2020 17th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)
Mexico City, Mexico. November 11-13, 2020

Predicting the BI-RADS Lexicon for Mammographic Masses Using Hybrid Neural Models

Juanita Hernández-López
Cinvestav-IPN
Unidad Tamaulipas
Ciudad Victoria, Mexico
juanita.hernandez@cinvestav.mx

Wilfrido Gómez-Flores
Cinvestav-IPN
Unidad Tamaulipas
Ciudad Victoria, Mexico
wgomez@cinvestav.mx

***Abstract*—The mammographic image analysis carried out by radiologists describes findings to infer the type of breast pathology to lead patients to adequate treatments. For such an analysis, the experts use a lexicon in the Breast Imaging-Reporting and Data System (BI-RADS), through which the shape and margin of the tumor are described, as well as the breast parenchymal density to determine the risk of developing breast cancer. In this sense, the development of technological tools to assist radiologists in the mammographic description is justified. In this article, two hybrid classification techniques based on neural models are proposed to predict the BI-RADS lexicon attributes for masses. These hybrid models comprise a pre-trained convolutional neural network (AlexNet) for feature extraction and neural models for pattern classification, namely, multilayer perceptron (MLP) and dendrite morphological neuron (DMN). The obtained classification results suggest that the hybrid neural model schemes obtain an accuracy of 91% for predicting the shape and density attributes and an 85% accuracy for predicting margin attributes. The Wilcoxon test indicates no significant differences between both hybrid models. Therefore, the proposed models can be useful for predicting attributes included in the BI-RADS lexicon for masses.**

***Index Terms*—Breast Cancer; BI-RADS Lexicon; Mammography Classification; Transfer Learning; Neural Networks.**

I. Introduction

Breast cancer is a chronic degenerative disease caused by the abnormal growth of breast cells. In Mexico, this pathology is the leading cause of death in adult women due to malignant neoplasms. It represents the second cause of death in the female population between 30 and 54 years of age [8].

Mammography (X-ray) is the most cost-effective imaging technique for early detection of breast cancer in screening programs. This technique achieves a detection rate of about 70% in patients with breast cancer [9].

Usually, radiologists specialized in breast imaging detect mammographic findings based on the Breast Imaging-Reporting and Data System (BI-RADS). The BI-RADS standardizes mammographic reports and defines the following clinical conducts: 1) recommend a routine study if the tumor has benign features, and 2) recommend a biopsy if it has malignant features [10].

The BI-RADS involves a lexicon for masses that describes the tumor's shape and margin and breast tissue density. Benign tumors are commonly of regular morphology with well-defined and circumscribed margins, while malignant tumors tend to have irregular shapes and ill-defined margins. Breast density represents a possible risk factor for breast cancer [11]. Fig. 1 shows the histopathological tendency of the three BI-RADS terms for masses: shape, margin, and density. Notice that each term has a set of attributes to describe the tumor and breast tissue.

Commonly, performing visual mammographic analysis leads to inter- and intra-observer variations [10], [11]. Furthermore, the high demand for mammographic studies may exceed the available human resources certified in breast cancer [12]. Therefore, computer-aided diagnosis (CAD) systems are adjunctive tools that analyze multiple mammograms using digital image analysis and artificial intelligence techniques. A CAD system generally classifies the tumors into benign and malignant classes, reducing the inter-observer variability [11]. A CAD follows the conventional pipeline of object recognition: image preprocessing, feature extraction, and classification. In this scheme, a critical part is the feature extraction to transform the image into quantitative values that describe its nature. Therefore, the CAD's accuracy depends on the quality of features extracted from the images and the classification scheme.

