mammograms.

This paper proposes a pipeline that includes a new preprocessing method for fat removal based on the breast's contour edge. The primary motivation for constructing the fat removal technique was the anatomical components present in most digital MLO mammograms of Mexican patients. Therefore, the method's applicability was evaluated using mammograms from Mexican women. Still, to explore the method's use for preprocessing mammograms from other countries, we selected the BCDR and INbreast databases, including mammograms from Portuguese women. Then, a comparative study between Mexican and Portuguese populations was made.

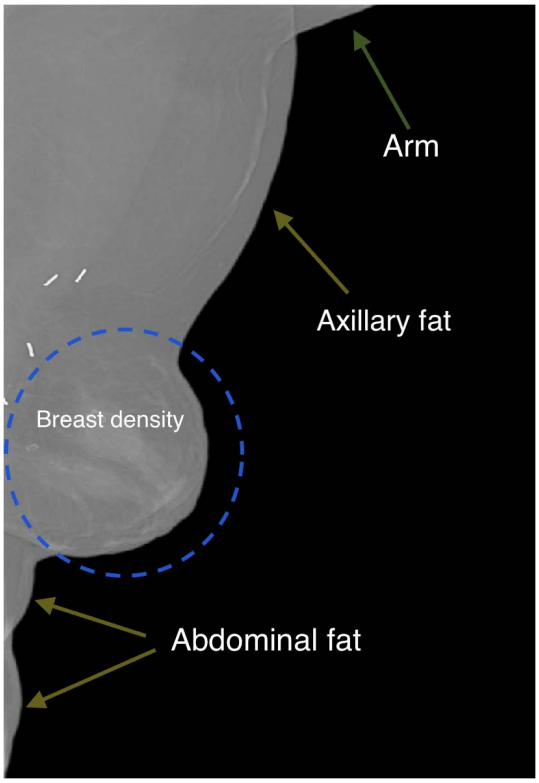

FIGURE 1. MLO view mammography with a part of the arm and axillary and abdominal fat.

The structure of this paper is as follows. Section two shows the related work regarding preprocessing techniques for MLO views in mammography. Section three describes the pipeline proposed for the preprocessing of mammograms; in this section, the fat-removal method is detailed. Section four describes the materials and methods used in the research. Section five presents the experimental results and a discussion of the results obtained. Finally, section six includes the conclusions and future work.

II. RELATED WORK

We have reviewed the most common MLO view mammography preprocessing techniques. Then, this section will discuss the techniques used for the preprocessing of mammograms. It has been decided to divide the methods into subsections depending on whether they are about breast tissue enhancement or pectoral muscle removal. Also, an additional subsection is for preprocessing approaches focused on unwanted anatomical components.

A. METHODS BASED ON ENHANCED AND DELIMITATION BREAST TISSUE

Accurate identification of the breast boundary on the mammogram presents two significant problems: 1) The contrast of the region near the breast boundary progressively decreases due to unbalanced compression of the breast tissue during mammography acquisition, so that the visibility of the breast skin line is poor. This complicates the correct detection of the breast boundary, 2) By having a non-uniform background that includes regions of high intensity, such as labels and annotations, the segmentation of the breast area can be negatively affected, which becomes challenging [14].

To improve the delineation of breast tissue, some techniques are applied. For example, mean filters, median filters, and normalization techniques are used to reduce the noise, and techniques such as adaptive histogram equalization (CLAHE), among others, are generally employed for image enhancement [1], [2].

For breast delineation, some techniques such as Gabor filters [18], the rolling ball algorithm [15], active contouring techniques [19], and methods based on neural networks [16] are applied. However, global or local thresholding techniques are frequently used for breast delineation [14]. The difficulty of the binarization techniques lies in selecting a strong threshold value [20] which can differentiate among regions (e.g., the area of the breast, pectoral muscle, and background [19]).

Also, a very notorious aspect of thresholding techniques is getting unsmoothed edges because, after removing noise or contrast enhancement, a thresholding technique is applied without considering the non-uniformity of the background [14], [19]. Considering those limitations, we include a breast delimitation method in addition to the fat removal method in the proposed pipeline containing a Gaussian filter to get mamograms with smoothed edges.

On the other hand, it is remarkable that most of the methods reported in state-of-the-art papers generally perform a qualitative evaluation performed by an expert. Moreover, few works consider an objective assessment of the results of a segmentation method using performance indexes such as Dice, Jaccard, and others [14].

B. METHODS BASED ON THE REMOVAL OF THE PECTORAL MUSCLE

The MLO projection includes the breast and the pectoral muscle because both structures are adjacent to each other [13]. Generally, the removal of the pectoral muscle is a necessary task because it contains high levels of intensity that can be easily confused by a classification method such as high density [21].

It has been mentioned in several publications that proper delineation of the pectoral muscle can be affected by different variables. For example, variability in the type of muscle shape (e.g., linear, concave, mixed), low contrast in the lower area of the muscle, confusion of the muscle border with the axillary cavity, and variability in size and contrast [2]. All these variations arise due to, among other reasons, the muscle's anatomy, the position of the patients during image acquisition, and skin folds [13].

Some works concerning removing the pectoral muscle are based on edge detection techniques [1], such as the Canny algorithm [6], [15] and contour maps [13]. These methods obtain the border of the pectoral muscle, but some limitations are slight distortions in the edge and disjoint areas.

Thresholding techniques are also applied to pectoral delimitation [22], [23]. Techniques such as region growth [19], [24], active contouring [25], and adaptive histogram equalization [26] among others [12], [27] are frequently used. However, some limitations in these kinds of techniques are unacceptable boundaries obtained for the under and over-segmentation.

Other kinds of methods are based on geometric rules [2], [28], which present an intuitive structure for removing pectoral muscles. However, these methods generally cut the pectoral muscle with a line without considering that the pectoral muscles vary in shape (e.g., curved and convex pectoral muscles).

Finally, methods based on learning features are based on artificial neural networks [14] and deep neural networks [21], [29], [30]. Some of the most relevant are based on semantic segmentation [31], [32], [33], [34], [35]. One of the main disadvantages of these techniques is the imbalance of classes to be segmented. Therefore, these techniques use data augmentation to improve their final performance to avoid that issue.

C. METHODS BASED ON ARMPIT EXCLUSION

Pérez et al. [16] proposed an armpit and pectoral muscle removal method. However, they did not differentiate between each exclusion process, but they mentioned that their method is defined by a robust procedure founded on negative gradient changes. In this paper, Perez et al. point out that removing the abdominal fat can improve their method's performance. However, no preprocessing task presented in the literature is focused on eliminating that unwanted anatomical component.

How we can see the delimitation of the pectoral muscle and breast tissue are two challenging tasks. However, other artifacts may affect the performance of a method [16].

III. PROPOSED APPROACH

Fig. 2 shows the proposed processing flow performed by the fat-removal-based pipeline method for digital mammograms.

The first block is for breast delineation. This block includes denoising using a median filter, followed by min-max normalization to translate the range of gray-level values to a smaller scale, [0,1]. Also, a Gaussian filter for smooth edges follows Otsu's method with a tolerance value established for digital mammograms to 0.1 under trial and error to obtain the breast segmentation with smooth edges. Finally, a gap-filling algorithm and a method for removing minor disconnected areas are applied to obtain a binary mask, BW.

