Lightweight CNN for Detecting Microcalcifications Clusters in Digital Mammograms

Ricardo Salvador Luna-Lozoya[1], Humberto de Jesús Ochoa-Domínguez[*,1],
Juan Humberto Sossa-Azuela[2], Vianey Guadalupe Cruz-Sánchez[1],
Osslan Osiris Vergara-Villegas[1]

[1] Universidad Autónoma de Ciudad Juárez,
Ciudad Juárez,
Mexico

[2] Instituto Politécnico Nacional,
Ciudad de México,
Mexico

al216618@alumnos.uacj.mx, hsossa@cic.ipn.mx
{hochoa, vianey.cruz, overgara}@uacj.mx

Abstract. Digital mammogram plays a key role in breast cancer screening, with microcalcifications being an important indicator of an early stage. However, these injuries are difficult to detect. In this paper, we propose a lightweight Convolutional Neural Network (CNN) for detecting microcalcifications clusters in digital mammograms. The architecture comprises two convolutional layers with 6 and 16 filters of 9×9, respectively at a full scale, a global pooling layer that eliminates the flattening and dense layers, and a sigmoid function as the output layer for binary classification. To train the model, we utilize the public INbreast database of digital mammograms with labeled microcalcification clusters. We used data augmentation techniques to artificially increase the training set. Furthermore, we present a case study that encompasses the utilization of a software application. After training, the resulting model yielded an accuracy of 99.3% with only 8,301 parameters. This represents a considerable parameter reduction as compared to the 67,797,505 used in MobileNetV2 with 99.8 % accuracy.

Keywords. Microcalcifications clusters detection, shallow convolutional neural network, deep learning.

1 Introduction

Breast cancer is a significant public health challenge, with the highest incidence among women [14].

The detection of small calcium deposits from 0.1 mm to 1 mm in length called Microcalcifications (MCs) [4], plays a vital role in identifying early breast cancer, leading to a 99% survival rate at 5 years or more [3]. Microcalcifications clusters (MCCs) are conformed by at least three MCs per cm^2. These lesions are present in up to 50% of the confirmed cancer cases [29, 36, 37].

The detection of MCs is a complex process due to their size, shape, and distribution [11]. Among the medical imaging techniques, mammography is the most widely used to detect MCCs [4, 6]. The use of Artificial Intelligence (AI) techniques is safe and reliable [9] and can be used to detect the initial signs of diseases [12].

Among these techniques, the Deep Learning (DL) models [21] have achieved high degrees of accuracy and Convolutional Neural Networks (CNNs) are being studied in the field of MCCs detection [4]. As CNN architectures evolve, they have become more complex and deeper.

Hence, the complexity has posed challenges, particularly in medical entities where resource-intensive models for diagnosis can be impractical. A solution is to develop lighter CNN architectures where training and/or retraining times can be minimized, making the network more accessible and efficient, all while requiring
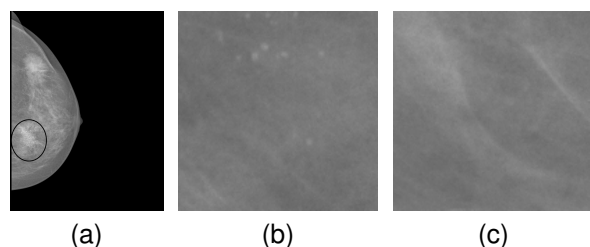
(a)   (b)   (c)

Fig. 1. Digital mammogram showing (a) a circled MCCs, (b) a patch of tissue with MCCs, and (c) with normal tissue

Table 1. Most representative architectures yielded by the Hyperband search algorithm

CL	Filter size	Number of filters	MPL	Parameters	Accuracy
2	5×5	CL1: 6 / CL2: 16	2	2,589	99.1%
2	5×5	CL1: 6 / CL2: 16	0	2,589	99.1%
2	5×5	CL1: 4 / CL2: 10	0	1,125	98.8%
1	5×5	CL1: 16	0	433	97.8%
6	5×5	CL1 - CL6: 4	0	2,129	99%
2	3×3	CL1 - CL2: 16	0	957	98%
2	7×7	CL1: 6 / CL2: 16	0	5,037	99.1%
2	11×11	CL1: 6 / CL2: 16	0	12,381	99.3%
2	**9×9**	**CL1: 6 / CL2: 16**	**0**	**8,301**	**99.3%**

fewer computational resources. In light of the challenges exposed, we present a novel approach incorporating a lightweight and shallow CNN for detecting the presence or absence of MCCs in digital mammograms.

This research builds upon the foundations laid in our prior work [19], representing a continuation and refinement of our previous findings. The paper makes significant contributions, which can be outlined as follows:

– A lightweight CNN specifically designed for the detection of MCCs in digital mammograms using a reduced number of parameters. The network's efficiency is attributed to its notably reduced number of parameters, making it an attractive and practical solution for medical entities seeking efficient MCCs detection.

– A case study of the proposed model. We are primarily concerned with the theoretical and practical applications of our model. Therefore, we developed a software application to detect MCCs. The application is being evaluated by expert radiologists.

The article is organized as follows: Section 2 reviews the related work. Section 3 outlines materials and methods. Section 4 presents the results. Section 5 discusses outcomes. Lastly, Section 6 offers conclusions.

