10th International Workshop on Innovative Simulation for Healthcare
18th International Multidisciplinary Modeling & Simulation Multiconference

ISSN 2724-0371 ISBN 978-88-85741-65-2 © 2021 The Authors.
doi: 10.46354/i3m.2021.iwish.009

Mammography CADx System based on Transfer Learning

J.A. Almaraz-Damian[1], Oscar Garcia-Avila[1], Volodymyr Ponomaryov[1,*], Rogelio Reyes-Reyes[1] and Clara Cruz-Ramos[1]

[1]Instituto Politecnico Nacional, Santa Ana Ave. # 1000, Mexico City, 04430, Mexico

[*]Corresponding author. Email address: vponomar@ipn.mx

Abstract

Early detection of breast cancer is crucial in the treatment of this desease, as the principal diagnoses tool mammography imaging is employed due to its no-invasive form. In this paper, a Computer-Aided Detection System (CADx) is presented for the analysis of digital mammogram images. The methods used in the proposed CAD system are Transfer Learning, Support Vector Machine as Classifier, and Feature Reduction based on Principal Component Analysis. The system has demonstrated improved performance in comparison with state-of-the-art methods in terms of quality metrics such as Accuracy, Specificity, Sensibility, and F1-Score.

Keywords: Mammograph; Pseudocolour; CADx; CNN; PCA; SVM; Transfer Learning

1. Introduction

Cancer is the name given to various diseases related to the excessive growth of cells caused by the damage in the reproduction cycle of cells. This alteration is caused by different factors such as lifestyle, environmental or hereditary, driving to the formation of a mass commonly referred to as *tumour* or *neoplasm*. These masses belong to two principal types, **malignant** or **benign** (de Cancerologia (INCan) Secretaria de Salud Gobierno de México; USA.gov).

The **benign** tumours grow at a non-accelerated rate. Additionally, they do not spread or infiltrate neighbouring tissues. On the other hand, **malignant** tumours can spread to nearby tissues or invade through the lymphatic and circulatory systems, creating new tumours in different parts of the body. Usually, when these tumours are extracted, they can reappear in the same place as a difference from the benign ones, which they do not reappear.

Breast cancer is one of the deadly diseases that affect women in Mexico and around the World. According to the World Health Organization (WHO) (OMS/OPS), every year: 1.38 million new cases are detected and 458,000 die from this disease. In Mexico, it represents the leading cause of death in women for every 100000 women diagnosed, 17 die due to complications related (INEGI). Therefore, the importance of early detection could help to eradicate this disease (de Salud Pública).

The imaging studies allow the physicians to make an early diagnosis of Breast cancer based on the perceptual view of the mass aided of a mammography (MG), which is considered the most practical method to obtain a visual representation of the suspicious area and is accessible to the general population due to its low cost and fast acquisition.

Table 1. BI-RADS System Summary

BI-RADS Category	*Definition*	*Action*
0	***Incomplete***	Additional imaging evaluation and/or comparison to prior mammograms is needed
1	***Negative***	There's no significant abnormality to report.
2	***Benign***	This is also a negative mammogram result, but the specialist chooses to describe a finding known to be benign.
3	***Probably benign finding***	Follow-up in a short time frame is suggested
4	***Suspicious abnormality***	Biopsy should be considered
5	***Highly suggestive of malignancy***	Appropriate action should be taken
6	***Known biopsy-proven malignancy***	Appropriate action should be taken

If a specialist requires a more detailed analysis of the suspicious area, complementary imaging studies are performed such as Ultrasound (US), Computed Tomography (CT), Magnetic Resonance Imaging (MRI). If those are inconclusive, a *biopsy* is performed, where a portion of breast tissue is extracted to be examined through histopathological studies to determine the grade of malignancy Society.

The **Breast Imaging Reporting and Data System** (BI-RADS) is a scheme that allows classifying the findings obtained in an MG image in a standardized way. Allowing to determine if the breast tissue contains any mass and its type also recognizes the degree of malignancy or benignity of the lesion. Providing for a non-invasive diagnosis and therefore the following process to be established by the specialist for its treatment Aibar et al. (2011); of Radiology (acr), the categorization of the BI-RADS system is detailed in Table 1.