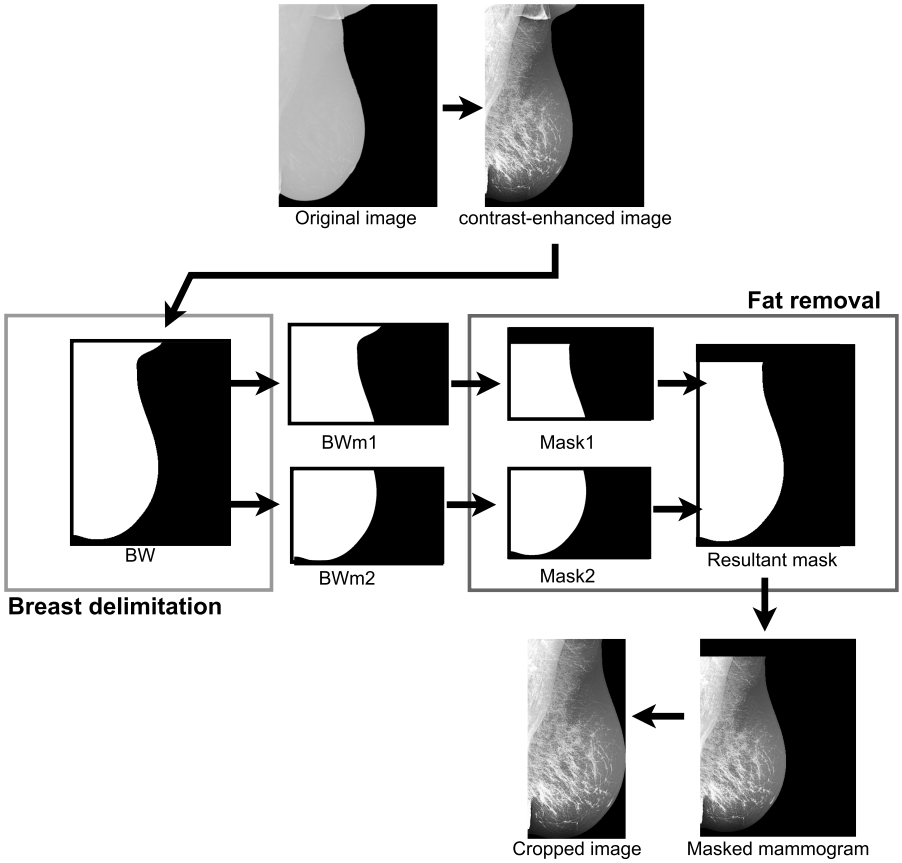

FIGURE 2. Processing flow performed for the proposed method for fat removal.

The second block is the fat removal method. This block removes the axillary and abdominal fat. The fat removal block receives the BW mask in two halved segments, BWm1 and BWm2, corresponding to the upper and lower part of the mask. Subsequently, a contour-edge method is calculated in each BWm1 and BWm2 area to obtain their coordinates (x,y). Then, the original coordinates are converted to quadratic coordinates (x^2,y^2) to obtain a contour-edge signature of the breast ($r = x^2 + y^2$) to get the minimum point in the function in a more robust way; Fig. 3 shows the difference between using the original signature coordinates and quadratic coordinates.

The minima of r are obtained using a peak detection function in Matlab (findpeaks). Then, we select the cut points (Ps) in the upper and lower segments. The average minimum point is the value chosen as the cut point (Ps) in the upper segment. The cutoff point (Ps) established in the lower segment is the last minimum point found in the quadratic function. Once the Ps is selected in the quadratic function r, the coordinates of Ps are located in the Cartesian plane to generate a cut line that goes from Ps(x,y) to Ps(0,y) in BWm1 and BWm2, respectively. The cut line is generated in both segments, then a connected area detection algorithm is applied, and the segment with the largest area in both BWm1 and BWm2 is selected. To ensure that the findpeaks method does not find a local minimum in BWm1, we considered a constraint; the region with the maximum area should be above the half area in BWm1; if the condition is not satisfied, no area is taken. Fig. 4 shows in red the minimum point found for the findpeaks function in a mammogram and with a dotted line the halfway point of the BWm1 area. At the end of the process, BWm1 and BWm2 are combined, and the BW mask information is updated. Finally, the grayscale mammogram is masked with the obtained binary mask to get the preprocessed mammogram without fat information.

Notice that the findpeaks function in Matlab returns a vector with local maxima points of an input signal vector. So, the r signal needs to be inverted to make the minimum points maxima.

Algorithm 1 shows the pseudo-code of the fat removal block.

IV. METHODS AND MATERIALS

This section is divided into four subsections. The first describes the datasets used in the evaluation of the proposed approach. The second describes the experimentation carried out to validate the use of the proposed method. The third describes the performance index used. Finally, the fourth subsection describes the software used.

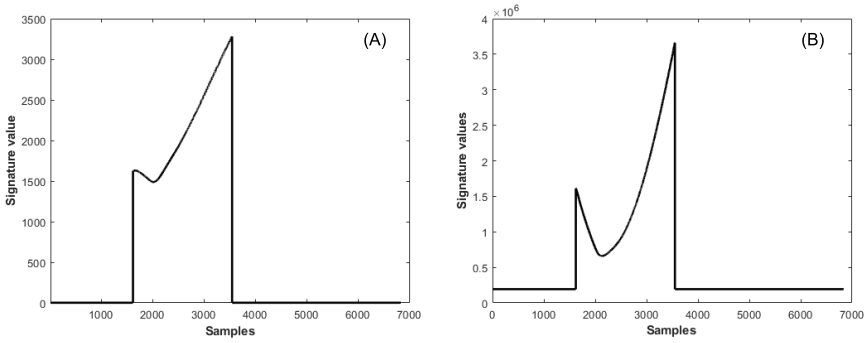

FIGURE 3. Original contour signature (A), and contour signature when using quadratic coordinates (B).

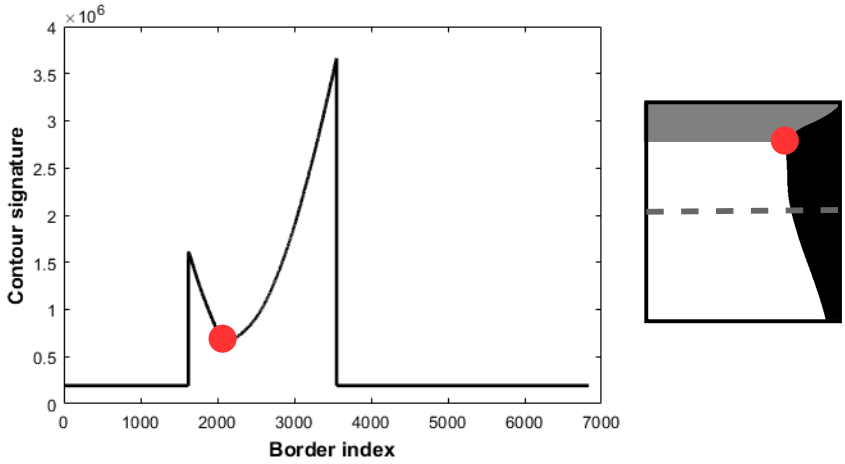

FIGURE 4. Minimum point found by the findpeaks function.

A. DATABASES

The mammograms used in this study were obtained from a hospital in San Luis Potosi, Mexico. The Private mammograms include a total of 256 mammograms with mediolateral oblique views.

Additionally, we used 337 mammograms from public databases from a hospital in Portugal. In total, 207 mammograms from the INbreast database [36] and 130 mammograms from the BCDR database [37] were selected.

B. EVALUATION OF THE PROPOSED METHOD

To measure the proposed methods' effectiveness, we evaluate the proposed preprocessing techniques following the subsequent considerations:

A comparison between the segmentation results and the manual segmentation was carried out to evaluate the performance of the breast delimitation stage. A human performed the manual segmentation, corroborated by an oncologist, using the software VIA .[1] The Jaccard and Dice index was used to do a quantitative evaluation.

Additionally, the breast delimitation method was compared with two methods proposed by Zebari et al. [14], and Taghanaki et al. [2].

On the other hand, to evaluate the performance of the fat removal method, a qualitative comparison was made. According to [38], the classes considered in the comparison were suitable, acceptable, and unacceptable. This method is not compared to others because similar methods do not exist in the literature.

Additionally, to see if it is convenient to apply the fat removal method before pectoral muscle removal. We performed complementary experimentation using a semantic segmentation method using a SegNet architecture for pectoral muscle delimitation. We selected this technique because there are recent methods used to remove pectoral muscles [39], so we found an opportunity to propose a new method for pectoral muscle removal in mammography.