Currently, the tendency in the development of CAD systems involves describing mammographic findings according to the radiological BI-RADS lexicon, such as the tumor attributes of shape, margin, and density [13]. Besides, convolutional neural networks (CNNs) allow learning features automatically from raw images, reducing the dependency on finding adequate

978-1-7281-8987-1/20/$31.00

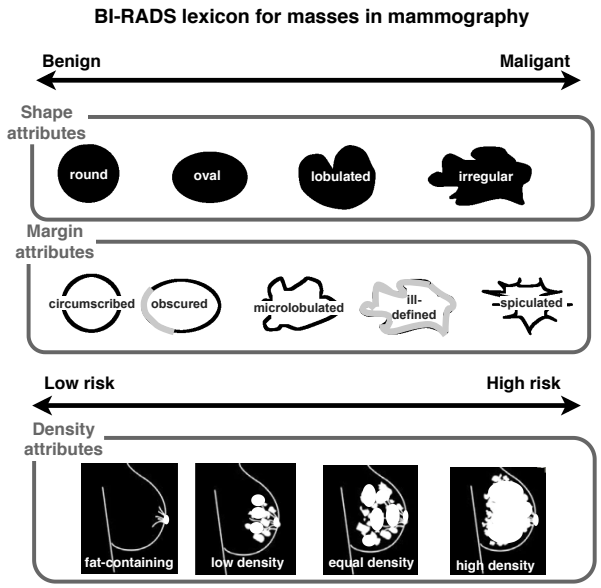

Fig. 1. Histopathological tendency of the BI-RADS lexicon according to their corresponding attributes.

hand-crafted features.

CNNs have an architecture based on layers that simulate human neural abstraction, obtaining low, medium, and high-level feature maps. CNNs training requires large volumes of data to learn the parameters of each layer. However, medical data with expert annotations are often scarce. So, training a CNN from scratch with a small dataset risks overfitting. In this case, it is convenient to reuse a pre-trained CNN model with millions of images to solve a new classification task through transfer learning.

A CNN can be used as an encoder to extract learned features to feed a conventional classifier, forming hybrid schemes. This approach was applied by Fonseca *et al.* [4] for classifying breast density. They used a CNN with HT-L3 architecture for feature extraction and a support vector machine (SVM) for classification. The SVM settings include a linear kernel and penalization hyperparameter $C = 1$. The method obtained an accuracy of $73.05 \pm 3.68\%$. In another work, Fonseca *et al.* [1] combined the CNN AlexNet and an SVM for classifying mammographic density and obtained a kappa performance equal to 0.8. Despite these efforts to predict breast density with hybrid schemes, it is still missing a complete scheme to predict the BI-RADS lexicon for masses by including the shape and margin terms in addition to the breast density term.

Because CNNs simulate the human visual abstraction process, hybrid schemes can use other classification algorithms inspired by neural interconnections, such as artificial neural networks (ANN). A widely used ANN model for pattern classification is the multilayer perceptron (MLP), whose basic unit is the perceptron, which is a linear discriminant. By interconnecting several perceptron units in several layers, nonlinear decision boundaries are generated. Dendrite morphological neuron (DMN) is another type of artificial neuron used for pattern classification. A dendrite is defined by a hyperbox that encloses patterns of the same class and uses the min and max operators as activation functions. A single DMN neuron is capable of solving nonlinearly separable problems [3].

In this paper, we propose using CNNs to extract learned features to describe the three terms of the BI-RADS lexicon for masses. Besides, the CNNs outputs are used to predict each lexicon term's attributes by using MLP and DMN models. Hence, two hybrid neural models are built: CNN-MLP and CNN-DMN.

TABLE I
MAMMOGRAPHY DATASET CHARACTERISTICS.

Lexicon	Attribute	# Images
Shape	Round	271
	Oval	431
	Lobulated	420
	Irregular	182
Margin	Non-spiculated	1363
	Spiculated	402
Density	Fatty	191
	Glandular	238
	Dense	73

II. MATERIALS AND METHODS

A. Mammography dataset

The set of mammograms used in this study was obtained from the Digital Database for Screening Mammography (DDSM)[1]. This dataset includes mammographic lesions labeled according to the BI-RADS lexicon for masses. It includes manual delineations of the tumors (i.e., binary masks) made by at least one radiologist. Besides, extreme points of the binary mask define the minimum bounding box that encloses the tumor, which is used to crop a region of interest (ROI). Table I summarizes the number of cases by attribute used for modeling the BI-RADS lexicon for masses.

B. Hybrid neural models

This study explores the use of hybrid neural models that combine CNNs for automatic description of the BI-RADS lexicon attributes with neural models for pattern classification, as shown in Fig. 2. The necessary steps for training the proposed approaches are described in the next [14].