Computer Aided Diagnosis (CAD) systems are capable of performing data processing to assist physicians. These systems are considered intelligent because they use feedback to acquire new knowledge and thus improve their performance. Two types of these systems are considered, those dedicated to the detection of an anomaly, normally those systems are called *Computer Aided Detection* (CADe), or those that are dedicated to determinate a diagnosis called *Computer Aided Diagnosis* (CADx) (Yanase and Triantaphyllou, 2019; Backfrieder and Zwettler, 2015; Zwettler and Backfrieder, 2014).

A CADx system consists of four main stages:

1. **Preprocessing**: At this stage, improvements are made to the image such as contrast adjustments, noise reduction, among others.
2. **Segmentation**: In this stage, the anomaly or lesion to be analyzed should be found, this area is commonly named the region of interest (ROI).
3. **Feature extraction**: Based on the region of interest, features or patterns are extracted to carry out their subsequent classification. Some of the features are based on shape, colour, texture, among others.
4. **Classification and evaluation**: From the extraction of the established features, at this stage, the system must be able to determine the outstanding features for each class, and thus be able to discriminate between one class or another.

In this work, a *CADx System* is proposed for the analysis of mammography images, which classify whether the image presents a malignant or benign lesion. The methods used in the proposed CADx system are Transfer Learning, Support Vector Machines (SVM) and the reduction of features through Principal Component Analysis (PCA). The system obtained favourable results compared to state-of-the-art methods using quality metrics such as *Accuracy*, *Specificity*, *Sensitivity and F-Measure.*

2. Related Works

Tsochatzidis, et al. (Tsochatzidis et al., 2019) used two schemes for binary classification (*benign* and *malignant*) of mammographic images. From the manual extraction of the *Region of interest* (ROI) of the images obtained from the DDSM-400 database, then a resizing of the images is applied. As a feature extractor they used architectures based on *Convolutional Neural Networks* (CNN), such as AlexNet (Krizhevsky et al., 2012), VGG-16/19 (Simonyan and Zisserman, 2014), ResNet (He et al.), GoogLeNet (Szegedy et al., 2015a) and Inception-BN v2 (Szegedy et al., 2015b). In the first scheme, the classifier is replaced and *trained from scratch*, and in the second scheme the classifier is *Fine-Tuned* , that means that uses pre-trained weights and in each image those are updated. The best results were obtained in the second scheme with the ResNet-101 network resulting in *AUC* = 0.859 and *Accuracy* = 0.785.

Ragab, et al. (Ragab et al., 2019b) proposed a CADx system using AlexNet architecture based on *Convolutional Neural Network* attempting to extract features of the DDSM and CBIS-DDSM databases for binary classification of lesions contained in the breast. These authors extract the ROI manually and through a thresholding technique based on the value of the pixels in the image. In the classification stage, they replaced the AlexNet classifier with a *Support Vector Machine* classifier. The best results obtained are: *Accuracy* = 0.872, *AUC* = 0.94, *Sensitivity* = 0.862, *Specificity* = 0.877, *Precision* = 0.88,

F1-Measure= 0.871.

Arora, et al. (Arora et al., 2020) proposed a methodology using the CBIS-DDSM database to classify masses as malignant and benign, analyzing 1318 ROI images contained in this database. They implemented the following CNN architectures: AlexNet, VGG16, GoogLeNet, ResNet-18, InceptionResNet (Szegedy et al., 2017) as feature extractors that are subsequently concatenated to create a feature vector. Finally, they use an *Artificial Neural Network* to classify those features obtaining *Accuracy* = 0.88, *AUC* = 0.88, *Precision* = 0.85, *Completeness* = 0.91.

Lévy, et al. (Lévy and Jain) proposed a CADx system based on AlexNet and GoogLeNet CNN architectures, using the DDSM database. Employing the *Transfer Learning* method and the *Data Augmentation* technique, they achieved *Accuracy* = 0.929 *Precision* = 0.924 and *Completeness* = 0.934.

Comparing the aforementioned systems, most of them use features extracted by a Convolutional Neural Networks with the help of the Transfer Learning technique based on state-of-the-art architectures (Krizhevsky et al., 2012; Simonyan and Zisserman, 2014; He et al.; Szegedy et al., 2015a,b, 2017) used for the classification of Imagenet task (Russakovsky et al., 2015). However, the principal drawback of these methods is the manual extraction of the Region of Interest, which is detrimental to the evaluation of the systems because the knowledge of a specialist is required to determine this region.