SegNet is a deep convolutional encoder-decoder architecture for semantic segmentation. This segmentation network consists of an encoder and a decoder for pixel-wise classification. The SegNet architecture includes the 13 convolutional layers and is topologically identical to the VGG16. The role of the decoder is to map the low-resolution encoder feature

[1] https://www.robots.ox.ac.uk/~vgg/software/via/

Algorithm 1 Fat Removal Algorithm

Input : Mammography and binary mask (BW)
Output : Preprocessed mammography

1 Divide BW into two segments, upper BWm1, and lower BWm2.
2 For each segment
3 Calculate the contour in BW.
4 Obtain the Cartesian coordinates (x,y) from the contour.
5 Convert the Cartesian coordinates (x,y) to quadratic coordinates (x^2,y^2).
6 Find the minimum points (Ps) in the edge function $r=x^2+y^2$
7 For BWm1, select the minimal point (P) as the midpoint in Ps.
8 For BWm2, select P as the last minimum value in Ps.
9 Generate the cut area from P(x,y) to P(0,y).
10 Apply the connected component detection algorithm.
11 Select the largest connected area in BWm1 and BWm2.
12 Merge the resulting masks (BWm1 and BWm2) into BW.
13 Mask the mammography with the binary mask obtained.

maps to full input resolution feature maps for pixel-wise classification. The fully connected layers are tuned during the training process to produce the class probabilities for each pixel independently using the new class label [40], [41].

The SegNet model is shown in Fig. 5. The result image label obtained by SegNet is shown in Fig. 6. In red is the background of the mammogram, in blue is the pectoral muscle, and in gray is the breast tissue, according to their class label. The background was labeled as class 0, the breast region was labeled as class 1, and the pectoral muscle region was labeled as class 2.

The SegNet model was trained using the BCDR, INbreast, and Private datasets. Each dataset was evaluated with the original mammograms and preprocessed mammograms using the fat removal method. Resulting in the following variants:

- S-O-BCDR: SegNet trained with the BCDR dataset using the original images without preprocessing.
- S-P-BCDR: SegNet trained with the BCDR dataset using the images preprocessed with the fat removal method.
- S-O-INbreast: SegNet trained with the INbreast dataset using the original images without preprocessing.
- S-P-INbreast: SegNet trained with the INbreast dataset using the images preprocessed with the fat removal method.
- S-O-Private: SegNet trained with the Private dataset using the original images without preprocessing.
- S-P-Private: SegNet trained with the Private dataset using the images preprocessed with the fat removal method.

The SegNet settings are learning rate $1e{-}3$, mini-batch size 4, momentum 0.9, L2Regularization 0.0005, the maximum number of epochs 260, shuffle every epoch, and training optimizer SGDM. Also, we use cross-entropy loss as a loss function. To avoid overfitting, we use data augmentation with random reflection, random translation in the range of [−10,10], and random rotation in the range of [−180, 180] in both the x and y axes. Also, all images were resized to 256 × 256 RGB images and converted to PNG format to be consistent for the semantic segmentation method. The network input receives normalized images in the range of [0,1]. To do this, the gray level is divided by 255.

All SegNet variants were performed using a cross-validation method with ten folds. The results were compared to the original and preprocessed mammography segmentation using the Accuracy and IoU index. In addition, the Kruskal-Wallis and Bonferroni statistical tests were applied.

C. PERFORMANCE INDEXES

According to the ground truth, accuracy is the ratio of correctly classified pixels to the total number of pixels in the class. The range of accuracy is between [0,1]. Equation (1) shows the calculated Accuracy [42].

$$\text{Accuracy} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{1}$$

TP is the number of true positives and *FN* is the number of false negatives.

Mean accuracy is the average accuracy of all classes in all images in a particular dataset.

Global accuracy is the ratio of correctly classified pixels to the total number of pixels in the mammogram. The global accuracy metric estimates the percentage of correctly classified pixels.

Intersection over union (IoU) is the ratio of correctly classified pixels to the total number of ground truth and predicted pixels in that class. IoU is the Jaccard similarity coefficient and is the most commonly used metric for semantic segmentation. The IoU metric measures statistical accuracy by penalizing false positives. The range in which it is found is between [0,1]. Equation (2) shows the calculated IoU [42].

$$\text{IoU} = \frac{\text{TP}}{(\text{TP} + \text{FP} + \text{FN})} \tag{2}$$

TP is the true positives, *FP* is the false positives, and *FN* is the false negatives.

Mean IoU is the average IoU score of all classes in a particular dataset.

Weighted IoU is the average IoU of each class weighted by the number of pixels in that class. This metric is appropriate for images with disproportionately sized classes to reduce the impact of errors in the small classes.

The Dice index is related to the Jaccard index according to (3).

$$\text{Dice} = 2 * \text{IoU}/(1 + \text{IoU}) \tag{3}$$

Additionally, the Kruskal-Wallis statistical test and the Bonferroni post-hoc test were used.

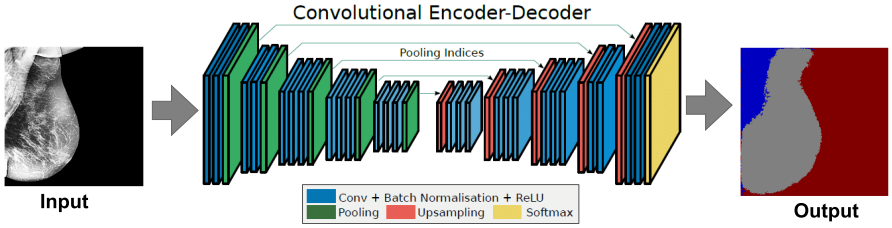

FIGURE 5. SegNet model trained to segment the background, breast tissue, and pectoral muscle on mammography.

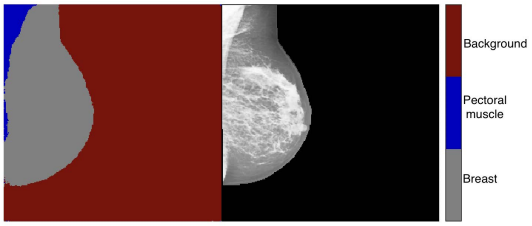

FIGURE 6. Example of the response labeling generated by the SegNet method.

D. SOFTWARE

The experimentation was carried out on a Windows 7 computer with an Intel Core i7 processor and 32 GB RAM. In addition, the computer has an NVIDIA GeForce 1080 Ti. All programs were developed in Matlab R2019a.

V. EXPERIMENTAL RESULTS AND DISCUSSION

This section includes the results of the proposed method and the results obtained from complementary experimentation using a SegNet architecture for pectoral muscle removal. As we introduced in Section IV-B, this additional experiment aims to see if it is convenient to apply the fat removal technique before or after the pectoral muscle removal.

A. PROPOSED METHOD

Table 1 shows the median and median absolute deviation (MAD) of Jaccard and Dice obtained for the three methods, the proposed method and the breast delimitation methods of Zebari et al. [14], and Taghanaki et al. [2], during the evaluation of the BCDR, INbreast, and Private datasets.

TABLE 1. Median and MAD performance of the breast tissue delimitation method.

Dataset	Method	Jaccard	Dice	p-value
BCDR	Proposed method	0.948 ± 0.014	0.972 ± 0.007	< 0.001
	Zebari et al. [14]	0.916 ± 0.012	0.955 ± 0.006	
	Taghanaki et al. [2]	$\mathbf{0.976 \pm 0.005}$	$\mathbf{0.988 \pm 0.002}$	
INbreast	Proposed method	0.956 ± 0.010	0.977 ± 0.005	< 0.001
	Zebari et al. [14]	0.945 ± 0.006	0.972 ± 0.003	
	Taghanaki et al. [2]	$\mathbf{0.983 \pm 0.003}$	$\mathbf{0.991 \pm 0.001}$	
Private	**Proposed method**	$\mathbf{0.944 \pm 0.012}$	$\mathbf{0.971 \pm 0.006}$	< 0.001
	Zebari et al. [14]	0.873 ± 0.010	0.922 ± 0.005	
	Taghanaki et al. [2]	0.812 ± 0.008	0.856 ± 0.004	

The results show that the proposed method is competitive with the methods in the literature. However, the methods obtain different results because mammograms from public and private banks, coming from different scanners, vary in dynamic range. The method of Taghanaki et al. [2] for example, obtains the best segmentation performance with public datasets; however, when segmenting mammograms from the Private dataset, it has the lowest performance. On the other hand, the proposed method adequately segments the set of mammograms of Mexican patients while ranking second when segmenting public datasets. In general, the segmentation performance of the proposed method is above 94%, according to the Jaccard and Dice index.