1) Image preprocessing: The images are contrast-stretched, in which the minimum and maximum gray levels are defined by the 5% and the 95% percentiles of the intensity histogram. This process normalizes the mammograms' intensity ranges since they were acquired by different scanners with distinct dynamic ranges [15].

On the other hand, each mammography is transformed into a useful representation according to the lexicon that will be described:

[1]http://marathon.csee.usf.edu/Mammography/Database.html

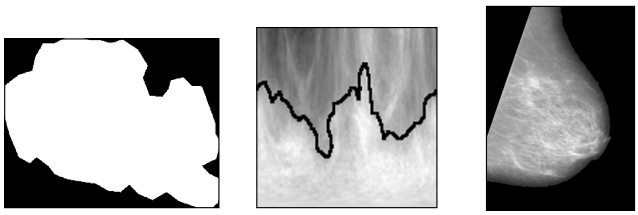

Fig. 2. Example of images included in the dataset for modeling the a) tumor shape, b) tumor margin, and c) breast density.

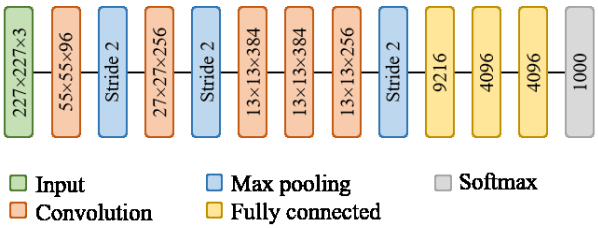

Fig. 3. AlexNet architecture.

- Shape: The global appearance of the tumor shape is extracted to describe its morphology. It is used the ROI of the tumor binary mask defined by the manual delineation of the radiologist, as shown in Fig. 2a.
- Margin: The tumor contour signature extracts information of its margin. The ROIs of the binary mask and the grayscale image are transformed to polar coordinates from the tumor's centroid. Edge detection is applied to the binary image to mask the grayscale image. The resultant image shows intensity variations radially outward from the tumor's centroid, and its signature is captured at the border of the tumor, as shown in Fig. 2b.
- Density: Intensity variations within the breast parenchyma are useful to describe its density. The pectoral muscle and the background of the mammogram are masked to enhance only the breast tissue region's intensity levels, as shown in Fig. 2c.

2) Learned features: Automatic feature extraction is performed by the CNN called AlexNet, whose architecture is shown in Fig. 3. The idea is to use a pre-trained model for using transfer learning on new data. AlexNet has an input layer of size $227 \times 227 \times 3$ (i.e., an RGB image), followed by five convolutional layers, three of them with max-pooling, and three fully connected layers. The output layer is a Softmax function configured to distinguish between 1000 classes of objects of the Imagenet dataset. AlexNet won the ILSVRC (Imagenet Large Scale Visual Recognition Challenge) competition in 2012.

The transfer learning process considers the following aspects:

- Resizing: The AlexNet input layer receives color images of size $227 \times 227 \times 3$. Therefore, the image dataset is adjusted to that size, and each grayscale image is replicated twice to obtain the three channels of an RGB image [2].
- Data augmentation: In order to reduce model overfitting, artificial data augmentation is performed. Scaling and rotations are randomly applied to training images in the range $[-30°, 30°]$, and scale factors of $[0.9, 1.1]$. This procedure assumes that more information can be extracted from the original dataset through augmentations.
- Training settings: the pre-trained convolution layers in the AlexNet are reused, and only the final fully connected layer is fine-tuned for the new classification problem. The training parameters are a mini-batch size of 200, a learning rate of 3×10^{-4}, a momentum value of 0.75, and a regularization factor of 0.05.

