DIAGNOSING BREAST MASSES USING ICA AND NON-IMAGE FEATURES

A. García-Manso, R. Gallardo-Caballero, C.J. García-Orellana,
*H.M. González-Velasco, M. Macías-Macías**

Abstract: One of the most challenging task for Computer Aided Diagnosis (CADx) systems designed to diagnose breast cancer is to be able to differentiate between benign and malignant masses. In this work we present a study made as part of an ongoing project whose aim is to develop an image-based CADx system for diagnosing mass lesions. Our system is based on *image-based* and *non-image* features. Image-based features are obtained using Independent Component Analysis (ICA), and both age and mammogram density are tested as non-image features. Performance results are provided for all the valid masses in a public database, obtaining a statistically significant improvement by adding age to image-based features. However, the addition of the density of the mammogram does not improve the system performance.

1.   Introduction

Over a million women worldwide are diagnosed with breast cancer every year, accounting for a tenth of all new cancers and 23% of all female cancer cases. Breast cancer incidence rates vary considerably, with the highest rates in North America and the lowest in Africa and Asia. Around 332,670 new cases of breast cancer occurred in the European Union in 2008 [17] and it is estimated that 182,460 occur in the USA each year [25].

The best way to fight against breast cancer is early-stage diagnose, and mammography seems to be the most effective test to diagnose breast cancer [22] since lesions can be detected before they are even felt by the patient. Early-stage detection greatly increases cure chances though, unfortunately, even expert radiologists can miss a significant proportion of abnormalities [3]. Moreover, an important

---

*Antonio García-Manso – Corresponding author, Ramón Gallardo-Caballero, Carlos J. García-Orellana, Horacio M. González-Velasco, Miguel Macías-Macías, Department of Electrical Engineering, Electronics & Automatics, Instituto de Computación Científica Avanzada de Extremadura (ICCAEx), University of Extremadura, Avda. Elvas S/N, 06006 Badajoz, Spain, E-mail: antonio@capi.unex.es

number of mammographic abnormalities are diagnosed as benign after biopsy [12] obliging patients to undergo an invasive procedure and wait for the corresponding result.

Many methods have been proposed in the last two decades to achieve a robust mammography-based CADx system. Although there are various types of mammographic abnormality, they can be primarily grouped in either masses or microcalcifications [19]. In the literature we can find different proposals focused on detection and segmentation of masses in mammograms, as found in [21]. Nevertheless, studies described in papers addressing both detection and diagnosis of masses usually use a proprietary database of small size or, if using a public database, the cases selected are not specified. This makes it difficult to compare results between different techniques. Rangayyan [23] provides a comparison of selected computer aided methods indicating the size of the used dataset and its accuracy results. The number of mammograms used in the listed works is as low as 25 in the work of Brzakovic [6], being the highest 322 in the work of Zheng and Chan [33] that uses all the mammograms provided by the Mini-MIAS database. Horsch [14] analyzes recent studies in mammography CADx and concludes that, in view of the observed variability in the used data sets, currently the only mammography database that is public and sufficiently large to allow for a meaningful and reproducible evaluation of a CADx system is the Digital Database for Screening Mammography [13] (DDSM).

Mass diagnosis is performed using very different techniques and datasets, most of which are considered to be non reproducible [14]. In CADx system, the performance is usually expressed by the Area Under the ROC Curve [10] (AUC or A_z). Restricting ourselves to works which provide a reproducible selection of mammograms, the number of papers that address this problem is really reduced [14]. Some studies use human provided features, as Elter [8], where the BI-RADS® (Breast Imaging-Report and Database System) [1] description of abnormalities is used to predict malignancy considering three types of classifiers. This technique was applied to 2,100 masses from DDSM, reporting an AUC value of 0.89. Similarly, Kim [18] also uses BI-RADS® features from DDSM but adds 14 gray level features extracted from masses. This study analyzes the dependence of the system's performance on the features selected and the origin of the case, i.e., the institution where the case was collected from. The authors report AUC values from 0.76 to 0.94. Finally, Yoon [30] adds the assessment of the cancer as a parameter to be considered in the diagnosis of a mass. The AUC values reported by this author are between 0.92 and 0.97.