2 Related Work

Efforts to improve accuracy are the main driver behind recent trends in the detection of MCCs. Here, we briefly review the works we consider the most significant because they put our work into context. Gómez et al. [10] proposed a methodology for preprocessing 832 digital mammograms specifically from the mini-MIAS [31] and the UTP [7] databases.

This CNN model comprises seven Convolutional Layers (CL) with a kernel size of 3×3. Following each CL, a Max Pooling Layer (MPL) and a layer of Rectified Linear Unit (ReLU) activation functions were incorporated. The CNN achieved a testing accuracy of 95.83%.

Rehman et al. [25] proposed a Fully Connected Deep-Separable CNN (FC-DSCNN) for detecting and classifying MCCs as benign or malignant. The system involves four steps including image processing, grayscale transformation, suspicious region segmentation, and MCCs classification.

They tested the system on 6,453 mammograms from the public DDSM [27] dataset and from the private Punjab Institue of Nuclear Medicine (PINUM) database, achieving results with 99% sensitivity, 82% specificity, 89% precision, and 82% recall.

Hsieh et al. [11] implemented a VGG-16 network to detect MCCs in 1586 mammograms from the Medical Imaging Department of the Chung-Shan Medical University. They used a Mask R-CNN for MCC segmentation and InceptionV3 for MCC classification (benign or malignant).

Table 2. Tuned hyperparameters and optimal values via Hyperband method

	Hyperparameter	Evaluated values	Best value
CL1	Number of filters	4, 6, 8, 10, 12	6
	Filter size ($n \times n$)	n = 3, 5, 7, 9, 11	9
CL2	Number of filters	16, 20, 24, 28, 32	16
	Filter size ($n \times n$)	n = 3, 5, 7, 9, 11	9
	Batch size	16, 32, 64, 128	64
	Learning rate	0.01, 0.001, 0.0001	0.001

The method achieved a 93% accuracy for classification and detection, 95% for MCs labeling, and 91% for MCC classification. The overall precision, specificity, and sensitivity were 87%, 89%, and 90%, respectively.

Valvano et al. [35] developed two CNNs for the detection and segmentation of Regions of Interest (ROIs) or patches containing MCs. They employed a private database consisting of 283 mammograms with a resolution of 0.05 mm.

Each patch was labeled positive if it contained MCs and negative if it did not. The presence or absence of MCs in each patch was then detected using a CNN. Both CNNs were constructed with six CLs. They achieved an accuracy of 98.22% for the detector and 97.47% for the segmenter.

The most intuitive idea to improve accuracy is to use deeper CNNs. This requires a lot of time to train and use it. There is a clear sacrifice of computational complexity and, in some cases, an incipient gain in precision. Recently, Luna et al. [19] showed that, for MCCs detection, very deep CNN performed similarly to the shallow ones.

They compared different CNNs, in the state-of-the-art, used for classification purposes and found that the networks yielded accuracies between 99.71% and 99.84%. Therefore, for this type of lesion, shallow networks with a reduced number of parameters can be designed to be accommodated in little hardware.

To the best of our knowledge, among these networks, only the VGG-16 architecture has been employed for MCCs detection [11]. Nevertheless, the authors did not report any comparison with other DL networks or structures, lacking sustain the use of this network for this type of lesions.

3 Materials and Methods

In this section, we present an overview of the materials utilized and the methods adopted to investigate MCCs detection in digital mammograms using CNNs.

3.1 Data

We used the INbreast database [22] for training, validating, and testing the model. It comprises 410 grayscale digital mammograms of 2,560 × 3,328 and 3,328 × 4,084 pixels, each pixel is 70 microns. The mammograms are labeled with various types of lesions. In this study, we selected exclusively the ten mammograms labeled as MCC in the database.

3.1.1 Data Preparation

We converted the Digital Imaging and Communication In Medicine (DICOM) images database into Portable Network Graphic (PNG) format. The labeling and coordinates of the breast lesions were available in separate Extensible Markup Language (XML) files and independently associated with the images.

In order to accurately mark the MCCs on the digital mammograms, we developed a custom software, in Python 3.0, to read and extract the MCCs coordinates from XML files for precise localization and annotation of these lesions within mammograms.

3.1.2 Patch Extraction

The proposed model processes mammograms in patches of 1 cm^2 equivalent to 144 × 144 pixels as those shown in Figs. 1 (b) and (c). We developed another dedicated computer program in Python 3.0 to extract annotated patches from the mammograms.