3. Methodology

3.1. Proposed CADx

The proposed CADx system is illustrated in Fig. fig:bd1. In the first instance, the mammographic image is *segmented* manually, eliminating artefacts such as labels among others. Subsequently, a *Pseudocolour* technique is performed to highlight the parts of greatest interest contained in the mammographic image. Subsequently, the features are obtained based on a Convolutional Neural Network architecture, which was trained for the Imagenet classification task. Finally, these features are reduced by the *Principal Component Analysis* and classified by a Support Vector Machine. The details of each stage of the proposed system are described below.

3.2. Preprocessing

3.2.1. *Segmentation*

Nowadays, obtaining a mammogram image is a fast and accurate process, however, these images are affected by artefacts such as labels, medical instrumentation or, in a specific case, the appearance of the pectoral muscle. To eliminate these artefacts a manual segmentation was carried out, preserving as much of the breast area, eliminating the aforementioned artefacts (Fig. fig:seg). This process was performed for all images used in this work.

3.2.2. *Pseudocolour*

Pseudo-colouring of greyscale images is a process used to complement visual information from X-ray images, improving the detection of perceptual features, structures or patterns. The principal objective of pseudocolour is to exploit the perceptual capabilities of the Human Visual System (Haindl and Remeš, 2019; Haindl, 2019; Mery and Saavedra, 2020). In this work, a pseudocolour method is applied, improving the image of the lesion denoted by $I_m(x,y)$, obtained in the previous step.

The pseudocolour operation of the lesion image $I_m(x,y)$ is based on the equations:

$$R(x,y) = \left| sin\left(2\pi\left(\frac{I_m(x,y)}{255} + \frac{\pi}{2}\right)\right)\right|, \quad (1)$$

$$G(x,y) = \left| sin\left(2\pi\left(\frac{I_m(x,y)}{255} + \frac{\pi}{4}\right)\right)\right|, \quad (2)$$

$$B(x,y) = \left| sin\left(2\pi\left(\frac{I_m(x,y)}{255} + \frac{\pi}{6}\right)\right)\right|, \quad (3)$$

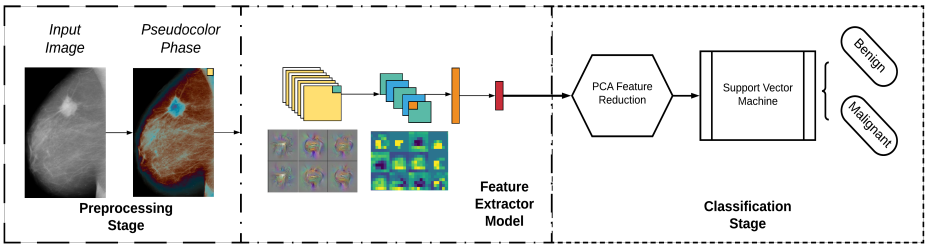
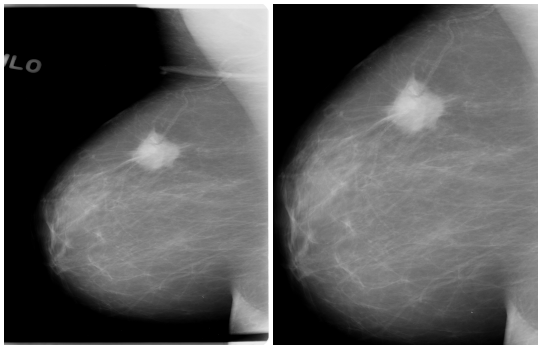


Figure 1. Block diagram of the implemented CADx system.



(a) **(b)**

Figure 2. Results of the segmentation stage: a) Original Image b) Region of interest

Finally, the previous images are concatenated to generate the final image $I_p(x, y)$ as:

$$I_p(x, y) = [R(x, y), G(x, y), B(x, y)]. \quad (4)$$

3.3. Feature Extraction

3.3.1. *Convolutional Neural Network*

Convolutional Neural Network (CNN) is a *Deep Learning* method that consists of two generic blocks: the convolutional block, where the convolutional layers, the pooling layers and the proposed activation functions are found, followed by the classification block, which consists of an Artificial Neural Network differentiating the features obtained in the convolutional block. Normally the first layers of a CNN can detect basic features such as circles, lines or edges and the deeper ones, should detect complex and specific patterns for each class (Albawi, 2017).