3) BI-RADS lexicon prediction: Figure 4 shows the two neural models used for pattern classification. The first model is a multilayer perceptron (MLP) network with a single hidden layer [6]. The MLP response in the kth output neuron is

$$z_k(\mathbf{x}) = \sigma\left(w_{j0} + \sum_{j=1}^{h} w_{kj}\sigma\left(w_{k0} + \sum_{i=1}^{d} w_{ji}x_i\right)\right), \quad (1)$$

where $\mathbf{x} = [x_1, \ldots, x_d]^T$ is the input pattern, w_{ji} is a weight between the input and hidden layers, w_{kj} is a weight between the hidden and output layers, $\sigma(\cdot)$ is the sigmoid activation function, and w_{j0} and w_{k0} are the bias values added in the input and hidden layers, respectively. The optimum number of hidden units is searched in the range $h \in [5, 100]$, with steps of 5 units. MLP is trained during 1000 epochs by gradient descent with momentum factor equal to 0.9 and an adaptive learning rate, which considers the next conditions. For each epoch, if the performance decreases, then the learning rate is increased by the factor 1.05. If performance increases by more than the factor 1.04, the learning rate is decreased by the factor 0.70. The initial learning is set to 0.01. These MLP settings are commonly used [16].

The second neural model is a hyperbox-based dendrite morphological neuron (DMN) [3]. The DMN output is the predicted class to the input pattern $\mathbf{x}$, which is expressed as

$$y(\mathbf{x}) = \arg\max_{k=1,\ldots,c} \left(r_k(\mathbf{x})\right), \quad (2)$$

where r_k is the dendrite cluster response for the kth class:

$$r_k(\mathbf{x}) = \min_{j=1,\ldots,l_k} \left(\min(\min(\mathbf{x} - \mathbf{w}, \mathbf{w} + \mathbf{b} - \mathbf{x}))\right), \quad (3)$$

where $\mathbf{w}_{j,k} = (w_{1,j,k}, \ldots, w_{d,j,k})^T$ is the coordinate of the lowest extreme point of the hyperbox (i.e., the dendrite), and $\mathbf{b}_{j,k} = (b_{1,j,k}, \ldots, b_{d,j,k})^T$ is a vector containing the side lengths of the hyperbox for each dimension. DMN is trained with stochastic gradient descent during 800 epochs and a learning rate of 0.01.

Two types of hybrid neural models, namely, CNN-MLP and CNN-DMN, are proposed by combining a CNN with a neural model for pattern classification. Figure 5 shows the basis hybrid neural model proposed for predicting the BI-RADS lexicon for masses. The shape model addresses a multiclass

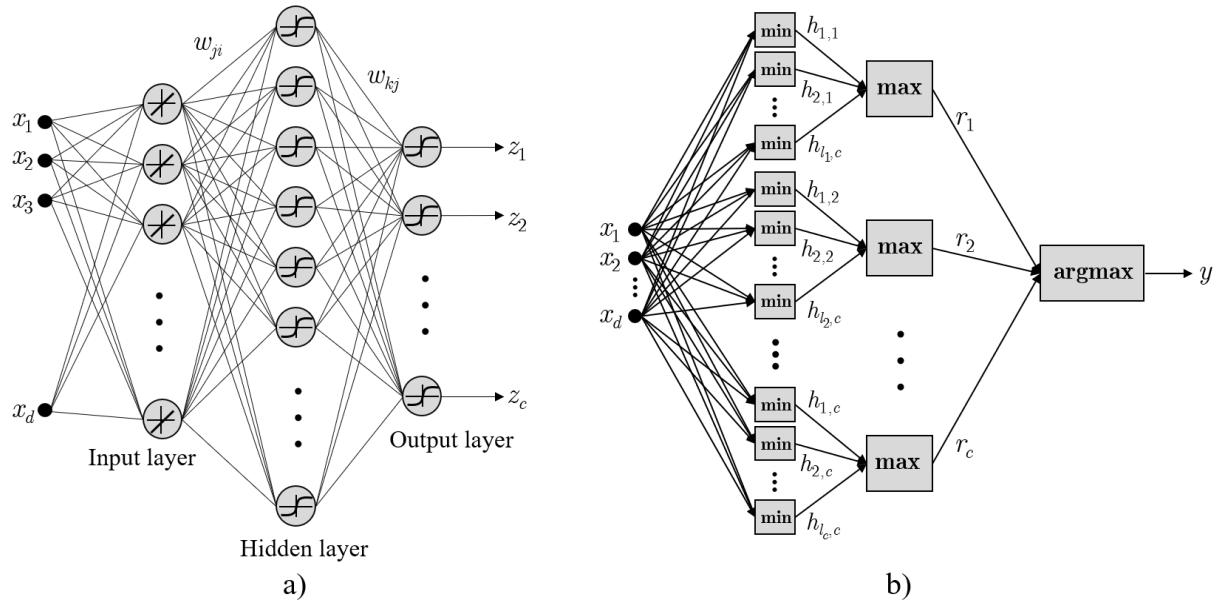

Fig. 4. Neural models for pattern classification: a) multilayer perceptrón, and b) dendrite morphological neuron.