In total, 1,576 patches with MCCs and 1,692 patches without lesions were selected. The initial CNN training sessions were conducted using the dataset [22] as it is. The results were not as expected in all the tested architectures [19]. We asked an expert radiologist to clean our database. She noticed that some patches, labeled as MCCs,
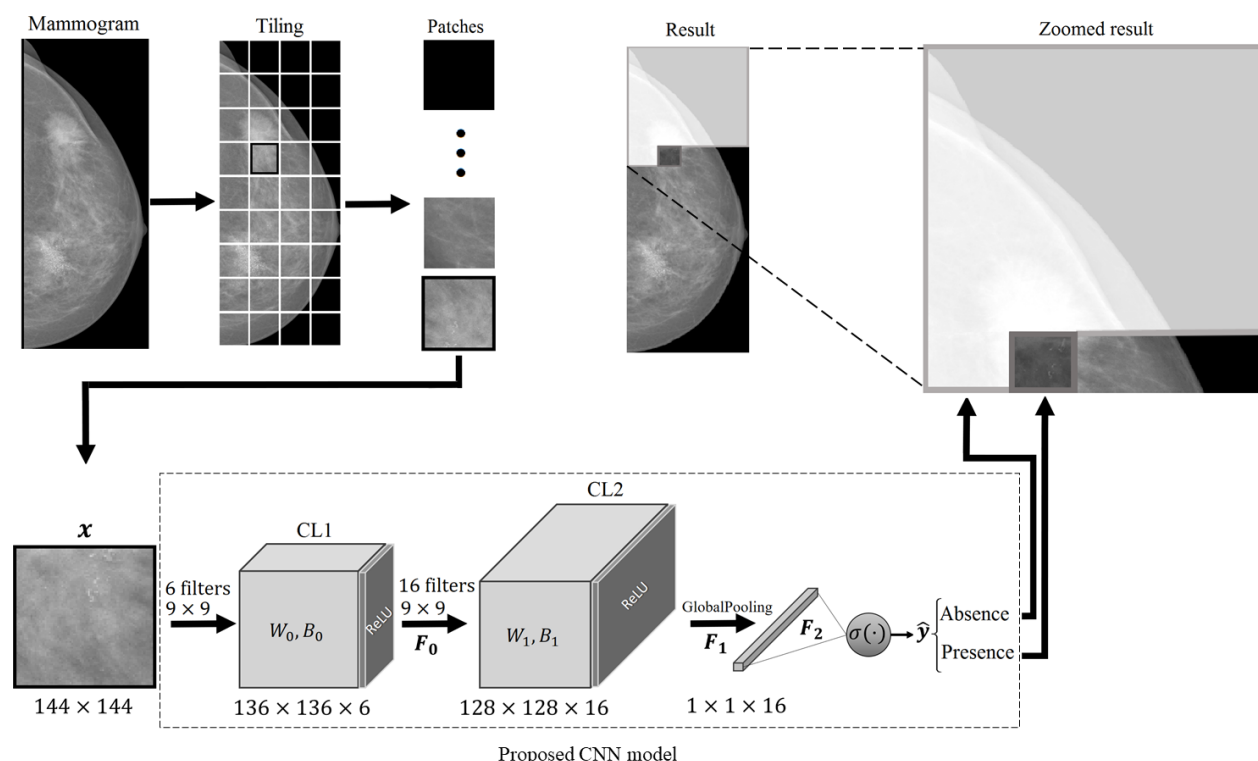
Fig. 2. Case study of the proposed CNN model. If the patch is classified as absence of MCCs, it becomes lighter, and if it is classified as presence the patch darkens

did not contain MCCs, and some unlabeled ones did contain them. Now, with the cleaned database the results exceeded 98% on accuracy [19].

3.1.3 Data Augmentation

The availability of mammograms labeled with MCCs in the INbreast database is limited. Since DL models depend on the quantity and contextual meaning of training data, we artificially increased the number of examples in the database by applying reflection, 180° turn, reflection and 180° turn, and 90° turn, to each patch to obtain 6,304 extra patches with MCCs and 6,768 extra patches without MCCs.

Notice that only geometric transformations were applied to preserve the original features. Consequently, we ended up with a total of 7,880 patches with MCCs and 8,460 patches without MCCs, resulting in a comprehensive dataset of 16,340 patches.

3.1.4 The Datasets

When training a DL model, it is very important to have a dataset with almost the same number of samples in each class. This prevents the model from becoming biased toward one class.

Hence, 7,880 patches with MCCs and 7,880 patches with normal tissue from the database were used. By Pareto's Principle [2], from the dataset we assigned 80% of the data for both training and validation, while the remaining 20% for testing purposes.

More specifically, we utilized 64% (10,088 patches) for training and 16% (2,520 patches) for validation, and for testing, we reserved the remaining 20% (3,152 patches).

To ensure consistency, all patches were normalized by dividing their pixel values by 255. Notice that the data augmentation process was applied to each dataset individually to avoid overfitting.
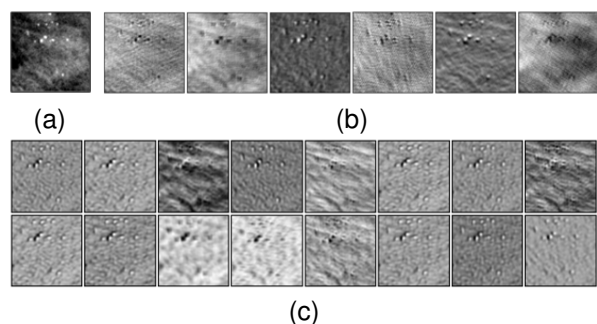
Fig. 3. Patch x (a) with MCCs and feature maps (b) F_0 and (c) F_1 (see Fig. 2)
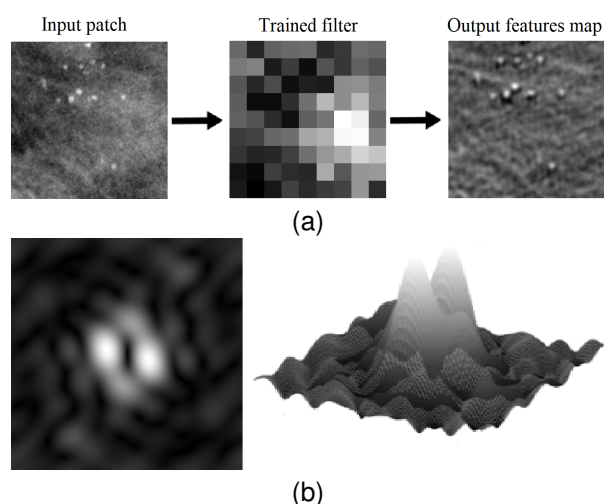

Fig. 4. Filtering process of (a) the third trained filter of $CL\,1$ and (b) magnitude response of the filter

3.1.5 The Proposed Architecture

The proposed architecture was conceived on the premise that biological models of MCs and their surrounding tissue exhibit a reduced number of features [38]. The MC is modeled as a sum of Gaussian functions [38] with limited frequency support (from 0.1 to 1 millimeter) [4].