More recently, some other papers describing classification of masses in mammograms (benign and malignant) have been published [32, 28, 29]. Though these papers were not analyzed in [14], in our view only [32] would have been listed as reproducible. In that study, they used mass shape features and an ensemble system formed by four specialist classifiers based on patient age (young/old) and region of interest (ROI) size (large/small). To train and test the system, they used the DDSM database, selecting the mammograms from Lumisys scanner. From these selected mammograms, they removed instances with extreme digitization artifacts and extremely large size (over 2000×2000 pixels). They also removed instances with mixed BI-RADS descriptors, and those ROI images that displayed only a por-

tion of a mass. Finally, they considered 543 mass prototypes (272 benign and 271 malignant) and obtained a performance of 56% with a single classifier trained with all mass prototypes, and a performance of 72% with the ensemble system of specialist classifiers. On the other hand, Tahmasbi [28] proposed an approach based on Zernike moments as descriptors of shape and margin characteristics, and a Multilayer Perceptron (MLP) trained with Back-Propagation (BP) or Opposition-based Learning (OBL) algorithms. They used the MIAS database (67 ROIs with benign lesions and 54 ROIs with malignant lesions) to train and test their system. As it was stated in [14] the MIAS database is valid for a meaningful and reproducible evaluation of a CADx system. Finally, Verma [29] presents an approach which relies on a *soft clustered based direct learning classifier* (SCBDL). They selected 200 mammograms (100 malignant and 100 benign) from the USF's Digital Database for Screening Mammography (DDSM) database, but they did not specify how they made this selection, and therefore a meaningful and reproducible evaluation cannot be made. In our study, we use a feature extractor based on Independent Component Analysis (ICA) to diagnose the masses provided in the mammograms of the DDSM. And as a secondary objective, we analyze the effect on the performance of our system obtained by the addition of non-image parameters such as age and breast density of the mammogram. The previously cited works used mass shape and margin as a base for the diagnosis being both features provided by human DDSM radiologists. As we mainly considered computer extracted features, a direct comparison of our results with those presented in the previously cited works does not provide an accurate picture of the performance of the different proposals. On the other hand, an approach similar to ours was presented in [7], where they also used an ICA-based feature extractor to obtain a number of texture features from regions of interest, which were later classified as abnormal o normal tissue using a supervised probabilistic classifier. However, they used the MIAS database for the mammographic images, and therefore we cannot compare the results, considering that they should be obtained over the same set of data to be comparable, as stated in [20] and [14].

The rest of our paper is organized as follows. Section 2 introduces the general concepts related to feature extraction and dataset characteristics used in our experiments. Next, Section 3 describes our methodology in detail. The results obtained by applying the system to a concrete subset of the DDSM database can be found in Section 4, and finally, our conclusions and possible improvements of our work are presented in Section 5.

2. Materials and methods

In this section, we provide a brief description of the mammogram database utilized. Furthermore, we describe the procedure implemented to build a set of mass prototypes and the main characteristics of the image feature extractor. Finally, in the last subsection we outline our distributed implementation of the neural classifier training process.

2.1 Data and prototype creation

The DDSM [13] is a resource available to the mammographic image analysis research community. Contains a total of 2,620 cases, each of which provides four screening views, mediolateral oblique (MLO) and craniocaudal (CC) projections of left and right breasts. Therefore, the database has a total of 10,480 images.

The cases are categorized in four major groups: *normal, cancer, benign and benign without callback*. In addition to the regular DDSM volumes, DDSM website provides the so called "DoD BCRP Mammography Datasets at USF" composed of four sets. They have been created for preliminary evaluation of the performance of CADx algorithms. Two of the four datasets focus on spiculated masses (BCRP_MASS_0 and BCRP_MASS_1), while the other two focus on clustered microcalcifications (BCRP_CALC_0 and BCRP_CALC_1). Each case in the BCRP datasets contains at least one malignant lesion of the corresponding type.