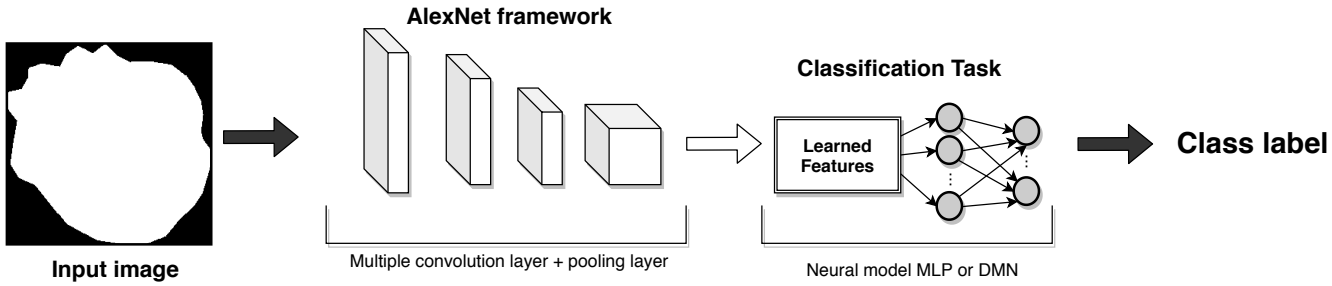

Fig. 5. Hybrid neural scheme for predicting the attributes of the three BI-RADS lexicon terms for masses: shape, margin, and density.

classification problem for differentiating between round, oval, lobulated, and irregular attributes. For the margin model, a binary classification problem is solved to distinguish between the spiculated and non-spiculated attributes, where the latter encompasses the circumscribed, undefined, obscured, and microlobulated attributes. Finally, the density model deals with a multiclass classification problem to discriminate between fatty, glandular, and dense densities, where glandular comprises the union of low and medium densities, also known as equal density.

C. Classification performance evaluation

The k-fold cross-validation method is used to generate independent training and test sets, with $k = 10$. In every experiment, one group is the test set, and the remaining nine groups comprise the training set.

For a classification problem with c classes, the corresponding confusion matrix $\mathbf{C}$ is a square matrix of size c-by-c whose ijth entry $\mathbf{C}_{ij}$ is the number of elements of the true class i that have been assigned to class j by the classifier [7].

The accuracy is probably the most often used measure to evaluate the overall effectiveness of a classifier, which is expressed by

$$\mathrm{ACC} = \frac{\mathrm{tr}(\mathbf{C})}{n}, \tag{4}$$

where $\mathrm{tr}(\cdot)$ is the trace operator and n is the total number of test observations.

Because the dataset presents unbalanced classes, the accuracy tends to be optimistic, due to the high hit rate of the majority class. In order to deal with unbalanced classes, the Matthews correlation coefficient (MCC) is used, which is expressed by

$$\mathrm{MCC} = \frac{n \cdot \mathrm{tr}(\mathbf{C}) - \sum_{kl} \mathbf{C}_k \mathbf{C}_l}{\sqrt{n^2 - \sum_{kl} \mathbf{C}_k (\mathbf{C}^T)_l} \sqrt{n^2 - \sum_{kl} (\mathbf{C}^T)_k \mathbf{C}_l}}, \tag{5}$$

where $\mathbf{C}_k$ is the kth row of $\mathbf{C}$, $\mathbf{C}_l$ is the lth column of $\mathbf{C}$, and $\mathbf{C}^T$ is transpose of $\mathbf{C}$.

