Algorithm for the Detection of Breast Cancer in Digital Mammograms Using Deep Learning

Natalia Pirouzbakht, Jose Mejía , *Eléctrica y Computación, IIT/UACJ.*

Abstract—Breast cancer is one of the most frequent malignant tumors in women worldwide, the detection of this disease in time increases the possibility of receiving a less aggressive treatment and increases the survival rate. In this paper, we developed a cancer detection system that could be beneficial to help radiologists in cancer detection. To this end, we used a deep-learning network architecture. The proposed network consists of three convolutional layers followed each by pooling, and finally, four full connected layers provided the output of the network. Here, we also proposed to feed up the net with contrast-enhanced images to improve performance.

Index Terms—Deep learning, mammography, breast cancer, convolutional neural network.

1 Introducción

Nowadays, breast cancer is the most frequent malignant tumor causing the highest number of deaths in women worldwide [14]. In Mexico, in 2014, of the total number of cancer cases diagnosed in the population over 20 years of age, the breast is the one with the greatest impact with 19.4%. In the same year, the mortality rate per malignant breast tumor is 15 deaths per 100,000 women over 20 years of age. In 2015, the incidence of malignant breast tumor is 14.80 new cases per 100,000 people. Globally, an estimated 1.38 million new cases and 458,000 deaths are detected each year [8]. The women who come to perform a mammography annually, can detect this disease in time and therefore the possibility of receiving a less aggressive treatment. Although this test has been effective in early detection, there is still a high percentage of false positives and false negatives, which causes patients to undergo more invasive unnecessary treatment and / or testing causing anxiety, increased costs, and long-term psychosocial damage. Young women are more likely to get false negatives and positives. The main cause is the density of the breast, the denser it is, the greater is the probability of obtaining erroneous results since the visualization of the neoplasm is more difficult. Also, false positives often occur when women take estrogen, when they have had biopsies or when they have a family history of breast cancer. According to the federally funded Breast Cancer Surveillance Consortium in the United States, for every 1,000 women who undergo the test, 100 are further tested, but only 5 have breast cancer [2], [4], [5], [13].

Computer-aided detection in the field of medicine, was developed among other things to assist radiologists in the interpretation of mammograms [6]. In 2014, M. Tan et al. [19], worked on reducing false positives recalls using a computerized mammographic image feature analysis scheme, where they analyzed the global mammogram texture and density characteristics calculated from four-view images with the help of the technique of artificial neural networks.

In the same year, X. Liu and Z. Zeng [11] proposed a new automatic mass detection method for breast cancer with false positive reduction using support vector machines, where they obtained a sensitivity of mass detection of 78.2% with a specificity of 1.48 false positives per image. Finally, in 2016, T. Kooi et al. [10], worked on large scale deep learning for computer aided detection of mammography lesions. Their research, offered a direct comparison between an advanced mammography CAD system, based on a set of manually designed features and a convolutional neural network, with the aim of having a system that can, ultimately, read mammograms independently. Later in 2016, S. Suzuki et al. [11] adopted a convolutional neural network architecture (DCNN) that consisted of eight layers with weight, including 5five convolutional layers, and three fully-connected layers in their study. They first trained the DCNN using about 1.2 million natural images for classification of 1,000 classes. Then, they modified the last fully-connected layer of the DCNN and subsequently trained the DCNN using 1,656 regions of interest in mammographic image for two classes classification:mass and normal. The detection test was conducted on 198 mammographic images including 99 mass images and 99 normal images. The experimental results showed that the sensitivity of the mass detection was 89.9% and the false positive was 19.2%. J. Arevalo et al. [1], worked on a hybrid CNN method to learn image-based features in a supervised way for mammography mass lesion classifications. The developed method comprises two main stages: (i) preprocessing to enhance image details and (ii) supervised training for learning both the features and the breast imaging lesions classifier, as result, their method exhibited significant improved performance, such as histogram of oriented gradients (HOG) and histogram of the gradient divergence (HGD), increasing the performance from 0.787 to 0.822 in terms of the area under the ROC curve (AUC). Furthermore, in 2017, W. Sun, T.Tseng, J. Zhang and W. Qian [17], developed a graph based semi-supervised learning (SSL) scheme using deep convolutional neural network (CNN) for breast cancer diagnosis with a small portion of labeled data in training set. Four modules were included in the diagnosis system: data weighing, feature selection,

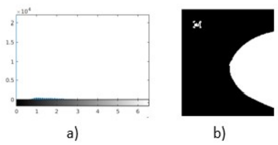

Figure 1. Preprocessing the image. a) Typical mammogram image histogram. b) Binary image with two objects: breast and label artifact.