All cases in the DDSM database were reported by experienced radiologists providing various BI-RADS parameters (density, assessment and subtlety), BI-RADS abnormality description and proven pathology. For each abnormality identified, the radiologist draws free form digital curves defining ground truth regions. Besides, each DDSM case includes additional information such as patient age, date of study and digitization or digitizer's brand.
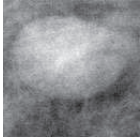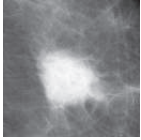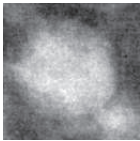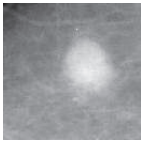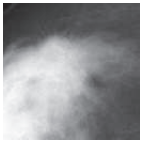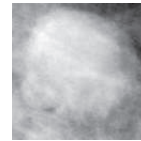
In our method, we use the regions provided by the radiologists to define squared ROIs, which are considered as prototypes for the mass classes. The DDSM database contains 2,582 regions reported as masses, both benign and malignant. From these regions, those located on the border of the mammograms were discarded, because it was not possible to take a square, with the mass centered, without stretching. Therefore, only 2,324 regions could be considered. Some mass prototype examples are shown in Tab. I.

2.1.1 Regions of interest
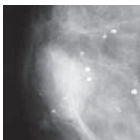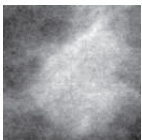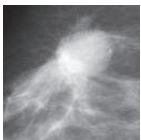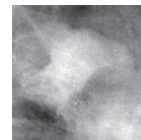
Ground truth regions are defined in the DDSM database by a chain code which generates a free hand closed curve. As stated above, we use the chain code to determine the smallest square region of the mammogram that includes the manually defined area. Therefore, if the mass is located near one edge of the mammogram, this procedure may not be able to obtain a squared region from the image and the mass is discarded as a valid prototype. Fig. 1 shows the ground truth region marked by the radiologist (solid line) and the area to be used as ROI (dot-dashed box).

DDSM mammograms were digitized with four different scanners for which optical density calibration [13] and spatial resolution is known. In addition, three scanners provide a linear optical response, and the fourth provides a logarithmic one. To eliminate the dependence of the origin of each digitized mammogram, all the ROIs considered were converted to optical density using the referenced calibration parameters.

The regions generated in this way have different sizes. However, the selected image feature extractor needs to operate on regions with the same size. Therefore, we need to reduce the size of the selected regions to common sizes. The reduction of ROIs to a common size has demonstrated to preserve mass malignancy information

Shape	Edges			
	Circumscribed	Ill-defined	Spiculate	Obscured
Oval	Case3090-RCC	Case0107-LCC	Case4178-LCC	Case0418-LCC
Round	Case3391-RCC	Case4021-RCC	Case0339-RCC	Case1576-RCC
Lobulate	Case0145-RCC	Case1908-RCC	Case0457-RMLO	Case0418-LMLO

Tab. I *Mass samples for each shape and margin combination. Each ROI has been resized to a common size of* 128×128 *pixels. Mammogram case and view is located over each ROI.*

.

[2]. To determine the optimum region size, we resized each ROI to three sizes: 32×32, 64×64 and 128×128 pixels. Resizing has been carried out using the bilinear interpolation algorithm provided by the OpenCV [5] library.

2.2 Feature vectors

As described above, we analyzed the performance variations obtained with the use of mixed feature vectors, combining image-extracted and non-image features available for each region of interest. The feature vector is mainly composed by image-extracted features using ICA. The non-image features that we considered for this study were age and mammogram density, obtained from each case specification. Both parameters were normalized when included as additional features in the feature vector of each ROI using the following expression:

$$x_i^{\mathrm{nor}} = \frac{x_i - \overline{x}}{\sigma},\tag{1}$$

where $\overline{x}$ and σ are the *mean* and the *sample standard deviation* calculated over the set of prototypes.
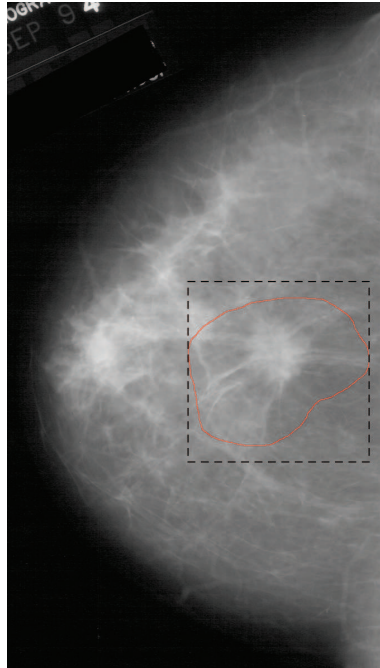
Fig. 1 *Ground truth region defined by a radiologist (solid line) and the ROI considered (dot-dashed box) on a DDSM mammogram.*

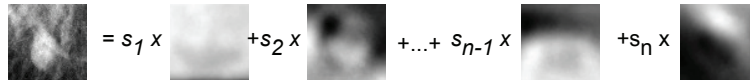## 2.3 Independent Components Analysis

Independent Component Analysis [15] is a statistical generative model whose objective is to explain the original data ($\mathbf{X}$) using statistically independent random vectors ($\mathbf{S}$), as it is indicated in Eq. (2). ICA defines a generative model for the observed multivariate data, typically given as a large sample database. In this model, it is assumed that the data are linear combinations of unknown latent variables, and the system whereby combined is also unknown. It is also assumed that non-gaussian latent variables are mutually independent and thus considered independent components of the observed data. These independent components, also called sources or factors, can be found by ICA. In our application, we can decompose or expand our image ($I(x,y)$) by using a base of images ($a_i(x,y)$, with $i = 1, \ldots, p$) multiplied by coefficients (s_i):

$$\mathbf{I}(x,y) = \sum_{i=1}^{p} \mathbf{A}_i(x,y)s_i. \tag{2}$$

The goal of ICA is to estimate the mixing matrix $\mathbf{A}$, in addition to the sources s_j. This technique can be used for feature extraction since the components of $\mathbf{X}$ can be considered as features representing the objects [15].

2.3.1   ICA feature extraction

Many basic models in image processing express an image as a linear superposition of some features or basis functions, where the coefficients are different for each image. Some linear transformations widely used in image processing are the Fourier, Haar or Gabor transforms [11]. Following this model, we can represent a ROI in a mammographic image as a series expansion in these functions. The coefficients of this series expansion can then be used as features characterizing the ROI to compose a feature vector which may be used to diagnose it. We can visualize this idea in Fig. 2.

Fig. 2 *A ROI expressed as a linear superposition of some ICA extracted basis functions.*

To estimate ICA basis from images we need to collect samples (patches) from the images to model. The collected patches are used to build a data matrix $\mathbf{X}$ which is the input to the FastICA [16] algorithm. In this algorithm, input data are centered first by subtracting the mean of each column of the data matrix $\mathbf{X}$. The data matrix is then *whitened* by projecting the data onto its principal component directions using a pre-whitening matrix ($\mathbf{K}$). The number of components (features) to extract is selected a priori. The ICA algorithm then estimates an un-mixing matrix $\mathbf{W}$ so that $\mathbf{XKW} = \mathbf{S}$, being $\mathbf{S}$ the estimated sources matrix.

Using the previous notation, being $\vec{R}$ the image associated to a previously-centered ROI, we can obtain the feature vector ($\vec{F}$) that characterizes the ROI as $\vec{F} = \mathbf{KW}\vec{R}$.

2.4   The classifier

We used an artificial neural network (ANN) to classify the obtained feature vectors. In particular, we consider a classical feed-fordward multilayer perceptron with a single hidden layer, which was trained using a variant of Back-Propagation (BP) algorithm named Resilient Back-Propagation (RPROP) [24]. Regarding the implementation, the Stuttgart Neural Network Simulator environment SDK [31] was used to generate and train the neural network classifiers. To avoid local minimum during the training process, each setting was repeated four times, changing the initial weights in the net at random. Furthermore, the number of neurons in the hidden layer was allowed to vary between 50 and 650 in steps of 50, selecting the network that provides the highest success rate over the test subset.