Besides, for the ith class, the true positives (TP_i), true negatives (TN_i), false positives (FP_i), and false negatives

(FN_i) are calculated as

$$\mathrm{FP}_i = \sum_{l=i} C_l - C_{il}, \quad (6)$$

$$\mathrm{FN}_i = \sum_{k=i} C_k - C_{ki}, \quad (7)$$

$$\mathrm{TP}_i = C_{ii}, \quad (8)$$

$$\mathrm{TN}_i = n - (\mathrm{TP}_i + \mathrm{FP}_i + \mathrm{FN}_i). \quad (9)$$

From counts in Eqs. 6 to 9, the performance indices of sensitivity (SEN), specificity (SPE), and area under the curve (AUC), also known as balanced accuracy, are calculated as

$$\mathrm{SEN} = \frac{1}{c}\sum_{i=1}^{c} \frac{\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FN}_i}, \quad (10)$$

$$\mathrm{SPE} = \frac{1}{c}\sum_{i=1}^{c} \frac{\mathrm{TN}_i}{\mathrm{TN}_i + \mathrm{FP}_i}, \quad (11)$$

and

$$\mathrm{AUC} = \frac{1}{2c}\sum_{i=1}^{c} \left(\frac{\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FN}_i} + \frac{\mathrm{TN}_i}{\mathrm{TN}_i + \mathrm{FP}_i} \right). \quad (12)$$

All the indices mentioned above should tend toward unity to indicate an adequate classification performance.

The Wilcoxon statistical test ($\alpha = 0.05$) determines if there is a significant difference between the CNN-MLP and CNN-DMN methods.

The computing platform considered includes an Intel i7 processor with four cores at 3.60 GHz, 32 GB of RAM, and a graphic card NVIDIA GeForce GTX 1050Ti. All the programs were developed in MATLAB R2019a (The Mathworks, Boston, Massachusetts, USA).

III. Results

Table II shows the results obtained by the CNN-MLP and CNN-DMN methods when predicting the BI-RADS terms of shape, margin, and density. The comparative results indicate that both methods obtained similar classification performances for all the indices.

In terms of accuracy, CNN-MLP and CNN-DMN obtained 91%, 85%, and 91% for shape, margin, and density models, respectively. The results obtained by the Wilcoxon test indicate that the p-values are higher than 0.05 (i.e., 95% confidence interval) for all the performance indices. Hence, there is no statistically significant difference between CNN-MLP and CNN-DMN methods for all models of the BI-RADS lexicon for masses.

Tables III to V show the confusion matrices obtained by CNN-MLP and CNN-DMN methods for the three models of the BI-RADS lexicon. Each entry in the matrices represents the mean of the ten cross-validation experiments. The diagonal of each matrix shows the number of cases correctly classified.

The shape model confusion matrices reveal that the most notable overlap between attributes is produced between oval and round attributes than other types of shapes. This effect is because of the similarity between round and oval shape patterns since both shapes represent benign attributes. Also, it is observed that the irregular attribute does not present overlap with round and oval shapes, which is useful to distinguish between benign and malignant tumors.

On the other hand, the margin model confusion matrices present a greater overlap between attributes than the shape and density models. This behavior is because of the low contrast in the interface between the tumor's margin and the adjacent glandular tissue. This aspect is mostly observed in tissue with high density, where the contrast between the tumor border and fibro-glandular tissue is shallow.

Finally, the density model confusion matrices present less overlap between attributes compared to the other models. There is no overlap between fatty and dense attributes since the contrast between both attributes is high. The overlap between density attribute transitions is also produced, particularly in classes: fatty-glandular and glandular-dense, although the patterns were classified correctly in most cases.

It is worth mentioning that the number of hidden neurons in the MLP network was for the shape and density models 11 ± 5, while for the margin model, it was 12 ± 1.

TABLE II
Classification results of the proposed approaches. The mean and standard deviation (in parentheses) of 10 folds of cross-validation are shown. The last column shows the p-value of the Wilcoxon test.