Therefore, we concluded that it is unnecessary to use a very deep CNN to classify MCCs. This was demonstrated in [19] where CNN models like LeNet-5 [16] with only 5 layers or AlexNet [15] with 8 layers can effectively detect MCCs with the same accuracy.

Besides, these two networks were specifically designed to classify numbers and natural images with a large set of features. Furthermore, in the literature, the current networks are pre-trained on natural images [20]. Hence, it is essential to capture a greater number of low- and high-level features. In the reported works on MCCs detection and classification [24, 18, 23, 26, 28], there is a notable absence of experiments.

The authors typically bring the knowledge of a pre-trained CNN to their own domain by retraining it to observe the prediction or classification results regardless of the depth of the network. However, models of MCCs proposed from biological analyses [38] report that these lesions have a limited number of features, often described as a sum of Gaussian functions.

Therefore, we decided to experiment with one convolutional and one MPL first. Then, we increased the number of layers and noticed that, after two or more layers, the performance was similar. Afterward, we experimented by suppressing the Pooling Layers (PLs) and noticed an improved performance.

Finally, we replaced the Flattening and FCLs with a Global Max Pooling Layer (GMPL) and noticed that the performance was not compromised. However, the number of parameters drastically decreased. Finally, for training, Hyperband search [17] was used to tune the hyperparameters. Table 1 shows the most representative combinations yielded by the algorithm. We propose the lightweight CNN depicted in the case study of Fig. 2.

Each model was trained using TensorFlow framework 2.0 [1] in Google Colaboratory [5]. The platform automatically adjusted the computer resources as needed. For instance, in the latest session, the model accessed a 108GB hard drive, an Intel Xeon (R) CPU @ 2.20GHz processor, and 13GB of memory.

Notice that, we will call the architecture to the structure of the CNN (number of layers, how they are connected, and the type of activation function) and the model to the function that the CNN is approximating after training. The architecture consists of two CLs with a ReLU layer at the output of each, followed by a GMPL.
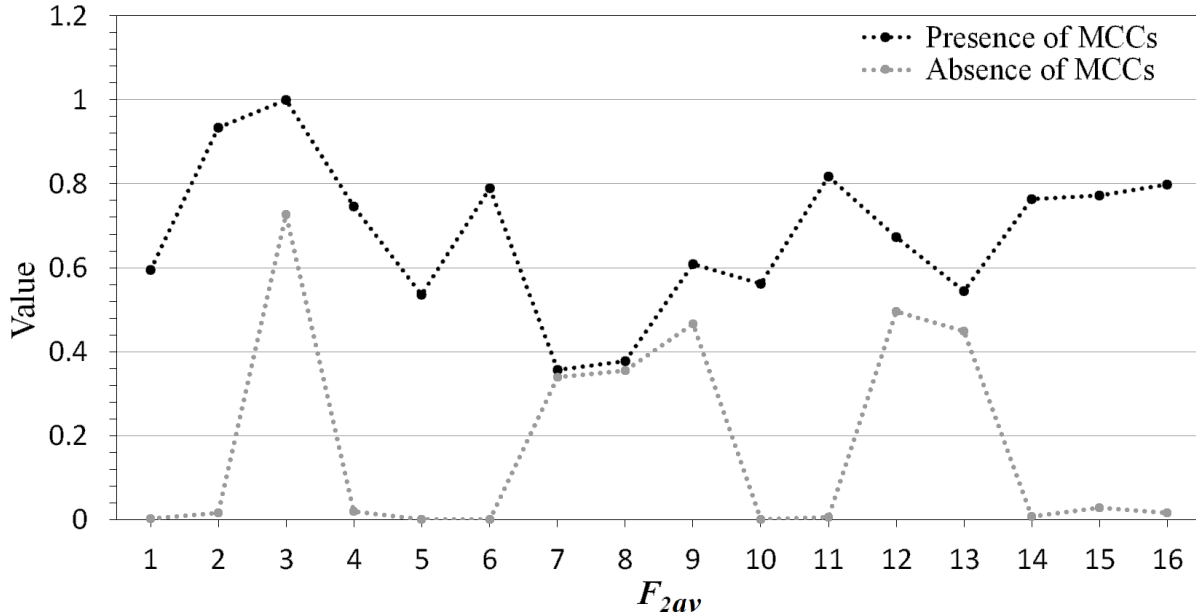
Fig. 5. Plots of the element-wise average of the components of the F_2 classified as presence (black dotted graph) and absence of MCCs (gray dotted graph)

The output layer consists of a sigmoid function. The two CLs are used at full scale, that is, no PLs are inserted to reduce dimensionality. The Binary Cross Entropy (BCE) cost function used is shown in Eq. (1):

$$L(y, \hat{y}) = -\frac{1}{m} \sum_{i=1}^{m} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)], \quad (1)$$

where $-\dfrac{1}{m} \sum_{i=1}^{m}$ is the average loss of the whole batch, m denotes the training set size, y_i is the label, taking binary values 0 or 1 and $\hat{y}_i$ is the predicted value. $-1/m$ ensures that the cost is always greater or equal to 0.