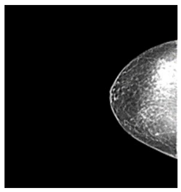

Figure 2. Processed image with the NSCT method.

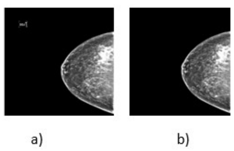

Figure 3. A mammogram image, a) with label artifact is in the left superior corner, b) without the artifact.

dividing co-training data labeling, and CNN. They achieved an area under the curve (AUC) of 0.8818, and the accuracy of CNN was 0.8243 using the mixed labeled and unlabeled data.

One of the difficulties facing the mammography study is that it generally has low contrast, making it difficult for radiologists to interpret results. In addition, it has been shown that the mammogram is susceptible to false positives and false negatives.

A study conducted in the United States in 2015 showed that women between 40 and 49 years of age constitute the highest percentage of false positive mammography results with the recommendation to perform other studies (33.1%). On average, 10% of 1,000 women who get a mammography will have to undergo further tests, but only 5 of that 10% actually have breast cancer. In the case of false negatives, 6% to 46% of women with invasive cancer will receive negative mammograms, especially if they are young or have dense breasts [3], [13].

The development of a cancer detection system could be beneficial to help radiologists in their interpretation and achieve a better diagnosis. In addition, the adoption of a system could reduce the workload of experts. Furthermore, in terms of economic benefit, a detection system could achieve a cost reduction as it could eliminate double reading, in addition to having a faster diagnosis.

Therefore, the development of an algorithm that by means of deep learning techniques can determine if a digital mammography presents or not breast cancer, could help radiologist in reducing the rate of false positives and negatives, being this of importance.

In this paper, an approach to detect mammograms with a possible tumor is presented, our approach is based on a Deep learning architecture. We proposed to preprocess the data using contrast enhancement in the contourlet domain. We expect that this prepossessing helps the network to generalize even with low data volumes in training the set. The preprocessing enhances several features in the images, such as microcalcifications that could help detect a cancer case more easily.

The contributions of this study are:

- Preprocessing of the mammogram images using the contourlet transform
- A new neural network topology of layers adapted to the task of breast cancer detection.

The rest of the paper is organized as follows. In section II we describe our proposed model to detect breast cancer cases, Section III experiments and results are showed, finally conclusions are provided in Section IV.

2 Methods

In this section, we describe the proposed algorithm which is composed of two stages. The first stage, described in section 2.1, consists in the preprocessing of data, where the images are prepared to be fed into the network. Finally, a second stage, which consists on feed the data to a convolutional neural network, is described in section 2.2 were we outline the proposed network topology.

2.1 Preprocessing of the data

The raw images from the data base of mammogram images are no suitable to be feed up directly into the network because they have a certain number of artifacts and because of the high dynamic range.

To alleviate this, we began the preprocessing of the images by first removing the label artifact that all images of the data base contain, see Figure 3a. For this end, we used binary image techniques. We obtained a binary image from the original in order to separate foreground (objects) from background, we selected a suitable threshold using the histogram of the image. The threshold is obtained as the value of intensity in the middle between the mean intensity of the background and the mean intensity of the object. Next, we assigned a "0" to the intensity of the pixels of the background or black value, while to the pixels in the objects or foreground we assigned a "1" or white value, see Figure 1b.