3.3.2. *Transfer Learning*

Transfer Learning is a *Machine Learning* technique where a model used to solve a given task is reused to solve a specific task (Yosinski et al., 2014; Pan and Yang, 2010; Weiss et al.). Commonly this technique is used in schemes based on Deep Learning due to the large amount of data required to train these methods from scratch, which is the main problem using these methods. Applying this technique allows developing models in such a way that they are accurate and fast. Two strategies should be considered when using this technique from a pre-trained model: The *Fine-Tuning* and the use of the model as *Feature Extractor*.

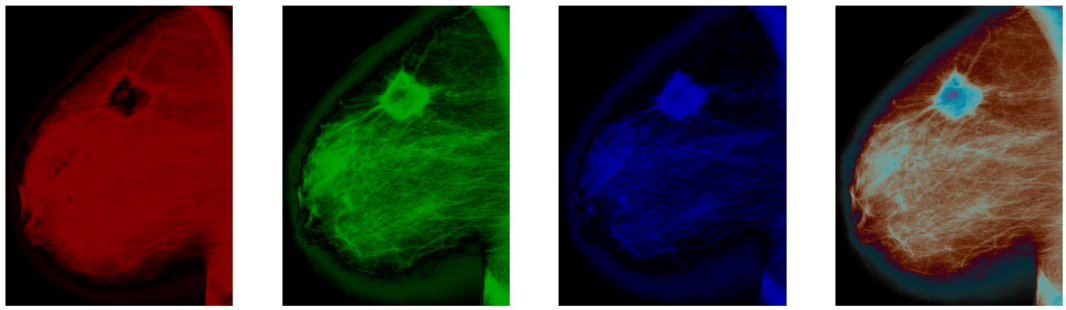
The Fine-Tuning of a model involves retraining it from a specific convolutional layer or convolutional block and replacing its classifier with one that fits a specific task, so this method adjusts the model to the new information that is provided.

Setting the model as a Feature Extractor involves removing the classification block and keeping the last convolutional layer of the model to obtain the features that the model identifies for each class before being classified. These features should be considered generic due to the model to detect them based on their knowledge of a similar task. Finally, to classify these features, a classifier suitable for this specific task must be used (see Fig. 4)

In this work, we employ the Feature Extraction method based on the NasNet (Zoph et al., 2018) architecture, which was trained in the ILSVRC (Russakovsky et al., 2015) classification task.

3.3.3. *NASNet*

NASNet is a Convolutional Neural Network architecture proposed by Zoph et al. (Zoph et al., 2018), for the CIFAR-10 database and latter modified for ILSVRC (Russakovsky et al., 2015) obtaining 82.7% *Accuracy*. This architecture contains two types of cells called normal and reduction cells. Normal cells contain convolutional operations, which return a feature map of the same dimension, while in the reduction cell, the feature map is reduced by a factor of two, reducing the width and height of the feature map obtained. In this study, the architecture *NASNet – A4@64* is used where the first number (4) is the number of cells that are stacked and the second (64) indicates the number of filters located in the penultimate layer of the network finally getting 4032 features.



(a) (b) (c) (d)

Figure 3. Results of the segmentation stage: a) Image generated from the equation 1 b) Image generated from the equation 2 c) Image generated from from the equation 3 d) Image generated from the equation 4

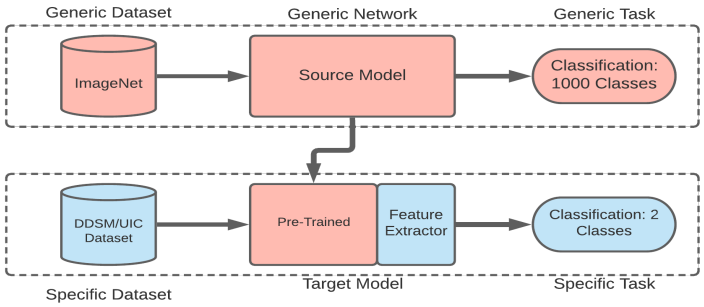


Figure 4. Conceptual view of Transfer of Learning.