3.1.6 Hyperparameter Tuning

Searching for optimal hyperparameters was a challenge because of the limited computational resources. Hence, we employed the Hyperband search method [17] for hyperparameters tunning by exploring the number and filter sizes, batch size, and learning rate within a relatively narrow range of options.

We used dropout regularization with a permanency of 80% throughout the training process and Adaptive Moment Estimation (ADAM) regularization. Table 2 shows the values of the hyperparameters evaluated by the method along with the best results.

3.2 The Proposed Model

From the previous section, the resulting CNN model consists of two CLs, each followed by a ReLU layer. The first layer has 6 filters of size 9×9, denoted by W_0 with biases B_0. The output is represented as:

$$F_0 = \max(0,\ W_0 \cdot x + B_0), \quad (2)$$

where $\max(0,\ z)$ denotes the largest value between zero and z. Similarly, the second layer comprises 16 filters of size 9× 9, denoted by W_1 with biases B_1. The output can be modeled as:

$$F_1 = \max(0, W_1 \cdot x + B_1). \quad (3)$$

The resulting 16 feature maps are sent to a GMPL to obtain the maximum value of each map

Table 3. Performance comparison of the proposed CNN versus MobileNetV2 and LeNet-5

Architecture	Accuracy	Parameters
MobileNetV2	**99.8%**	67,797,505
LeNet-5	99.3%	2,233,365
Proposed	99.3%	**8,301**

to yield a vector of 16 features represented as $F_2 = \max(F_1)$. The F_2 vector is sent to the output layer where a predicted value between 0 and 1 is assigned according to the vector values. The proposed CNN model is shown in Fig. 2.

3.2.1 Software Application

We developed a web-based software application to test the model's ability to analyze digital mammograms in real time with the domain used to train the network (INbreast database [22]). The user interface allows to import digital mammograms in a PNG format. The software extracts progressively 1 cm^2 patches from the mammogram scanning it from top to bottom and from left to right.

The patch undergoes analysis by the proposed model that yields results between 0 and 1. A near 0 result indicates the absence of MCCs, prompting the application to display the patch in a light gray color. Conversely, a result close to 1 indicates the presence of MCCs, displaying the patch as it is. The application can be configured to display the patch with a color depending on the class it belongs to.

Additionally, counters for each class are maintained to display the number of patches found with and without MCCs during the scanning. The application is hosted on a local server equipped with a 100GB hard drive, an Intel Xeon (R) CPU @ 2.20GHz processor, and 8GB of memory.

Debian [30] serves as the operating system, Apache 2 [32] as the HTTP server, and PHP 8 [34] as the backend. PHP handles tasks such as uploading mammograms to the server, removing the black background, and splitting images into patches for analysis.

Angular v14 [8] is used as the frontend, fetching patches from the backend and utilizing a web service to implement the proposed model. The application's aesthetic is styled using the Bootstrap library [33].

3.2.2 Case Study

Fig. 2 shows a case study implemented for the proposed model. The input mammogram is split into patches of 144 $\times$144 pixels. The coordinates of each patch are stored and the patch ***x*** is sent to the trained CNN model where it undergoes classification. The classified patch is seamlessly integrated back into the mammogram at its original location with a different grayscale that depends on the output classification result $\hat{\boldsymbol{y}}$.

The result is shown in a displayed mammogram with detected normal tissue in light gray and injured tissue in dark gray. The transformation can be inverted anytime to show the original image. This case study was implemented in a software application that is under test by the Centro de Imagen e Investigación (Medimagen) of Chihuahua, México [13].

4 Results

This section exposes the results of the proposed CNN. All the models were trained with 100 epochs. Fig. 3 shows (a) one patch with MCCs that undergoes prediction, (b) the six feature maps F_0 yielded by the first CL, and (c) the sixteen feature maps F_1 at the output of the second CL.

Fig. 4(a) shows the convolution process of the input patch with MCs with the third trained filter of the $CL\,1$. The brightest pixels represent the parts of the spectrum with the highest magnitude. Fig. 4(b) shows the magnitude response of this filter. Observe the limited frequency support.