Fig. 7 shows the visual results of the fat removal proposed method used on four images from the three datasets: BCDR, INbreast, and Private.

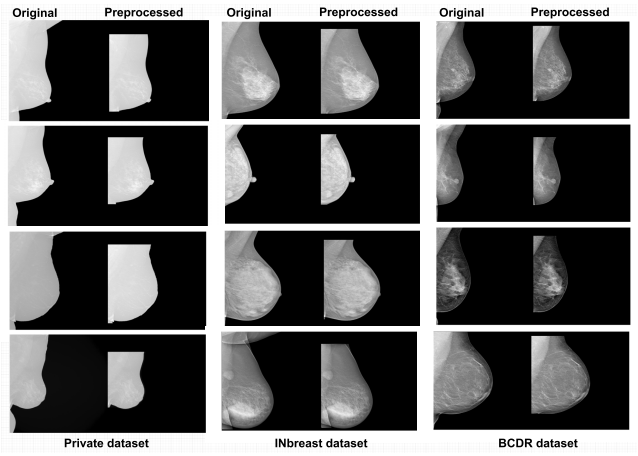

FIGURE 7. Results obtained from the proposed axillary and abdominal fat removal preprocessing technique.

The results showed mammograms included in the datasets and the preprocessed images obtained for the proposed method. A person, supervised by an oncologist, reviewed the preprocessed images and considered if the method was suitable for digital mammogram preprocessing. First, however, a qualitative evaluation was necessary. Table 2 shows the performance achieved by the fat removal method considering the suitable, acceptable, and unacceptable labels. The results show that 94.18% of the cases were suitable, 4.32% acceptable, and 1.50% unacceptable.

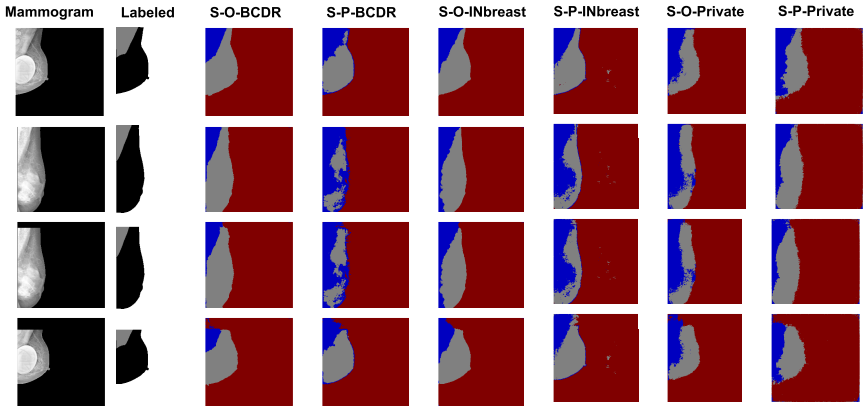

FIGURE 8. Segmentation results were obtained by SegNet when evaluating four mammograms. The first column (Mammogram) shows the original and preprocessed images obtained from two BCDR mammograms. The second column (Labeled) shows the manual delimitation performed by a human. Finally, the following columns show the results of the SegNet variants trained previously.

TABLE 2. Number of cases classified as suitable, acceptable, unacceptable labels.

Database	Suitable	Acceptable	Unacceptable
BCDR	119	8	3
INbreast	197	8	2
Private	250	10	4
Total	566 (94.18%)	26 (4.32%)	9 (1.50%)

B. PECTORAL MUSCLE REMOVAL METHODS

Table 3 shows the average results obtained for the semantic segmentation methods using a SegNet architecture.

The results of the statistical test of Kruskal-Wallis was a value $p<0.001$. The pairwise comparison results obtained by the Bonferroni method are shown in Table 4.

The results show that the variants trained with the BCDR and INbreast datasets present a significant difference ($p<0.05$) in all the semantic segmentation performance indices used. Furthermore, there is a difference between the variants trained with original images and those preprocessed using the fat removal method. The best results were obtained when training the network with the original images without preprocessing.

In the case of the variants trained with mammograms in the Private dataset, we can only observe a statistically significant difference in the mean accuracy index, but a statistically similar performance in other indexes. Therefore, the best variant is the S-P-Private that uses preprocessed images.

Table 5 shows the accuracy, and IoU, per class: pectoral muscle, breast, and background. The results per class show that in terms of accuracy, the pectoral, breast, and background classes reach values above 90% accuracy for all variants trained with BCDR and INbreast mammograms. For the S-O-Private and S-P-Private variants, the highest classification performance in terms of accuracy is for the S-P-Private variant, which considers the removal of axillary and abdominal fat before removing the pectoral muscle.

The breast and background classes reach higher segmentation performances than the pectoral class in the IoU index. The pectoral class performs better with the variants trained with INbreast and BCDR. On the other hand, the variants trained with the Private dataset reached 77 and 75% for the S-O-Private and S-P-Private variants, respectively.

From these results, we point out that for the INbreast and BCDR databases, it is not necessary to apply the fat removal method before removing the pectoral muscle. This is because axillary artifacts do not significantly influence preprocessing improvement. On the other hand, in the Private dataset, as there are more significant variations of undesired artifacts in the images, such as the arm and the axillary basin, preprocessing of fat removal is necessary. Therefore, removing fat will depend on how well the mammograms were taken, the position of the patient during the mammography, or directly related to the patient's complexion. In this case, the Private mammograms analyzed show more fat than the mammograms in the INbreast and BCDR datasets. This phenomenon may be directly related to the analyzed population.

Table 6 shows a comparison of the results obtained in contrast to other recently published semantic segmentation methods.

The results shown in Table 6 show that the performances achieved by the S-O-BCDR, S-O-INbreast, and S-P-Private methods are competitive with those of the literature.

In addition, we performed a generalization study to observe the generalization power of semantic segmentation methods using SegNet when using different datasets to observe the improvements and limitations of the methods because this kind of study is not usually done in the literature. The following subsection shows this kind of analysis done on the S-O-BCDR, S-O-INbreast, and S-P-Private proposed methods.

1) GENERALIZATION ANALYSIS OF SEMANTIC SEGMENTATION METHODS

A generalization analysis was considered to evaluate the generalization power of the trained variants. In this generalization study, the SegNet model was trained with one dataset and tested with the rest.

TABLE 3. General results of accuracy and IoU of the different variants analyzed. In bold the highest median results by dataset type.

Method	Global Accuracy	Mean Accuracy	Mean IoU	Weighted IoU
S-O-BCDR	**0.985 ± 0.002**	**0.980 ± 0.004**	**0.927 ± 0.016**	**0.971 ± 0.004**
S-P-BCDR	0.973 ± 0.003	0.967 ± 0.008	0.911 ± 0.009	0.950 ± 0.016
S-O-INbreast	**0.986 ± 0.0005**	**0.981 ± 0.003**	**0.925 ± 0.006**	**0.974 ± 0.0009**
S-P-INbreast	0.975 ± 0.003	0.968 ± 0.010	0.915 ± 0.008	0.954 ± 0.006
S-O-Private	0.968 ± 0.006	0.955 ± 0.006	**0.880 ± 0.021**	0.942 ± 0.009
S-P-Private	**0.973 ± 0.003**	**0.965 ± 0.001**	0.878 ± 0.014	**0.951 ± 0.006**

TABLE 4. P-value obtained from Bonferroni statistical test for an α of 95% confidence. In bold are the comparisons where there is a significant statistical difference.