Model	Index	CNN-MLP	CNN-DMN	p-value
Shape	MCC	0.887 (0.039)	0.888 (0.043)	$p > 0.05$
	ACC	0.917 (0.028)	0.918 (0.037)	
	AUC	0.975 (0.016)	0.985 (0.007)	
	SEN	0.894 (0.039)	0.895 (0.046)	
	SPE	0.970 (0.010)	0.970 (0.011)	
Margin	MCC	0.565 (0.102)	0.553 (0.087)	$p > 0.05$
	ACC	0.853 (0.032)	0.850 (0.027)	
	AUC	0.769 (0.057)	0.760 (0.047)	
	SEN	0.924 (0.015)	0.925 (0.017)	
	SPE	0.615 (0.105)	0.594 (0.089)	
Density	MCC	0.857 (0.053)	0.854 (0.059)	$p > 0.05$
	ACC	0.912 (0.032)	0.910 (0.036)	
	AUC	0.960 (0.023)	0.978 (0.013)	
	SEN	0.885 (0.055)	0.880 (0.059)	
	SPE	0.947 (0.020)	0.945 (0.022)	

TABLE III
Confusion matrices of the shape model: a) CNN-MLP and b) CNN-DMN.

A)

	I	L	O	R
Irregular (I)	26	1	0	0
Lobulated (L)	0	42	0	0
Oval (O)	0	1	38	3
Round (R)	0	0	5	13

B)

	I	L	O	R
I	26	1	0	0
L	0	42	0	0
O	0	1	38	3
R	0	0	5	13

TABLE IV
Confusion matrices of the margin model: a) CNN-MLP and b) CNN-DMN.

A)

	N-S	S
Non-spiculated (N-S)	126	10
Spiculated (S)	16	25

B)

	N-S	S
N-S	126	10
S	16	24

TABLE V
CONFUSION MATRICES OF THE DENSITY MODEL: A) CNN-MLP AND B) CNN-DMN.

A)

	F	G	D
Fatty (F)	17	2	0
Glandular (G)	1	23	0
Dense (D)	0	1	6

B)

	F	G	D
F	17	2	0
G	1	23	0
D	0	1	6

IV. DISCUSSION AND CONCLUSIONS

This paper introduces two hybrid neural models for predicting the three terms of the BI-RADS lexicon for masses: shape, margin, and density. A CNN, defined by the AlexNet architecture, was used for extracting learned features, whereas two neural models for pattern classification were utilized: multilayer perceptron (MLP) and dendrite morphological neuron (DMN). Hence, the CNN-MLP and CNN-DMN hybrid neural schemes were proposed.

In the literature, hybrid schemes have generally been used to classify mammographic lesions into benign and malignant classes [5]. Also, the classification of breast density using hybrid approaches has been addressed [1]. These hybrid schemes commonly use an SVM as the base classifier due to its high generalization in recognition tasks. However, the use of other neural models for creating hybrid schemes for predicting the BI-RADS lexicon for masses is underexplored.

The shape model distinguishes between four types of attributes: round, oval, lobulated, and irregular. On the other hand, the margin model discriminates between spiculated and non-spiculated margins. Finally, the density model classifies the breast tissue density as fat, glandular, and dense. The outputs of these models can provide responses associated with radiologists' radiological language to provide a second opinion regarding the description of the lesion and breast composition, which is useful for decision-making.

CNN-MLP and CNN-DMN obtained competitive classification performance for all the classification indices. In terms of accuracy, both methods obtained 91%, 85%, and 91% for the shape, margin, and density models, respectively. Furthermore, the Wilcoxon test indicated no significant differences between the results obtained by both hybrid models. Therefore, the use of these two hybrid neural models can be useful for predicting attributes included in the BI-RADS lexicon for masses.

The DMN architecture is versatile and straightforward since a single neuron can generate nonlinear decision boundaries. On the other hand, MLPs require at least one layer of hidden neurons to generate nonlinear decision boundaries. This aspect can increase the computational cost of training in an MLP, whereas, in a DMN, the cost of training can be reduced due to its geometric exploitation property.

Future work considers developing a CAD system that incorporates the BI-RADS lexicon's prediction for breast lesions classification in benign and malignant classes, intending to provide mammographic reports compatible with the BI-RADS standard. Besides, other pre-trained CNN architectures for feature extraction will be evaluated.