Fig. 5 shows two plots of the element-wise average output of the sixteen components of the vector F_2, the upper graph, represented by the dotted black line, is the average of the prediction of one hundred patches classified as presence. The light gray dotted line is the element-wise average of the prediction of one hundred patches classified as absent.
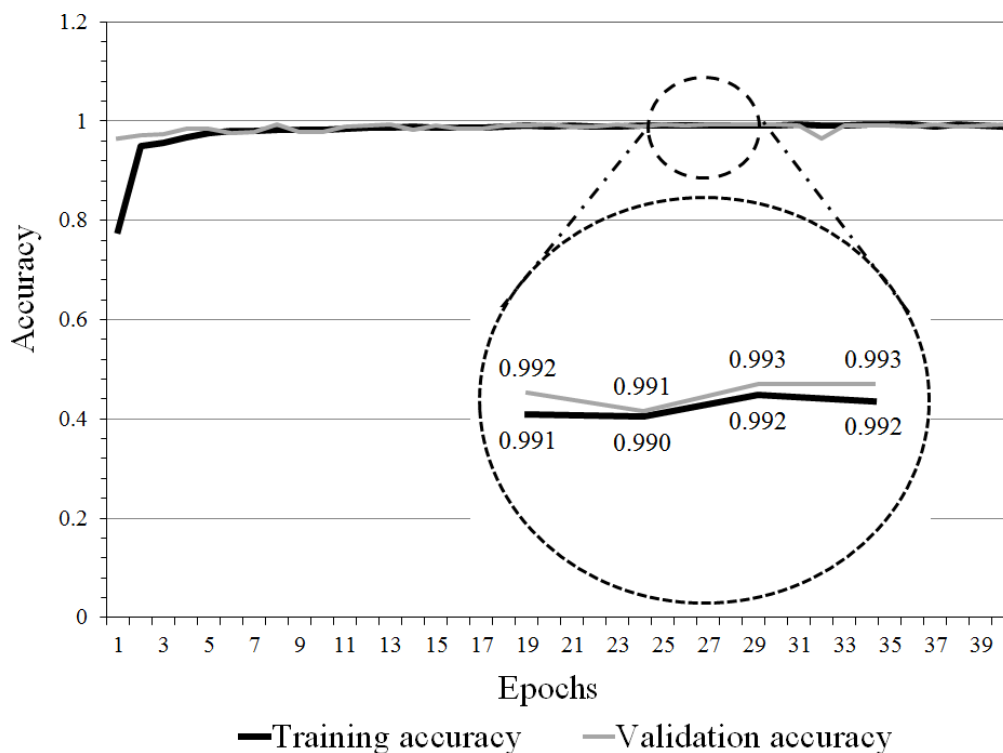
Fig. 6. Proposed model performance accuracy for (black line) training and (gray line) validation

In other words:

$$F_{2\,\mathrm{av}} = (F_{2,1} \cdot + F_{2,2} \cdot + \cdots + F_{2,100}) \cdot /100, \quad (4)$$

where $\cdot+$ and $\cdot/$ are the element-wise sum and division operations and $F_{2,i}$ the ith vector after each prediction. Table 3 presents a comparison of both accuracy and the number of trainable parameters among the proposed model and the MobileNetV2 and LeNet-5 networks.

In [19], MobileNetV2 demonstrated the highest precision in detecting MCCs, while the LeNet-5 network exhibited the fewest number of trainable parameters. Observe that both, the MobileNetV2 and the LeNet-5, were trained from scratch using the same datasets as in the proposed model was trained. Fig. 6 shows the accuracy performance throughout the configured epochs for both the training and the validation processes.

It is important to mention that an expert radiologist corroborated the testing results by using the software application developed.

5 Discussion

In Fig. 3 (b) we notice that, in the first, second, fourth, and sixth maps (from left to right), the MCs locations appear in a pitch black with a rounded feature. Smaller MCs locations are more noticeable in the first and second maps. However, larger MCs locations are detected on the second, fourth, and sixth maps.

These maps separate the MCs leaving only the information of the surrounding tissue. The third and fourth maps highlight the features of the MCs being more prominent on the third map. Besides, the surrounding tissue is attenuated leaving only the MCs features.

Furthermore, Fig. 3 (c) shows a higher level of features. However, we can still see that, from left to right and top to bottom, the third, fifth, eighth, eleventh, twelfth, and thirteenth maps carry the tissue features, and the remaining maps are the MCs features.

The proposed CNN identifies and separates in the feature maps the various characteristics in a patch. To save parameters, a GMPL is added to the output of the second layer. Fig. 5 shows two plots $F_{2\,\mathrm{av}}$ corresponding to the averaged elements of each output F_2 as explained in the previous section.

Notice how the two plots do not overlap each other, this means that on average, there is no overfitting in the network. It is important to observe that ten feature maps yield results close to zero when MCCs are absent and results greater than 0.5 when MCCs are present. Here, the third feature map yields a result greater than 0.5 when MCCs are absent.

However, the same map yields a value close to one when MCCs are present. Additionally, feature maps 7 and 8 give results close to the overlap. Nevertheless, on average, the results are separated. Fig. 6 shows that training and validation performance are not separated from each other.

In fact, they maintain the same tendency. This suggests that there is no overfitting. Table 3 shows that our network achieves comparable accuracy to LeNet-5 CNN with the notable advantage of being 268 times smaller. Moreover, observe that the MobileNetV2 CNN yields an accuracy that is only 0.5% higher than the proposed network. However, the proposed network is 8,167 times smaller.

The MCs range from 0.1 to 1 mm [4] and the scanner used to collect the INbreast database has a resolution of 70 microns per pixel in both directions (horizontal and vertical) [22].

Therefore, an MC varies in size from approximately 2 to 14 pixels which indicates a limited frequency support (from $|1/14|$ to $|1/2|$) as shown in Fig. 4(b) where the bandpass region is delimited by the size of the MC, which clearly indicates that this filter is trained to capture the support.

Moreover, within this region of MCs support, there are other signals that are not MCs as shown in the output features map of Fig. 4(a). Nevertheless, these extra features will be discriminated by the $\mathrm{CL}\,2$.