Once the binary image is obtained we found the objects in the image as sets of white pixels connected using an 8-neighborhood. Then we filtered the objects by area, that is, we only kept objects with a certain area, in our experiments an area of 1000 was sufficient to filter out the object that contains the chest area from the label artifacts that have less area, this value was obtained empirically from a set of 20 images, since the proportion of the area of the label regarding to the breast is almost constant in all images, the value found, worked for the entire database. We used the filtered binary image as a mask to further filter the original image in order to remove the label artifacts, an example of the result obtained is shown in Figure 3b.

The next step in the preprocessing was to equalize the intensity values in the image and reduce its dynamic range.

The original images in the database have a dynamic range of 0-65536 values of intensity, that, besides occupying much space, is not fully utilized, see the Figure 1a. This could affect the time or success of network training because only a portion of the dynamic range provides information. We reduced the dynamic range by first equalizing the image intensity using the technique of histogram equalization [7] and using a mapping to the range of 0 to 255.

The final preprocessing step was a contrast enhancement, for this end we used the technique used in [12], this improves the contrast of all structures in the mammogram, and improves visibility of small lesions such as microcalcifications, which are known to be an indicative of lesions such as tumors [15], [16]. We expected that this helped the network in learning specially improving the generalization when using small databases of images, which is the case of the mammogram database used.

Later on, we described the method to enhance the mammogram, for further details see [12]. The process begins by transforming the mammogram using the nonsubsampling contourlet transform.

$$Y = NSCT(I)$$

Where, I is a mammogram image, $NSCT(\cdot)$ is the nonsubsampling contourlet transform operator, and Y is the mammogram image in the transformed domain.

This transform decomposes the input image, I, in several subbands $y_{i,j}$, that is Y is a set of subbands $\{y_{1,1}, y_{2,1}, \ldots, y_{i,j}, \ldots\}$, where i is the number of level and j in the number of direction in the transform.

The subbands of Y, are then processed using

$$y'_{i,j} = \begin{cases} w_1 y_{i,j}(n_1, n_2) & \textit{if } b_{i,j}(n_1, n_2) = 0 \\ w_2 y_{i,j}(n_1, n_2) & \textit{if } b_{i,j}(n_1, n_2) = 1 \end{cases},$$

where y' is the processed subband, w_1 and w_2 are weights used for the tissue and microcalcifications respectively, $b_{i,j}$ is a binary image where points of high gradient are the foreground, and (n_1,n_2) are the coordinates of the subband processed. In this work, we used the values suggested in [12] for the weights. In Figure 2, it is show an example of an image processed with this technique.

2.2 Net Architecture

A Convolutional Neural Network consists of a number of convolutional, pooling, and fully connected layers. In our proposed network, see Figure 4, the first step is a convolutional layer, where we used 30 filters, with a kernel size of 5 x 5. To calculate the match of a feature to a patch of the image, each pixel in the kernel is multiplied by the value of the corresponding pixel in the image. To complete the convolution, we repeat the process, lining up the kernel with every possible image patch.

The feature map, it's a map where in the image the feature is found, and as result, we get a set of filtered images, one for each of the filters. It is possible to repeat this process as many times as wanted, therefore in this work we used 3 convolutional layers of the same size but with different filters, 30, 50 and 40 respectively.

The next step is the pooling layer, also known as max-pooling because we chose the maximum as statistic. This layer takes large images and shrink them down. We used three pooling layers of size 2×2 with a stride of 2 and the process consists of walking a small window across a filtered image of the convolution layer output and taking the maximum value from the window so it preserves the best fits of each feature within the window.

Finally, we used four fully connected layers, identified as ip1, ip2, ip3 and ip4, each of 105, 25, 7 and 2 neurons respectively which takes every single value and translate them into votes. We used the rectified linear unit (RelU) as nonlinearity activation function.

In this work, we only had two categories, images with and without cancer, so we ended ip4 with two neurons. The obtained votes are expressed as weights between each value. Then, the answer with the most votes wins and finally is declared the category of the input. The network was implemented using the Caffe framework described in [9].

3 Results

This section contains the results of training and testing the proposed network with the mammogram database. All experiments were performed on a computer with a Core i7-6700HQ, 2.6GHz $\times$ 8 processor and 31.3 GB of RAM, no GPU was used.