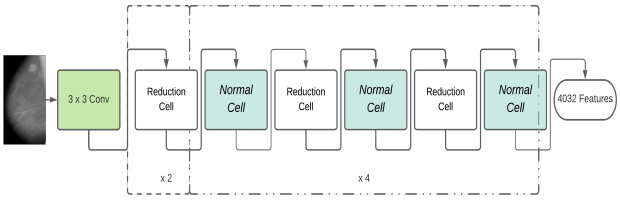


Figure 5. Block diagram of the Convolutional NASNet architecture.

3.3.4. *Grad-CAM*

Proposed by Selvaraju et al. (Selvaraju et al., 2017) allows visualizing the most relevant features of the Convolutional Neural Network after predicting an image. This method obtains the visual representation of the most salient regions involved in the classification phase, using the gradients of the last convolutional layer. This work employs this method to ensure that the architecture detects the extracted features belonging to the ROI.

3.4. Feature Reduction

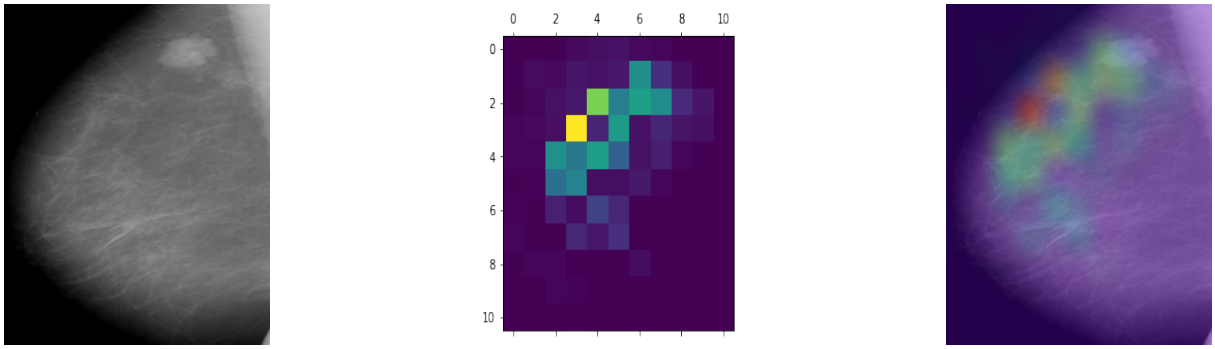
After extracting the deep features, it is necessary to reduce the number of them since a large number increases the computation time to make a prediction. To solve this problem the *Principal Component Analysis* (PCA) is used.

Principal Component Analysis is the statistical method used to transform a *p-dimensions* dataset into another *q-dimensions* dataset called components, projecting the original vector to a lower dimension (de Oliveira et al., 2010).

PCA is used to obtain the optimal attributes for the classification stage where the advantages are in avoiding of overfitting and, also it can improve the accuracy in the prediction (Ragab et al., 2019a).

The steps used in the principal components algorithm are:

- Normalize each characteristic in the dataset with a mean equal to zero and its unit variance.
- Get the covariance matrix of the training dataset.
- Obtain the Eigenvalues and Eigenvectors from the covariance matrix, where the Eigenvectors will provide us with the directions of the Principal Components.



(a) (b) (c)

Figure 6. Results of the visualization stage by Grad-CAM: a) Original Image, b) Activation Map obtained by Grad-CAM, c) Interpolation of the image a) and b).

- Project the data in the test data set to the address of the training dataset PCs.

After applying a variation of the Recursive Feature Elimination technique, which is a wrapper-style feature selection algorithm. It works by searching for a subset of features starting with all features and removes them until the desired number remains. In this work, instead of features, we reduce the number of components, obtaining the optimal number of components is 404.

3.5. Classifier

The classifier used in this work is the *Support Vector Machine*, which belongs to the supervised machine learning techniques. It is based on the concept of the decision planes, where a set of points, who represent the classes, shall be separated by a hyperplane (Cortes and Vapnik, 1995).

In many cases, the dataset cannot be separated by a hyperplane, so a function called kernel should be used. Commonly used kernels are the Linear, Polynomial, Radial Base Function (RBF), the latter assigns the original data to a new identity space, in which separability between classes can be found. In the present work, an SVM with RBF kernel is implemented with parameters $C = 1$ and $\gamma = \frac{1}{\#_{features}}$.