6 Conclusions

In this paper, we propose a lightweight CNN for detecting MCCs in digital mammograms. The input layer has 6 filters of size 9×9 with ReLU activation functions to have a 6-dimensional feature maps. The second layer performs a nonlinear mapping using 16 filters of size 9×9 with ReLU function.

No PL was added to reduce the dimensionality of the CLs. A GMPL is added to reduce the number of parameters and transform the last 16-dimensional feature maps into a 1D vector. For binary classification, the last layer is a sigmoid function. The resulting model comprises 8,301 parameters making it easily implementable across various frameworks. The achieved accuracy aligns with results from the LeNet-5 and the even more intricate MobileNetV2.

The application developed for our model is under test by the Centro de Imagen e Investigación (Medimagen) of Chihuahua, México. A noteworthy discovery by the expert radiologist, while using the application, was that the model can identify MCCs that initially were not labeled in the INbreast database. This is because the unmarked MCCs were challenging to observe without the support of the application, and the almost imperceptible MCCs often turn out to be malignant.

The ongoing aspect of this research involves developing a faster residual CNN with enhanced performance. Then, the proposed model in this research serves as a foundation for the new CNN. In addition, other types of layers such as the depthwise separable convolutional layers are also being tested. Because of the simplicity of our CNN, we are developing a framework to include explainability in the model. In addition, we are collecting a database of Mexican mammograms, labeled by expert radiologists with several types of lesions that can be used to train new models of DL to work in hospitals and clinics of the country.

Acknowledgments

References

1. **Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., et al. (2015).** TensorFlow: Large-scale machine learning on heterogeneous systems. DOI: 10.48550/arXiv.1603.04467.

2. **Abdelaziz-Ismael, S. A., Mohammed, A., Hefny, H. (2020).** An enhanced deep learning approach for brain cancer mri images classification using residual networks. Artificial Intelligence in Medicine, Vol. 102, pp. 101779. DOI: 10.1016/j.artmed.2019.101779.

3. **American Cancer Society (2023).** Tasas de supervivencia del cáncer de seno.

4. **Basile, T. M. A., Fanizzi, A., Losurdo, L., Bellotti, R., Bottigli, U., Dentamaro, R., Didonna, V., Fausto, A., Massafra, R., Moschetta, M., Tamborra, P., Tangaro, S., La-Forgia, D. (2019).** Microcalcification detection in full-field digital mammograms: A fully automated computer-aided system. Physica Medica, Vol. 64, pp. 1–9. DOI: 10.1016/j.ejmp.2019.05.022.

5. **Bisong, E. (2019).** Building machine learning and deep learning models on Google Cloud Platform: A comprehensive guide for beginners. Apress Berkeley, CA. DOI: 10.1007/978-1-4842-4470-8.

6. **Cronin, K., Scott, S., Firth, A., Sung, H., Henley, S. J., Sherman, R. L., Siegel, R., Anderson, R., Kohler, B., Benard, V., Negoitia, S., Wiggins, C., Cance, W., Jemal, A. (2018).** Annual report to the nation on the status of cancer, part i: National cancer statistics. Cancer, Vol. 128, No. 24, pp. 4251–4284. DOI: 10.1002/cncr.34479.

7. **Echeverry-Correa, J. D., Orozco-Gutiérrez, A. A., Cárdenas-Peña, D. A., Marín-Mejía, S. (2023).** Recuperación de información por contenido orientada a la clasificación de grupos de microcalcificaciones en mamografías - Protocam. Universidad Tecnológica de Pereira. DOI: 10.22517/9789587225174.

8. **Google (2023).** The web development framework for building the future.

9. **Henriksen, E. L., Carlsen, J. F., Vejborg, I. M. M., Nielsen, M. B., Lauridsen, C. A. (2018).** The efficacy of using computer-aided detection (CAD) for detection of breast cancer in mammography screening: a systematic review. Acta Radiologica, Vol. 60, No. 1, pp. 13–18. DOI: 10.1177/0284185118770917.

10. **Hernández-Gómez, K. A., Echeverry-Correa, J. D., Orozco-Gutiérrez, A. A. (2021).** Automatic pectoral muscle removal and microcalcification localization in digital mammograms. Healthcare Informatics Research, Vol. 27, No. 3, pp. 222–230. DOI: 10.4258/hir.2021.27.3.222.

11. **Hsieh, Y. C., Chin, C. L., Wei, C. S., Chen, I. M., Yeh, P. Y., Tseng, R. J. (2020).** Combining VGG16, Mask R-CNN and inception V3 to identify the benign and malignant of breast microcalcification clusters. IEEE International Conference on Fuzzy Theory and Its Applications (iFUZZY), pp. 1–4. DOI: 10.1109/iFUZZY50310.2020.9297809.

12. **IBM (2023).** DREAM challenge results: Can machine learning help improve accuracy in breast cancer screening?

13. **ImagenologIA (2023).** Microcalcification clusters detection model in real time.

14. **International Agency for Research on Cancer (2023).** Breast.

15. **Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012).** Imagenet classification with deep convolutional neural networks. 26th Annual Conference on Neural Information Processing Systems. Advances in Neural Information Processing Systems, pp. 1106–1114.

16. **Lecun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998).** Gradient-based learning applied

to document recognition. Proceedings of the IEEE, Vol. 86, No. 11, pp. 2278–2324. DOI: 10.1109/5.726791.

17. **Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., Talwalkar, A. (2017).** Hyperband: A novel bandit-based approach to hyperparameter optimization. The Journal of Machine Learning Research, Vol. 18, No. 1, pp. 6765–6816.