Method	p value	Global Accuracy	Mean Accuracy	Mean IoU	Weighted IoU
S-O-BCDR	S-P-BCDR	**p<0.001**	**p<0.001**	**p=0.007**	**p<0.001**
S-O-INbreast	S-P-INbreast	**p<0.001**	**p=0.003**	**p=0.012**	**p<0.001**
S-O-Private	S-P-Private	p=0.162	**p=0.017**	p=0.850	p=0.140

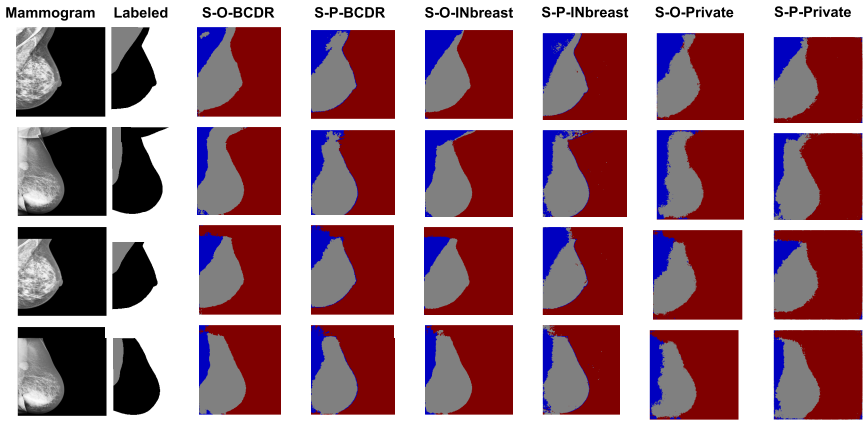

FIGURE 9. Segmentation results were obtained for SegNet when evaluating four mammograms. The first column (Mammogram) shows the original and preprocessed images obtained from two INbreast mammograms. The second column (Labeled) shows the manual delimitation performed by a human. Finally, the following columns show the results of the SegNet variants trained previously.

TABLE 5. Results of Accuracy and IoU per class of the different variants analyzed.

Class	Variant	Accuracy	IoU
Pectoral	S-O-BCDR	0.976	0.835
	S-P-BCDR	0.949	0.798
	S-O-INbreast	0.975	0.821
	S-P-INbreast	0.943	0.800
	S-O-Private	0.943	0.773
	S-P-Private	0.958	0.753
Breast	S-O-BCDR	0.975	0.957
	S-P-BCDR	0.975	0.967
	S-O-INbreast	0.977	0.964
	S-P-INbreast	0.976	0.968
	S-O-Private	0.941	0.890
	S-P-Private	0.960	0.906
Background	S-O-BCDR	0.989	0.988
	S-P-BCDR	0.979	0.970
	S-O-INbreast	0.991	0.990
	S-P-INbreast	0.985	0.978
	S-O-Private	0.981	0.978
	S-P-Private	0.978	0.976

Tables 7, 8, and 9 show the quantitative results of central tendency and dispersion achieved by the variants S-O-BCDR, S-O-INbreast, and S-P-Private during the generalization tests.

Table 7 shows the median and median absolute deviation (MAD) results of the S-O-BCDR method when evaluated with different datasets with and without the proposed fat removal preprocessing. The results are shown by class; breast, pectoral, and background. For the pectoral class, which is the one that concerns us, the S-O-BCDR method has an adequate generalization when evaluating the INbreast dataset without preprocessing for the rest of the datasets; its discrimination power is lower.

Fig. 8 shows the segmentation performance obtained by the six variants of SegNet when tested with two BCDR mammograms. The first two mammograms are without preprocessing, and the next two are preprocessed using the proposed fat-removal method. We can see that the S-O-BCDR method is the variant that achieves the best segmentation, followed by the S-O-INbreast variant; the rest of the variants over-segment the pectoral muscle.

Table 8 shows the generalization results of the S-O-INbreast method. The S-O-INbreast method achieves 86 ± 0.004% IoU for the BCDR dataset without preprocessing. However, for other datasets, it reaches a performance between 54-66%.

TABLE 6. Comparison between methods using the mean accuracy and IoU.

Method	Network	Dataset	Mean Accuracy	Mean IoU
Ali et al. (2020) [31]	Unet	INbreast	0.879 ± 0.045	0.95 ± 0.031
Xiang et al. (2022) [32]	PeMNet	INbreast & OPTIMAM	0.9746 ± 0.004	0.994 ± 0.029
Zhou et al. (2022) [33]	Deeplabv3+	INbreast	0.991	-
Leite et al. (2021) [34]	Mobilenet-unet	INbreast & MIAS	-	0.945 ± 0.073
S-O-BCDR proposed method	SegNet	BCDR	0.980 ± 0.004	0.927 ± 0.016
S-O-INbreast proposed method	SegNet	INbreast	0.981 ± 0.003	0.925 ± 0.006
S-P-Private proposed method	SegNet	Private	0.965 ± 0.001	0.878 ± 0.014

TABLE 7. Generalization results of the S-O-BCDR method when evaluating different datasets. The median and median absolute deviation (MAD) from the IoU index were obtained per class.

Class	Dataset	Median	MAD
Breast	Original BCDR	-	-
	BCDR preprocessed	0.955	0.008
	Original INbreast	0.961	0.010
	INbreast preprocessed	0.952	0.011
	Original Private	0.833	0.053
	Private preprocessed	0.868	0.053
Pectoral	Original BCDR	-	-
	BCDR preprocessed	0.648	0.083
	Original INbreast	0.831	0.064
	INbreast preprocessed	0.649	0.090
	Original Private	0.594	0.150
	Private preprocessed	0.457	0.159
Background	Original BCDR	-	-
	BCDR preprocessed	0.970	0.006
	Original INbreast	0.989	0.003
	INbreast preprocessed	0.968	0.007
	Original Private	0.984	0.003
	Private preprocessed	0.972	0.008

TABLE 8. Generalization results of method S-O-INbreast when evaluating different datasets. The median and median absolute deviation (MAD) from the IoU index were obtained per class.

Class	Dataset	Median	MAD
Breast	Original BCDR	0.967	0.010
	BCDR preprocessed	0.959	0.010
	Original INbreast	-	-
	INbreast preprocessed	0.962	0.008
	Original Private	0.853	0.047
	Private preprocessed	0.904	0.039
Pectoral	Original BCDR	0.864	0.049
	BCDR preprocessed	0.632	0.095
	Original INbreast	-	-
	INbreast preprocessed	0.667	0.075
	Original Private	0.647	0.113
	Private preprocessed	0.540	0.093
Background	Original BCDR	0.992	0.002
	BCDR preprocessed	0.971	0.005
	Original INbreast	-	-
	INbreast preprocessed	0.972	0.007
	Original Private	0.989	0.002
	Private preprocessed	0.972	0.008

Fig. 9 shows the segmentation performances obtained by the six SegNet variants when tested with two INbreast mammograms, with and without preprocessing. The variants with the best segmentation performance continue to be the S-O-BCDR and S-O-INbreast methods. The rest of the variants are over-segmented or under-segmented in most cases. The second image is complicated because it does not consider the arm as a part of the breast or pectoral muscle in the labeled image, and the methods find that the arm does not pertain to the background. Hence, some methods assign it to the pectoral or breast classes.

TABLE 9. Generalization results of method S-P-Private when evaluating different datasets. The median and median absolute deviation (MAD) from the IoU index were obtained per class.