4. Experimental Results

The described method was performed in the Google®Collaboratory platform based on Linux Operative System with 12 GB of RAM and an Nvidia®Tesla K80 GPU with 12 GB VRAM, Python 3.7 with Keras (Chollet et al., 2015), Scikit-learn (Pedregosa et al., 2011) and TensorFlow (Abadi et al., 2015) libraries were used.

4.1. Databases

In this work, two public datasets were used, which are described below:

Dataset CBIS-DDSM (Curated Breast Imaging Subset of DDSM) has 2,620 mammography studies. It contains normal, benign and malignant cases with pathological information verified by specialists. The images are in the *DICOM* format. Images containing mass lesions were used for this work. The number of Images used from the CBIS-DDSM database is shown in Table 2.

Table 2. Number of images contained in the CBIS-DDSM database

Database cases	Number of Images
Benign for Training	370
Malignant for Training	341
Benign for Test	121
Malignant for Test	80
Total	997

UIC Dataset contains 286 images, obtained retrospectively within a Protocol approved by the Review Board of the University of Chicago Medical Center. These images (Table 3) are in greyscale and PNG format (at Chicago (UIC)).

Table 3. Number of images contained in the UIC database

Database cases	*Number of Images*
Benign	111
Malignant	175
Total	286

4.2. Quality Criteria

To evaluate the performance of the proposed system we use the following quality metrics:

Accuracy is the total number of correct predictions

among the total number of samples, the precision is given by:

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}. \quad (5)$$

Rate of true positives or *Sensitivity* is the number of positive cases that were correctly predicted as positive with respect to all positive cases calculated as:

$$Sensibility = \frac{tp}{tp + fn}. \quad (6)$$

Rate of true negatives or *Specificity* is the number of negative cases that were correctly predicted as negative with respect to all negative cases calculated as:

$$Specificity = \frac{tn}{tn + fp}. \quad (7)$$

Precision is the number of correct positive results divided by the number of positive results predicted by the classifier.

$$Precision = \frac{tp}{tp + fp}. \quad (8)$$

F1-Measure characterizes the harmonic mean between precision and Sensitivity.

$$F1 = \frac{2tp}{2tp + fp + fn}. \quad (9)$$

Furthermore, in the classification task, the results can be expressed in a matrix called *Confusion Matrix*, which contains the following information: the classes of the dataset and the total number of elements. Additionally describes the performance of the classification model where the number of cases that the model correctly predicts the positive class are the true positives (*tp*), the number of cases that the model correctly predicts for the negative class are the true negatives (*tn*), the number of cases which the model incorrectly predicts for the positive class is the false positives (*fp*) and finally, the false negatives (*fn*) is the number of cases that the model incorrectly predicts the negative class.

In Fig. 7, a) the confusion matrix of the results obtained using the DDSM dataset is shown, in addition where the following metrics were obtained: *Accuracy* = 0.88, *Sensitivity* = 0.88, *Specificity* = 0.87, *Precision* = 0.87, and a *F1-Measure* = 0.87. In addition, the curve *Area Under the Curve* (AUC) was obtained obtaining a value of 0.94. (see Fig. 7)

For validation of results, the *K-fold* technique was used, which is a common technique to evaluate the system. Here, the data set is divided into *k-splits*, and the classifier is trained using $K - 1$ divisions to estimate an error value, which is the average of the errors made in each division where the classifier was evaluated. In Fig. 8, the Receiver Operating Characteristic is presented for the DDSM database evaluated with the *K-fold* validation technique where, the system shows a *AUC* = 0.97.

Moreover, the system was tested in the UIC database, where achieves an *Accuracy* = 0.97, *Sensitivity* = 0.97, *Specificity* = 0.97, *Precision* = 0.97 and *AUC* = 0.97 and for *k-fold* validation shows an *AUC* = 0.83 showing the robustness of the proposed model on different mammography databases.

In table 4 shows that our system achieves better performance compared to the methods found in the *state of the art*. From our point of view, this is due to three main factors: is the use of pseudocolour to exploit the attributes of Convolutional Neural Networks, second the use of the NASNet network as a feature extractor and its correct detection in the ROI, visualized by the Grad-CAM algorithm, and finally, the reduction of features obtained from the principal component analysis.