3.   Outline of the process

In this section, we provide an overview of the structure of our system, describing the main steps required to configure it and to obtain a diagnose for an unknown region of interest.

3.1 System setup

The implementation of the system consists of three stages. In the first stage we estimate ICA basis from ROI images. The second stage involves all the procedures required to optimize the neural classifiers for each considered feature vector composition. And finally, in the last stage we evaluate the diagnostic capability for each input setting to select the best system configuration.

We estimated ICA basis using mass prototypes (both benign and malignant) as input to the *FastICA* algorithm applying *log cosh* function to estimate neg-entropy. For of the selected prototype sizes (32×32, 64×64 and 128×128), we generated a set of ICA basis. Each ICA basis allows us to estimate a specific number of features from a ROI of a given size. In order to optimize the performance of the system, we obtained ICA basis to calculate from 10 to 60 features in steps of five, totalizing 33 different feature extraction configurations. Fig. 3 shows an ICA basis obtained from sample patches which provides 30 features.

Masses Prototypes   ICA Basis of 30 components
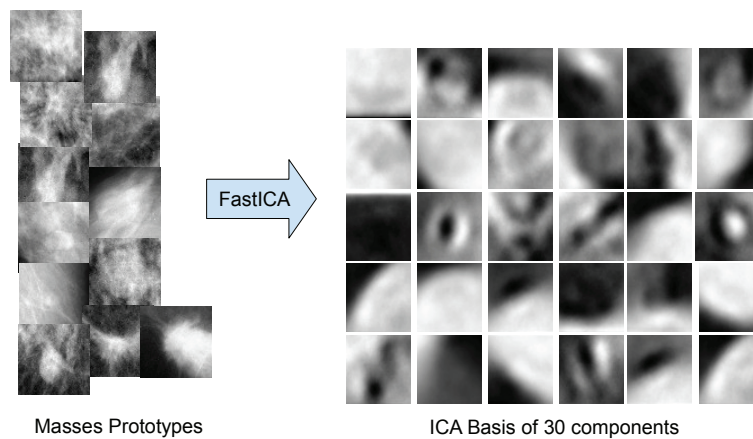
Fig. 3 *Obtention of ICA basis vectors of mammographic tissue from a set of prototypes.*

Any combination of *ICA window size* and *number of features to be extracted* generates an input configuration for the neural classifier. To train it, we built a data set of ICA-extracted feature vectors for each input configuration. Each data set was split into three disjoint subsets: learning, validation and test. We decided to use a validation subset to estimate the generalization error and improve generalization by avoiding the overtraining [4].

Once the subsets intended for learning, validation and testing were generated, the neural network classifier (a Multilayer Perceptron) was trained with the RPROP learning algorithm [24]. The Stuttgart Neural Network Simulator (SNNS) [31] environment was used to generate and train the network. The high number of input configurations to be optimized led us to use SNNS's kernel function facilities. These libraries allowed us to implement network training procedures in standalone executable that could be run in a distributed Beowulf cluster [27].

To achieve an optimal network configuration, we made a sweep of the number of neurons in the hidden layer. We trained several network configurations with a number of hidden neurons ranging from 50 to 650 in steps of 50. Furthermore, each configuration was repeated four times, performing a random initialization of the neuron weights in each repetition.

Finally, the test subset, which contains input vectors not used in the optimization of the neural classifier, was used to provide performance results of our system.

3.2   Experiments

Two experiments were configured using the same database of prototypes, but changing the procedure to build the sets of prototypes and the method to evaluate the performance of the system. This may let us study the robustness of our system to variations on the composition of the sets. The setup for this two experiments was as follows:

- In *experiment A* ten different training cycles were carried out building the corresponding prototype sets as follows. First, all the prototypes extracted from the BCRP_MASS_1 subset and fifty benign prototypes (randomly selected in each cycle) were included in the test subset. Then, the remaining prototypes were used to build training and validation subsets, assigning 80% of them to the training subset and the other 20% to the validation subset.