Class	Dataset	Median	MAD
Breast	Original BCDR	0.881	0.025
	BCDR preprocessed	0.910	0.017
	Original INbreast	0.908	0.022
	INbreast preprocessed	0.911	0.017
	Original Private	0.887	0.024
	Private preprocessed	-	-
Pectoral	Original BCDR	0.713	0.089
	BCDR preprocessed	0.626	0.103
	Original INbreast	0.691	0.126
	INbreast preprocessed	0.630	0.108
	Original Private	0.804	0.059
	Private preprocessed	-	-
Background	Original BCDR	0.960	0.008
	BCDR preprocessed	0.966	0.009
	Original INbreast	0.974	0.006
	INbreast preprocessed	0.963	0.009
	Original Private	0.977	0.004
	Private preprocessed	-	-

In the case of the S-P-Private method, Table 9 shows the generalization performance of the method. The approach adequately generalizes with 71% IoU to the BCDR dataset, marginally with 69% IoU to the INbreast dataset, and accurately with 80% IoU to the Private dataset. The results suggest that applying fat removal preprocessing to the mammography bank of Mexican women is effective and it is capable of generalizing datasets that have not been preprocessed, as is the case with the INbreast, BCDR, and Private datasets.

Fig. 10 shows the segmentation performances obtained by the six SegNet variants tested with three Private mammograms. The first three mammograms are without preprocessing, and the next three are preprocessed using the proposed method. The preprocessed mammograms were correctly segmented by the S-O-BCDR, S-O-INbreast, and S-P-Private methods. On the other hand, the original mammograms were only correctly segmented by the S-O-Private method for these examples. These results suggest that it is difficult to obtain a technique that adequately generalizes with all

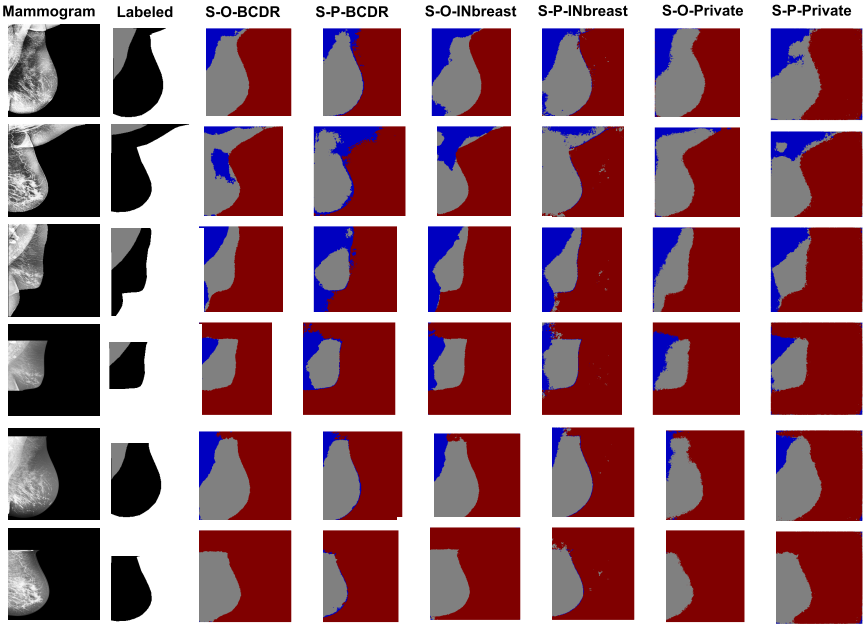

FIGURE 10. Segmentation results were obtained for SegNet when evaluating four mammograms. The first column (Mammogram) shows the original and preprocessed images obtained from two Private mammograms. The second column (Labeled) shows the manual delimitation performed by a human. Finally, the following columns show the results of the SegNet variants trained previously.

TABLE 10. Average IoU performance obtained by the S-O-BCDR, S-O-INbreast, and S-P-Private methods when evaluating different datasets.

Method	Training	Test	IoU
S-O-BCDR	BCDR	**BCDR**	**0.927 ± 0.016**
		BCDR preprocessed	0.857 ± 0.032
		INbreast	0.924 ± 0.026
		INbreast preprocessed	0.856 ± 0.036
		Private	0.804 ± 0.069
		Private preprocessed	0.766 ± 0.073
S-O-INbreast	INbreast	BCDR	0.941 ± 0.020
		BCDR preprocessed	0.854 ± 0.037
		INbreast	**0.925 ± 0.006**
		INbreast preprocessed	0.867 ± 0.030
		Private	0.833 ± 0.054
		Private preprocessed	0.805 ± 0.047
S-P-Private	Private preprocessed	BCDR	0.873 ± 0.041
		BCDR preprocessed	0.834 ± 0.043
		INbreast	0.858 ± 0.051
		INbreast preprocessed	0.831 ± 0.045
		Private	0.890 ± 0.029
		Private preprocessed	**0.878 ± 0.014**

datasets, particularly when they do not follow the same pattern with which they were trained. For example, although the S-O-BCDR and S-O-INbreast methods seem to generalize the mammograms accurately from the INbreast and BCDR datasets, they did not get the same results using the Private datasets without mammogram preprocessing. This behavior is because the Private dataset contains a more significant amount of tissue such as the arm, pectoral muscle, armpit basin, and abdominal fat. As a result, a high-intensity contrast is generated after applying a contrast enhancement technique. In this way, the tissue's high intensity and variety can confuse the methods and, therefore, result in over-segmentation, as we can see with the S-O-BCDR, S-O-INbreast, and S-P-Private methods when evaluating these mammograms.

The overall performance of the S-O-BCDR, S-O-INbreast, and S-P-Private methods, in terms of the IoU index, is shown in Table 10. Bold indicates the method's performance after being trained with the testing set from the same dataset. Note

that each method has been evaluated with different datasets to observe the generalization performance.

The S-O-BCDR method reaches 92% IoU after being tested with images from the BCDR dataset. Furthermore, it achieves a similar performance of 92% with INbreast mammograms. On the other hand, the S-O-INbreast method reaches 92% when tested with the INbreast dataset and performs 94% using the BCDR dataset. Contrastingly, the S-P-Private method achieves 87% IoU when tested with preprocessed Mexican mammograms, and it reaches 87%, 85%, and 89% when tested with the BCDR, INbreast, and Private datasets, respectively.

VI. CONCLUSION AND FUTURE DIRECTIONS

In this paper, a pipeline for mammography preprocessing was presented. This pipeline includes a block for breast delimitation and a method for abdominal and axillary fat removal.

The fat removal proposed method originated from a need found intrinsically in the mammograms of Mexican patients (the Private dataset); the objective was to leave the area of breast tissue as clean as possible for posterior analysis. The method was also evaluated using other public digital mammograms. The BCDR and INbreast datasets are composed of mammograms from Portuguese patients.

The proposed method is fully automatic and requires no interaction from the user. In addition, it is suitable for breast delimitation obtained above 94% according to the Jaccard and Dice index for the three datasets evaluated. Furthermore, the fat removal method suitably removes the axillary and abdominal fat in 94% of cases. Therefore, the proposed method allows obtaining a cleaner area of the breast, which may be helpful for classification tasks.

Furthermore, a comparative study was carried out using a SegNet to verify whether applying the fat removal method before pectoral muscle removal would be convenient. The results showed that in the case of public mammograms, since these are of patients who do not present significant variations in fatty tissue, they do not require the removal of fat before detecting the pectoral muscle. However, after the pectoral muscle removal, the proposed fat removal method can be used to improve other kinds of analysis, such as density analysis. On the other hand, for mammograms with significant fat variations like in the Private dataset that we evaluated, we found it is effective to apply the fat removal method before removing the pectoral muscle to get a cleaner breast area. Therefore, we conclude that a preprocessing technique like we propose can be beneficial depending on the population analyzed and its features.