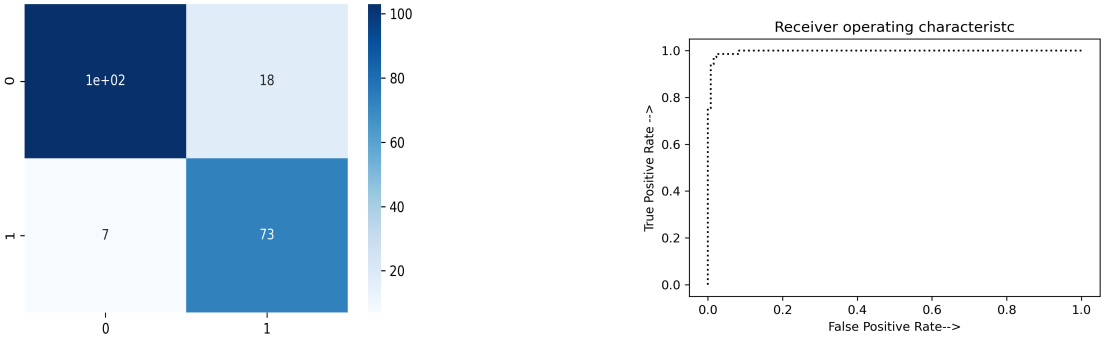
5. Conclusions

In this work, the design of a CADx system for the classification of breast cancer in mammography images was presented, where a Pseudocolour technique based on Sine function was employed to enhance the perceptual information of the Feature Extraction method based on Convolutional Neural Networks.

By evaluating this method using the quality metrics we obtained *Accuracy* = 0.88, *Sensitivity* = 0.88, *Specificity* = 0.87, *Precision* = 0.87 and *AUC* = 0.94 for the DDSM data set . In addition, to check the robustness of the novel method, we proceeded to classify another set of data, such as the UIC, achieving *Accuracy* = 0.97, *Sensitivity* = 0.97, *Specificity* = 0.97, *Precision* = 0.97 $F1 - Measure = 0.97$ and *AUC* = 0.97, which are competitive with the methods found in the *state-of-the-art* for classification of lesions from mammography images.

A possible limitation of the proposed method is the aggregation of new data and their nature, for example, if a physician requires to examine an image with a different view (obliques or horizontal) the system should be retrained, therefore, needs for a vast quantity of images with the same view for the training process. The same limitation applies if the physician requires to add a lesion or any medical condition not established in this work.

As future work consists of the development of a CADx system approach employing a Convolutional Neural Network trained from scratch to determine the BI-RADS stage where a breast lesion belongs and the implementation of those systems in public health hospitals.



(a) (b)

Figure 7. Results of the proposed CADx system: a) Confusion matrix, b) Area under the curve.

Table 4. Comparison with the state of the art.

	Tsochatzidis et al. (2019)	Ragab et al. (2019b)	Arora et al. (2020)	Lévy and Jain	**Proposed CADx System**		
Dataset	DDSM-400	DDSM	DDSM	DDSM	DDSM	UIC	AVG
Accuracy	0.78	0.872	0.88	0.929	**0.88**	**0.97**	**0.92**
Sensibility	N/A	0.862	0.91	0.934	**0.88**	**0.97**	**0.92**
Specificity	N/A	0.877	N/A	N/A	**0.87**	**0.97**	**0.92**
Precision	N/A	0.88	0.85	0.924	**0.87**	**0.97**	**0.92**
F1-Measure	N/A	0.871	N/A	N/A	**0.87**	**0.97**	**0.92**
AUC	0.859	0.94	0.88	N/A	**0.94**	**0.97**	**0.95**

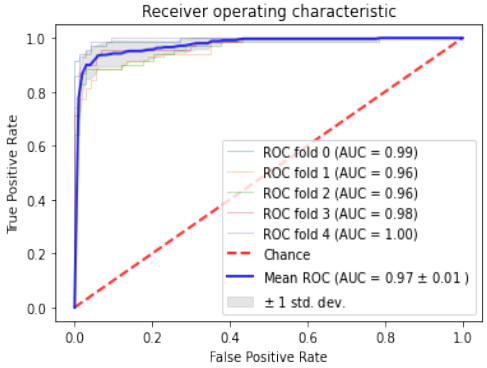


Figure 8. *K-fold* for the DDSM database with $k = 5$.

6. Acknowledgements

Authors would like to thank Instituto Politecnico Nacional (Mexico) and Consejo Nacional de Ciencia y Tecnologia (Mexico) for their support in this work.