18. **Liu, H., Chen, Y., Zhang, Y., Wang, L., Luo, R., Wu, H., Wu, C., Zhang, H., Tan, W., Yin, H., Wang, D. (2021).** A deep learning model integrating mammography and clinical factors facilitates the malignancy prediction of BI-RADS 4 microcalcifications in breast cancer screening. European Radiology, Vol. 31, No. 8, pp. 5902–5912. DOI: 10.1007/s00330-020-07659-y.

19. **Luna-Lozoya, R. S., Ochoa-Domínguez, H. J., Sossa-Azuela, J. H., Cruz-Sánchez, V. G., Vergara-Villegas, O. O. (2023).** Comparison of deep learning architectures in classification of microcalcifications clusters in digital mammograms. Mexican Conference on Pattern Recognition, pp. 231–241. DOI: 10.1007/978-3-031-33783-3_22.

20. **Mahardi, Wang, I. H., Lee, K. C., Chang, S. L. (2020).** Images classification of dogs and cats using fine-tuned VGG models. IEEE Eurasia Conference on IOT, Communication and Engineering, pp. 230–233.

21. **Miotto, R., Wang, F., Wang, S., Jiang, X., Dudley, J. T. (2018).** Deep learning for healthcare: review, opportunities and challenges. Briefings in Bioinformatics, Vol. 19, No. 6, pp. 1236–1246. DOI: 10.1093/bib/bbx044.

22. **Moreira, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J., Cardoso, J. S. (2012).** INbreast: Toward a full-field digital mammographic database. Academic Radiology, Vol. 19, No. 2, pp. 236–248. DOI: 10.1016/j.acra.2011.09.014.

23. **Mota, A. M., Clarkson, M. J., Almeida, P., Matela, N. (2022).** Automatic classification of simulated breast tomosynthesis whole images for the presence of microcalcification clusters using deep CNNs. Journal of Imaging, Vol. 8, No. 9, pp. 231. DOI: 10.3390/jimaging8090231.

24. **Rasool, E., Anwar, M. J., Shaker, B., Hashmi, M. H., Rehman, K. U., Seed, Y. (2023).** Breast microcalcification detection in digital mammograms using deep transfer learning approaches. Proceedings of the 9th International Conference on Computing and Data Engineering, pp. 58–65.

25. **Rehman, K. U., Li, J., Pei, Y., Yasin, A., Ali, S., Mahmood, T. (2021).** Computer vision-based microcalcification detection in digital mammograms using fully connected depthwise separable convolutional neural network. Sensors, Vol. 21, No. 14, pp. 4854. DOI: 10.3390/s21144854.

26. **Sabani, A., Landsmann, A., Hejduk, P., Schmidt, C., Marcon, M., Borkowski, K., Rossi, C., Ciritsis, A., Boss, A. (2022).** BI-RADS-Based classification of mammographic soft tissue opacities using a deep convolutional neural network. Diagnostics, Vol. 12, No. 7, pp. 1564. DOI: 10.3390/diagnostics12071564.

27. **Sawyer-Lee, R., Gimenez, F., Hoogi, A., Rubin, D. (2016).** Curated breast imaging subset of digital database for screening mammography (CBIS-DDSM) [data set]. The cancer imaging archive.

28. **Shiri Kahnouei, M., Giti, M., Akhaee, M. A., Ameri, A. (2022).** Microcalcification detection in mammograms using deep learning. Iranian Journal of Radiology, Vol. 19, No. 1. DOI: 10.5812/iranjradiol-120758.

29. **Sickles, E., D'Orsi, C., Bassett, L., et al. (2013).** ACR BI-RADS® mammography. In: ACR BI-RADS® atlas, breast imaging reporting and data system. American College of Radiology.

30. **Software in the Public Interest (2023).** The universal operating system.

31. Suckling, J. (2023). The mini-MIAS database of mammograms.

32. The Apache Software Foundation (2023). HTTP server project.

33. The Bootstrap Team (2023). The most popular HTML, CSS, and JS library in the world.

34. The PHP Group (2023). PHP: Hypertext preprocessor. www.php.net/.

35. Valvano, G., Santini, G., Martini, N., Ripoli, A., Iacconi, C., Chiappino, D., Della Latta, D. (2019). Convolutional neural networks for the segmentation of microcalcification in mammography imaging. Journal of Healthcare Engineering, Vol. 2019, pp. 1–9. DOI: 10.1155/2019/9360941.

36. Wang, J., Nishikawa, R. M., Yang, Y. (2017). Global detection approach for clustered microcalcifications in mammograms using a deep learning network. Journal of Medical Imaging, Vol. 4, pp. 024501. DOI: 10.1117/1.JMI.4.2.024501.

37. Wang, J., Yang, Y. (2018). A context-sensitive deep learning approach for microcalcification detection in mammograms. Pattern Recognition, Vol. 78, pp. 12–22. DOI: 10.1016/j.patcog.2018.01.009.

38. Yang, Y., Yang, Y., Liu, Z., Guo, L., Li, S., Sun, X., Shao, Z., Ji, M. (2021). Microcalcification-based tumor malignancy evaluation in fresh breast biopsies with hyperspectral stimulated raman scattering. Analytical Chemistry, Vol. 93, No. 15, pp. 6223–6231. DOI: 10.1021/acs.analchem.1c00522.