Also, we did a generalization analysis to test the generalization capacity of the proposed semantic segmentation methods: S-P-Private, S-O-BCDR, and S-O-INbreast. This study was necessary to evaluate the effectiveness of the semantic segmentation methods. Unfortunately, the generalization study is not usually carried out in the literature. Still, it was essential to know the scope and limitations of the method and observe the generalization power of the SegNet methods. Among the results obtained, the S-P-Private method, which was trained with mammograms of Mexican women preprocessed by the proposed fat removal method, was able to generalize the pectoral muscle segmentation in the INbreast, BCDR, and Private datasets with a 69%, 71%, and 80% IoU index, respectively. On the other hand, the S-O-BCDR and S-O-INbreast methods did not adequately segment the pectoral muscles using the Private dataset; their performance was less than 65% of the IoU index. This behavior is because both datasets come from a hospital in Portugal. Therefore the population analyzed is similar and differs from the Mexican people.

We concluded that datasets that include unwanted tissues (e.g., abdominal and axillary fat, and the arm) in a small proportion, as in the INbreast and BCDR datasets, current methods can remove the pectoral muscles without fat removal. However, for mammograms with unwanted tissue covering a large proportion of the area, as in the Private dataset, we recommend that the proposed fat removal method be used before detecting the pectoral muscle to remove structures such as the arm, which can be confused with the pectoral muscle.

The S-O-BCDR, S-O-INbreast, and S-P-Private methods obtain an overall segmentation performance in three classes (breast, pectoral muscle, and background) of 92.7%, 92.5%, and 87.8%, respectively, according to the mean of the IoU index. In terms of average accuracy, these methods obtain 98.0%, 98.1%, and 96.5%, respectively. These methods were compared with other recent semantic segmentation methods. We determined that the proposed methods are competitive with other recent approaches in the literature.

In future work, we desire to evaluate the fat removal method on film-based mammograms because, in this research work, only digital mammograms were used. In addition, we will aspire to improve the pectoral muscle segmentation method by evaluating other semantic segmentation networks and applying an object detection method to avoid over-segmentation problems. Finally, we want to analyze the mammographic density performance using mammograms of Mexican patients preprocessed with the proposed method.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] P. S. Vikhe and V. R. Thool, "Intensity based automatic boundary identification of pectoral muscle in mammograms," *Proc. Comput. Sci.*, vol. 79, pp. 262–269, Jan. 2016.

[2] S. A. Taghanaki, Y. Liu, B. Miles, and G. Hamarneh, "Geometry-based pectoral muscle segmentation from MLO mammogram views," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 11, pp. 2662–2671, Nov. 2017.

[3] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA, Cancer J. Clinicians*, vol. 55, pp. 74–108, Apr. 2005.

[4] S. M. Shah, R. A. Khan, S. Arif, and U. Sajid, "Artificial intelligence for breast cancer analysis: Trends & directions," *Comput. Biol. Med.*, vol. 142, Mar. 2022, Art. no. 105221.

[5] M. Oladele, A. Adekiigbe, and T. Adepoju, "A two-stage model for breast cancer detection and classification from mammograms," *SSRN Electron. J.*, pp. 1–13, Mar. 2022.

[6] S. Sreedevi and E. Sherly, "A novel approach for removal of pectoral muscles in digital mammogram," *Proc. Comput. Sci.*, vol. 46, pp. 1724–1731, Jan. 2015.
[7] P. Haria, M. Thakur, P. Popat, A. Katdare, and S. Chauhan, "Breast imaging: Past and present," *Indian J. Med. Paediatric Oncol.*, vol. 43, pp. 1–5, Feb. 2022.
[8] C. K. Glide-Hurst, N. Duric, and P. Littrup, "A new method for quantitative analysis of mammographic density," *Med. Phys.*, vol. 34, no. 11, pp. 4491–4498, Nov. 2007.
[9] S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel, and M. T. Rickard, "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," *IEEE Trans. Med. Imag.*, vol. 23, no. 9, pp. 1129–1140, Sep. 2004.
[10] S. Bacha and O. Taouali, "A novel machine learning approach for breast cancer diagnosis," *Measurement*, vol. 187, Jan. 2022, Art. no. 110233.
[11] A. Marchesi, A. Bria, C. Marrocco, M. Molinara, J.-J. Mordang, F. Tortorella, and N. Karssemeijer, "The effect of mammogram preprocessing on microcalcification detection with convolutional neural networks," in *Proc. IEEE 30th Int. Symp. Computer-Based Med. Syst. (CBMS)*, Jun. 2017, pp. 207–212.
[12] S. J. S. Gardezi, F. Adjed, I. Faye, N. Kamel, and M. M. Eltoukhy, "Segmentation of pectoral muscle using the adaptive gamma corrections," *Multimedia Tools Appl.*, vol. 77, no. 3, pp. 3919–3940, Feb. 2018.
[13] H. Ture and T. Kayikcioglu, "Accurate detection of distorted pectoral muscle in mammograms using specific patterned isocontours," *IEEE Access*, vol. 8, pp. 147370–147386, 2020.
[14] D. A. Zebari, D. Q. Zeebaree, A. M. Abdulazeez, H. Haron, and H. N. A. Hamed, "Improved threshold based and trainable fully automated segmentation for breast cancer boundary and pectoral muscle in mammogram images," *IEEE Access*, vol. 8, pp. 203097–203116, 2020.
[15] A. R. Beeravolu, S. Azam, M. Jonkman, B. Shanmugam, K. Kannoorpatti, and A. Anwar, "Preprocessing of breast cancer images to create datasets for deep-CNN," *IEEE Access*, vol. 9, pp. 33438–33463, 2021.
[16] F. J. Pérez-Benito, F. Signol, J.-C. Perez-Cortes, A. Fuster-Baggetto, M. Pollan, B. Pérez-Gómez, D. Salas-Trejo, M. Casals, I. Martínez, and R. Llobet, "A deep learning system to obtain the optimal parameters for a threshold-based breast and dense tissue segmentation," *Comput. Methods Programs Biomed.*, vol. 195, Oct. 2020, Art. no. 105668.
[17] V. De la Cruz-Góngora, E. Chiquete, H. Gómez-Dantés, L. Cahuana-Hurtado, and C. Cantú-Brito, "Trends in the burden of stroke in Mexico: A national and subnational analysis of the global burden of disease 1990–2019," *Lancet Regional Health*, vol. 10, pp. 682–692, Jun. 2022.
[18] P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangeri, M. L. Pepe, and R. M. Rangayyan, "Estimation of the breast skin-line in mammograms using multidirectional Gabor filters," *Comput. Biol. Med.*, vol. 43, no. 11, pp. 1870–1881, Nov. 2013.
[19] A. Rampun, P. J. Morrow, B. W. Scotney, and J. Winder, "Fully automated breast boundary and pectoral muscle segmentation in mammograms," *Artif. Intell. Med.*, vol. 79, pp. 28–41, Jun. 2017.
[20] G. Toz and P. Erdogmus, "A novel hybrid image segmentation method for detection of suspicious regions in mammograms based on adaptive multithresholding (HCOW)," *IEEE Access*, vol. 9, pp. 85377–85391, 2021.
[21] A. Rampun, K. López-Linares, P. J. Morrow, B. W. Scotney, H. Wang, I. G. Ocaña, G. Maclair, R. Zwiggelaar, M. A. González Ballester, and I. Macía, "Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network," *Med. Image Anal.*, vol. 57, pp. 1–17, Oct. 2019.
[22] S. S. Fadhil and F. A. A. Dawood, "Automatic pectoral muscles detection and removal in mammogram images," *Iraqi J. Sci.*, vol. 62, pp. 676–688, Feb. 2021.
[23] S. Rahimeto, T. G. Debelee, D. Yohannes, and F. Schwenker, "Automatic pectoral muscle removal in mammograms," *Evolving Syst.*, vol. 12, no. 2, pp. 519–526, Jun. 2021.
[24] E. M. H. Saeed and H. A. Saleh, "Pectoral muscles removal in mammogram image by hybrid bounding box and region growing algorithm," in *Proc. Int. Conf. Comput. Sci. Softw. Eng. (CSASE)*, Apr. 2020, pp. 146–151.
[25] I. A. Lbachir, R. Es-Salhi, I. Daoudi, and S. Tallal, "A new mammogram preprocessing method for computer-aided diagnosis systems," in *Proc. IEEE/ACS 14th Int. Conf. Comput. Syst. Appl. (AICCSA)*, Oct. 2017, pp. 166–171.
[26] M. Mustra and M. Grgic, "Robust automatic breast and pectoral muscle segmentation from scanned mammograms," *Signal Process.*, vol. 93, no. 10, pp. 2817–2827, Oct. 2013.
[27] M. A. Rahman and R. K. Jha, "Multidirectional Gabor filter-based approach for pectoral muscle boundary detection," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 6, no. 4, pp. 433–445, Apr. 2021.
[28] A. Shrivastava, A. Chaudhary, D. Kulshreshtha, V. P. Singh, and R. Srivastava, "Automated digital mammogram segmentation using dispersed region growing and sliding window algorithm," in *Proc. 2nd Int. Conf. Image, Vis. Comput. (ICIVC)*, Jun. 2017, pp. 366–370.
[29] D. A. Zebari, D. A. Ibrahim, and A. Al-Zebari, "Suspicious region segmentation using deep features in breast cancer mammogram images," in *Proc. Int. Conf. Comput. Sci. Softw. Eng. (CSASE)*, Mar. 2022, pp. 253–258.
[30] T. Nagalakshmi, "Breast cancer semantic segmentation for accurate breast cancer detection with an ensemble deep neural network," *Neural Process. Lett.*, vol. 54, no. 6, pp. 5185–5198, Dec. 2022.
[31] M. J. Ali, B. Raza, A. R. Shahid, F. Mahmood, M. A. Yousuf, A. H. Dar, and U. Iqbal, "Enhancing breast pectoral muscle segmentation performance by using skip connections in fully convolutional network," *Int. J. Imag. Syst. Technol.*, vol. 30, no. 4, pp. 1108–1118, Dec. 2020.
[32] X. Yu, S.-H. Wang, J. M. Górriz, X.-W. Jiang, D. S. Guttery, and Y.-D. Zhang, "PeMNet for pectoral muscle segmentation," *Biology*, vol. 11, no. 1, p. 134, Jan. 2022.
[33] K. Zhou, W. Li, and D. Zhao, "Deep learning-based breast region extraction of mammographic images combining pre-processing methods and semantic segmentation supported by deeplab v3," *Technol. Health Care*, vol. 30, pp. 173–190, Feb. 2022.
[34] R. L. Prates, W. Gomez-Flores, and W. Pereira, "Semantic segmentation of mammograms using pre-trained deep neural networks," in *Proc. 18th Int. Conf. Electr. Eng., Comput. Sci. Autom. Control (CCE)*, Nov. 2021, pp. 1–6.
[35] V. Tiryaki and V. Kaplanoglu, "Deep learning-based multi-label tissue segmentation and density assessment from mammograms," *IRBM*, vol. 43, no. 6, pp. 538–548, 2022.
[36] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, "INbreast: Toward a full-field digital mammographic database," *Acad. Radiol.*, vol. 19, no. 2, pp. 236–248, 2013.
[37] D. C. Moura and M. A. Guevara López, "An evaluation of image descriptors combined with clinical data for breast cancer diagnosis," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 8, no. 4, pp. 561–574, Jul. 2013.
[38] Y. J. Kim, E. Y. Yoo, and K. G. Kim, "Deep learning based pectoral muscle segmentation on mammographic image analysis society (MIAS) mammograms," *Precis. Future Med.*, vol. 5, no. 2, pp. 77–82, Apr. 2021.
[39] Y. Mo, Y. Wu, X. Yang, F. Liu, and Y. Liao, "Review the state-of-the-art technologies of semantic segmentation based on deep learning," *Neurocomputing*, vol. 493, pp. 626–646, Jul. 2022.
[40] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
[41] T. Ghosh, L. Li, and J. Chakareski, "Effective deep learning for semantic segmentation based bleeding zone detection in capsule endoscopy images," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 3034–3038.
[42] G. Csurka, D. Larlus, and F. Perronnin, "What is a good evaluation measure for semantic segmentation?" in *Proc. Brit. Mach. Vis. Conf.*, 2013, pp. 1–11.