References

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., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Aibar, L., Santalla, A., Criado, M. L., González-Pérez, I., Calderón, M., and Gallo, J.&; Parra, J. (2011). Clasificación radiológica y manejo de las lesiones mamarias. *Clínica E Investigación En Ginecología Y Obstetricia*, 38(4):141–149.

Albawi, S. (2017). Tareq abed mohammed; saad al-zawi. In *"Understanding of a convolutional neural network"*. Conference on Engineering and Technology (ICET).

Arora, R., Rai, P. K., and Raman, B. (2020). "deep feature–based automatic classification of mammograms". *International Federation for Medical and Biological Engineering*.

at Chicago (UIC), T. U. I. Uic dataset. 2020.

Backfrieder, W. and Zwettler, G. (2015). Rotated principal components for fuzzy segmentation szinitigraphic time series in individual dose planing. pages 33–37. Conference Code:116063.

Chollet, F. et al. (2015). Keras.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.

de Cancerologia (INCan) Secretaria de Salud Gobierno de México, I. N. " ¿qué es el cáncer?" actualizado.

de Oliveira et al., J. E. E. (2010). Mammosys: A content-based image retrieval system using breast density patterns,. In *Comput. Methods Programs Biomed.*

de Salud Pública, I. N. Dale la mano a la prevención del cáncer de mama.

Haindl, M. and Remeš, V. (2019). Pseudocolor enhance-

ment of mammogram texture abnormalities. *Mach. Vision Appl.*, 30:785–794.

Haindl, M.&; Remeš, V. (2019). Pseudocolor enhancement of mammogram texture abnormalities. *Machine Vision and Applications*, 30(4):785–794.

He, K., Zhang, X., Ren, S., and Sun, J. "deep residual learning for image recognition,". 2015.

INEGI. "estadísticas a propósito del día mundial de la lucha contra el cáncer de mama (19 de octubre)", comunicado de prensa núm. 462/20.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "imagenet classification with deep convolutional neural networks". Advances in Neural Information Processing Systems.

Lévy, D. and Jain, A. "breast mass classification from mammograms using deep convolutional neural networks", stanford university, [cs.

Mery, D. and Saavedra, D. &; Prasad, M. . (2020). X-ray baggage inspection with computer vision: A survey. *IEEE Access*, 8(13):45620–14563.

of Radiology (acr), A. C. "acr bi-rads atlas® 5th edition, ii. *REPORTING SYSTEM, B. ASSESSMENT CATEGORIES, pagina*, 135.

OMS/OPS. Breast cancer: prevention and control.

Pan, S. J. and Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Ragab, D. A., Sharkas, M., Marshall, S., and Ren, J. (2019a). Breast cancer detection using deep convolutional neural networks and support vector machines. 7:2167–8359.

Ragab, D. A., Sharkas, M., Marshall, S., and Ren, J. (2019b). "breast cancer detection using deep convolutional neural networks and support vector machines".

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., and Khosla, A. a. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition.

Society, A. C. "cuando se comunican con usted después del mamograma".

Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. (2017). Inception-v4 inception-resnet and the impact of residual connections on learning. AAAI Conference on Artificial Intelligence.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., and Anguelov, D. . R. (2015a). Going deeper with convolutions. pages 1–9. IEEE conference on computer vision and pattern recognition.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2015b). Rethinking the inception architecture for computer vision. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, NV, USA, 26 Jun–1 Jul 2016;. Las Vegas.

Tsochatzidis, L., Costaridou, L., and Ioannis (2019). "deep learning for breast cancer diagnosis from mammograms—a comparative study",. *Journal of Imaging*.

USA.gov, N. C. I. What is cancer?

Weiss, K., Khoshgoftaar, T. M., and Wang, D. A survey of transfer learning. *J. Big Data*, 2016:3.

Yanase, J. and Triantaphyllou, E. (2019). "a systematic survey of computer-aided diagnosis in medicine: Past and present developments",. *Expert Systems with Applications*, 138(112821).

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? *arXiv preprint arXiv:1411.1792*.

Zoph, B., Vasudevan, V., Shlens, J., and Le, Q. V. (2018). Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8697–8710.

Zwettler, G. and Backfrieder, W. (2014). Automated domain-specific feature selection for classification-based segmentation of tomographic medical image data. pages 26–35. Conference Code:108894.