The database used is publicly-available provided by the group Health Cooperative for "The Digital Mammography DREAM challenge". The dataset is composed by 500 mammogram images, in different sizes ranging from 3328x2560 to 5928x4728 pixels in DICOM format. The database also includes annotated files to identify normal from cancer cases.

To speed up the training process, we changed the original format to portable network graphics (png), and reduced the size of all images to 208 x 208, with one channel or gray scale.

Since the cases with cancer were only 32 of 500 cases, we selected the training set as $29 + 41 = 80$ images, with 29 of the images presenting cancer cases and the rest normal cases, we used a test set composed of 3 images with cancer and 7 without cancer.

The training phase consisted of 4000 iterations, which were completed in 1 hour and 20 minutes approximately. We tested the resulting network in the test set, obtaining 100% of accuracy.

In Figure 5, is shown the final filter weights of the first convolution layer, we note that it is difficult to visually determine a predominant pattern or characteristic in data, that could have used by the network in its classification task.

4 Conclusion

In this paper, a novel algorithm for detecting breast cancer is presented. We preprocessed the mammogram image to remove artifact, and enhanced contrast by means of the NSCT, subsequently we fed up the image to a deep neural network. We obtained favorable results, which we attributed to the preprocessing of the images in the database that helps to enhance the structure of the mammogram. Thus, this preprocessing facilitated that the filters in the convolutional layers were able to adapt and obtain characteristics of importance to classify correctly these images, even though the

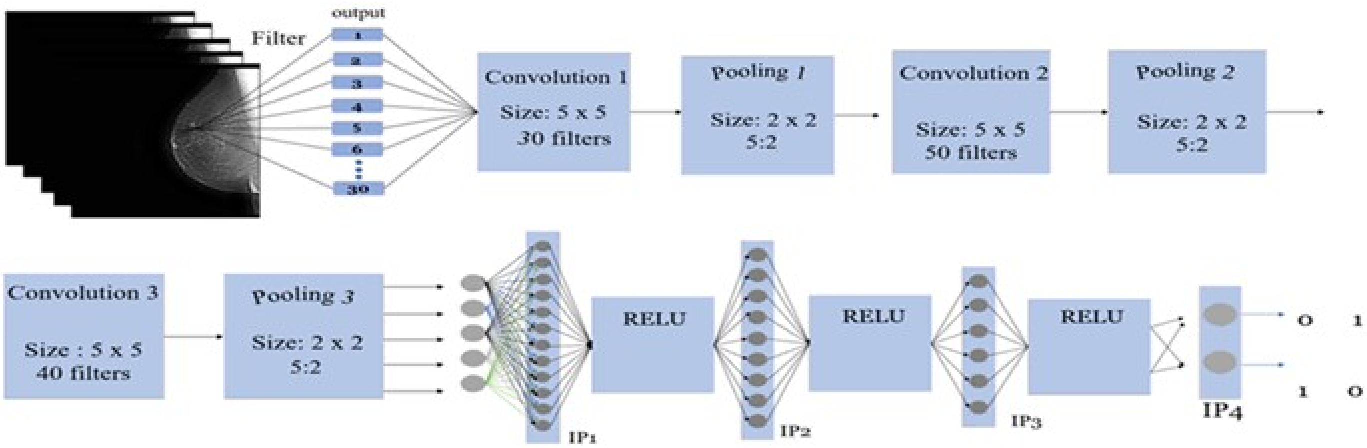

Figure 4. Network architecture used in this work.

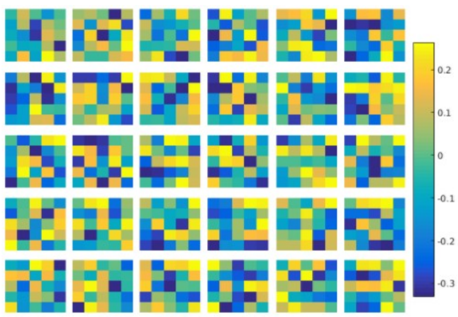

Figure 5. Filter weights for the first convolutional layer, after the network is fully trained using 4000 iterations.

training data base was small. As a further work, we suggest to test the algorithm with a larger database, in order to have a better idea of its performance, and avoid a possible over-fitting.