JUANITA HERNÁNDEZ LÓPEZ received the Computer Systems Engineering degree from the Instituto Tecnológico Superior of El Mante, in 2014, the master's degree in computer science from the Centro de Investigación y de Estudios Avanzados del I.P.N. (Cinvestav Unidad Tamaulipas), and the Ph.D. degree in engineering sciences and computer technologies from the Cinvestav Unidad Tamaulipas, in 2021. Currently, she works on a Conacyt project related to mammography analysis. Her main research interests include medical image analysis, pattern recognition, and computational intelligence.

JUAN HUMBERTO SOSSA AZUELA (Senior Member, IEEE) received the B.Sc. degree in electronics from the University of Guadalajara, in 1981, the M.Sc. degree in electrical engineering from CINVESTAV-IPN, in 1987, and the Ph.D. degree in informatics from the National Polytechnic Institute of Grenoble, France, in 1992. He is a full time Professor at the Centre for Computing Research, National Polytechnic Institute of Mexico. His main research interests include pattern recognition, artificial neural networks, image analysis, and robot control using image analysis.

ALBERTO SALVADOR NUÑEZ VARELA received the Ph.D. degree in computer sciences from the Autonomous University of San Luis Potosi. He is currently a Professor the Autonomous University of San Luis Potosi. His current research interests include software engineering, grammatical inference, natural language processing, and machine learning.

CÉSAR AUGUSTO RAMÍREZ GAMEZ received the master's degree in computer engineering in 2012. He is currently pursuing the Ph.D. degree in computer science from the Faculty of Engineering, University of San Luis Potosi. He is also a Professor at the Faculty of Engineering, University of San Luis Potosi. His research interests include computer vision, image processing, and machine learning.

VIRGINIA CANSECO GONZÁLEZ, photograph and biography not available at the time of publication.

FRANCISCO EDUARDO MARTÍNEZ PÉREZ graduated from the Autonomous University of Baja California. He received the master's degree in computer engineering and the Computer Engineering degree from the Autonomous University of San Luis Potosi, and the Doctor of Science degree. He is a full-time Research Professor at the Autonomous University of San Luis Potosi. His current research project is funded by PRONACES (Strategic National Programs by CONACYT) and called, "Automatic system for the immediate interpretation of mammograms for the determination of cancer risk." This project focuses on the development of image processing and intelligent algorithms. His work experience includes nine years in various public and private sector positions. Those that occurred for the San Luis Potosi State Government are: the Director of Information Technology at the State Attorney General's Office and the Deputy Director of Information Technology at the Secretary of Economic Development. His private sector position was that of Project Manager at a company called Quad-Tree. He is currently participating in the Dispersal of Science and Technology Network as a Representative of the Interinstitutional Committee of Open Government (CIGA) spearheaded by COPOCyT (Council of Science and Technology of the State of San Luis Potosi). Additionally, he is a Diffuser and a Disseminator on Science Communications. Currently, Ph.D. Martínez belongs to the National System of Researchers in México. His research interests include image processing, ambient intelligence (AmI), ubiquitous computing, human–computer interaction, and medical informatics. He has been a Professional Member of the IEEE and the ACM, since 2012.

SANDRA EDITH NAVA MUÑOZ received the Ph.D. degree in science from UABC, Mexico. She has been a Research Professor at UASLP, Mexico, since 2002. Her research interests include software engineering, human–computer interaction, context aware computing, and medical informatics.

HÉCTOR GERARDO PÉREZ GONZÁLEZ, photograph and biography not available at the time of publication.

JOSÉ IGNACIO NÚÑEZ VARELA received the M.Sc. degree in computer science from the University of Pittsburgh, in 2006, and the Ph.D. degree in computer science from the University of Birmingham, U.K., in 2013. Since 2014, he has been a Faculty Member at the Computer Science Department, Autonomous University of San Luis Potosi, Mexico, where he coordinates the Intelligent Systems Engineering Undergraduate Program in 2017.