ACKNOWLEDGMENT

Juanita Hernández-López thanks to the National Council of Science and Technology (CONACyT, Mexico) for the research scholar grant (No. 463795). This research was supported by a Fondo SEP-Cinvestav 2018 grant (No. FidSC2018/145).

REFERENCES

[1] P. Fonseca, B. Castañeda, R. Valenzuela, and J. Wainer, "Breast Density Classification with Convolutional Neural Networks", Lecture Notes in Computer Science, vol 10125. Springer, Cham, 2017.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", in Advances in Neural Information Processing Systems, pp. 1097-1105, Curran Associates, Inc., 2012.

[3] E. Zamora and H. Sossa, "Dendrite morphological neurons trained by stochastic gradient descent", Neurocomputing, vol. 260, pp. 420-431, 2017.

[4] P. Fonseca, J. Mendoza, J. Wainer, J. Ferrer, J. Pinto, J. Guerrero, and B. Castañeda, "Automatic breast density classification using a convolutional neural network architecture search procedure", Proc. SPIE, Medical Imaging 2015: Computer-Aided Diagnosis, vol.9414, pp. 941428, 2015.

[5] Z. Jiao, X. Gao, Y. Wang, and J. Li, "A deep feature based framework for breast masses classification", Neurocomputing, vol. 197, pp. 221-231, 2016.

[6] S. O. Haykin, Neural Networks and Learning Machines. Prentice-Hall, 3rd ed., 2009.

[7] A. Rodríguez-Cristerna, W. Gómez-Flores, and W. C. de Albuquerque Pereira, "A computer-aided diagnosis system for breast ultrasound based on weighted bi-rads classes", Computer Methods and Programs in Biomedicine, vol. 153, no. Supplement C, pp. 33-40, 2018.

[8] J. Rodrǵuez-Rojas, M. Garza-Montemayor, V. Treviño-Alvarado, and J. G. Tamez-Pena, "Predictive features of breast cancer on mexican screening mammography patients", Proc. SPIE, Medical Imaging 2013: Computer-Aided Diagnosis, vol. 8670, pp. 867023, 2013.

[9] M. Gargouri Laroussi, N. G. Ben Ayed, A. Damak Masmoudi and D. Sellami Masmoudi, "Diagnosis of masses in mammographic images based on Zernike moments and local binary attributes", IEEE, 2013 World Congress on Computer and Information Technology (WCCIT), pp. 1-6, 2013.

[10] A. Redondo, M. Comas, F. Maciá, F. Ferrer, C. Murta-Nascimento, M. T. Maristany, E. Molins, M. Sala, and X. Castells, "Inter- and intra radiologist variability in the bi-rads assessment and breast density categories for screening mammograms", The British Journal of Radiology, vol. 85, no. 1019, pp. 1465-1470, 2012.

[11] R. Rass Winkel, M. V. Euler-Chelpin, M. Nielsen, P. Diao, M. Bachmann Nielsen, W. Yao Uldall, and I. Vejborg, "Inter-observer agreement according to three methods of evaluating mammographic density and parenchymal pattern in a case control study: impact on relative risk of breast cancer.", BMC Cancer, vol. 15, no. 274, pp. 1-14, 2015.

[12] E. Gaona, S. I. Osorio and V. Arenas, "Efficiency indicators of mammography in the detection of breast cancer early stages: Exploratory study in mexico", International Journal of Applied Science and Technology, vol. 7, no. 4, pp. 32-36, 2017.

[13] L. Tsochatzidis, K. Zagoris, N. Arikidis, A. Karahaliou, L. Costaridou, and I. Pratikakis, "Computer-aided diagnosis of mammographic masses based on a supervised content-based image retrieval approach", Pattern Recognition, vol. 71, pp. 106-117, 2017.

[14] R. C. González and R. E. Woods, Digital Image Processing, Prentice-Hall, 3rd ed., 2008.

[15] C. Tromans, and M. Brady, "An alternative approach to measuring volumetric mammographic breast density", Springer, International Workshop on Digital Mammography (IWDM), Lecture Notes in Computer Science, vol. 4046, 2006.

[16] T. Hastie, R. Tibshirani, and J. Friedman, Neural Networks, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer New York, pp. 389–416.