References

[1] Arevalo J. et al. (2016). *Representation learning for mammography mass lesion classification with convolutional neural networks*, *Elsevier*, vol. 37, pp-248-257.

[2] Brodersen J., and Siersma V. D. (2013). *Long-Term Psychosocial Consequences of False-Positive Screening Mammography*, *Annals of Family Medicine*, vol. 11, no. 2, pp. 106-115.

[3] National Cancer Institute. (2017, February 22). *Breast Cancer Screening (PDQ®)–Health Professional Version [Online]*. Available: https://www.cancer.gov/types/breast/hp/breast-screening-pdq.

[4] American Cancer Society. (2017). *Limitations of Mammograms [Online]*. Available: https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/limitations-of-mammograms.html.

[5] Elmore J. G., Barton M. B., Moceri V. M., Bolk S., Arena P. J., and Fletcher S. W. (1998). *Ten-Year Risk of False Positive Screening Mammograms and Clinical Breast Examinations*, *The New England Journal of Medicine*, vol. 338, no. 16, pp. 1089-1096.

[6] Fenton J. et al. (2007). *Influence of Computer-Aided Detection on Performance of Screening Mammography*, *The New England Journal of Medicine*, vol. 356, no. 14, pp. 1399-1409.

[7] Gonzalez R., Eddins S., and Woods R. E. (2004). *Digital image processing using MATLAB*.

[8] Instituto Nacional de Estadísticas y Geografía (2016). *Estadísticas a propósito del día mundial de la lucha contra el cáncer de mama*.

[9] Jia Y., Shelhamer E., Donahue J., Karayev S., Long J., Girshick R., and Darrell T. (2014). *Caffe: Convolutional architecture for fast feature embedding. Proceedings of the 22nd ACM international conference on Multimedia*, pp. 675-678.

[10] Kooi T. et al. (2016). *Large scale deep learning for computer aided detection of mammographic lesions*, *Medical Image Analysis*, vol. 35, pp. 303-312.

[11] Liu X., and Zeng Z. (2014). *A new automatic mass detection method for breast cancer with false positive reduction*, *Neurocomputing*, vol. 152, pp. 388-402.

[12] Mejia J., Domínguez H. D. J. O., Villegas O. O. V., Sánchez V. G. C., and Maynez L. O. (2009). *The nonsubsampled contourlet transform for enhancement of microcalcifications in digital mammograms*, In Mexican International Conference on Artificial Intelligence, *Springer*, Berlin, Heidelberg, pp.292-302.

[13] Myers E. R. et al. (2015). *Benefits and Harms of Breast Cancer Screening: A Systematic Review*, *Jama*, vol. 314, no. 15, pp. 1615-1634.

[14] Reynoso-Noverón N., Villaseñor-Navarro Y., Hernández-Ávila M., and Mohar-Betancourt A. (2013). *Carcinoma in situ e infiltrante identificado por tamizaje mamográfico oportunista en mujeres asintomáticas de la Ciudad de México*, *Salud pública de México*, vol. 55, no. 5, pp. 469-477.

[15] Sickles E. A. (1984). *Mammographic Features of Early Breast Cancer*, *American Journal of Roentgenology*, vol. 143, pp. 461 - 464.

[16] Sickles E. A. (1986). *Mammographic Features of 300 Consecutives Nonpalpable Breast Cancers*, *American Journal of Roentgenology*, vol. 146, pp. 661 - 663.

[17] Sun W., Tseng T., Zhang J. and Qian W. (2017). *Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data*, *Elsevier*, vol. 57, pp.4-9.

[18] Suzuki S. et al. (2016). *Mass Detection Using Deep Convolutional Neural Network for Mammographic Computer–Aided Diagnosis*, in Proceedings of the SICE Annual Conference 2016, pp. 1382-1386.

[19] Tan M., Pu J., and Zheng B. (2014). *Reduction of false-positive recalls using a computerized mammographic image feature analysis scheme*, *Physics in Medicine and Biology*, vol. 59, no. 15.