- In *experiment B* a 10-fold cross-validation test [4] was carried out using all the collected prototypes. In each fold, the prototypes selected to adjust the model were divided into two parts, using 80% of the prototypes to train the neural network and the remaining to avoid overtraining.

Tab. II shows the average number of benign and malignant masses on each subset (training, validation and test) for both experiments. In experiment A, since prototypes were included in training or validation subsets randomly and independently of its pathology, we provide mean values for the ten cycles. As can be seen, the selection process led to a balanced distribution of malignant and benign prototypes in each subset. In average, an 80.7% of the malignant prototypes were included in the training subset and the remaining 19.3% into the validation subset. In the same way, a 79.3% of the benign prototypes where included in the training subset and a 20.7% into the validation subset. Furthermore, the 10-fold process carried out in experiment B generates a maximum imbalance among malignant and benign prototypes number lower than two percent for the three subsets.

Fig. 4 shows the performance values obtained with classifiers that used exclusively ICA feature extractors with a ROI size of 64×64, over their corresponding test set on experiment B. The best median success rate value was obtained for an ICA feature extractor which operates over 64×64 pixel prototypes and provides 20 features from the ROI. Therefore, as the best performances was obtained on prototypes of 64×64 pixel, making further tests was discarded, both with prototypes smaller than 32×32 pixels and larger than 128×128 pixels.

To evaluate the improvement obtained adding non-image features, we decided to compare only those input configurations that include the selected ICA feature extractor and other non-image features.

Subset	Pathology	Experiment A	Experiment B
Training	Malignant	900.3	823.3
	Benign	857.7	802.7
Validation	Malignant	215.7	251.3
	Benign	223.3	217.7
Test	Malignant	76	121.2
	Benign	51	115.8

**Tab. II** *Average number of malignant and benign masses in each subset of the two experiments.*
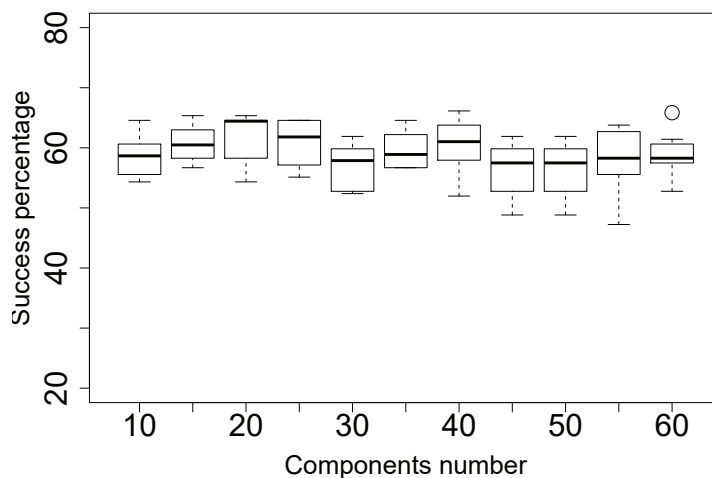
Fig. 4 *Success rate over the test set of classifiers that used exclusively 64×64 ICA feature extractors and the considered features number.*

4. Results and discussion

Neural classifiers' performance was evaluated by means of Receiver Operating Characteristics (ROC) graphs [9] and accuracy tests. Specifically, we choose the Area Under the ROC Curve (AUC) [10] to compare the performance of the considered classifiers. ROC curves were generated using the ROCR package of the R language environment [26].

Fig. 5 provides ROC curves obtained for the four feature combinations considered in experiment A. Graph (a) in Fig. 5 corresponds to a system whose input feature vector only considers the 20 features provided by the common ICA feature extractor. Graph (b) shows the response of the system when adding the normalized age to the previous feature vector. Graph (c) shows the performance obtained combining ICA features and normalized mammogram density. And finally, graph (d) provides the response of the system that considers ICA features, normalized age and density to generate its input feature vector.
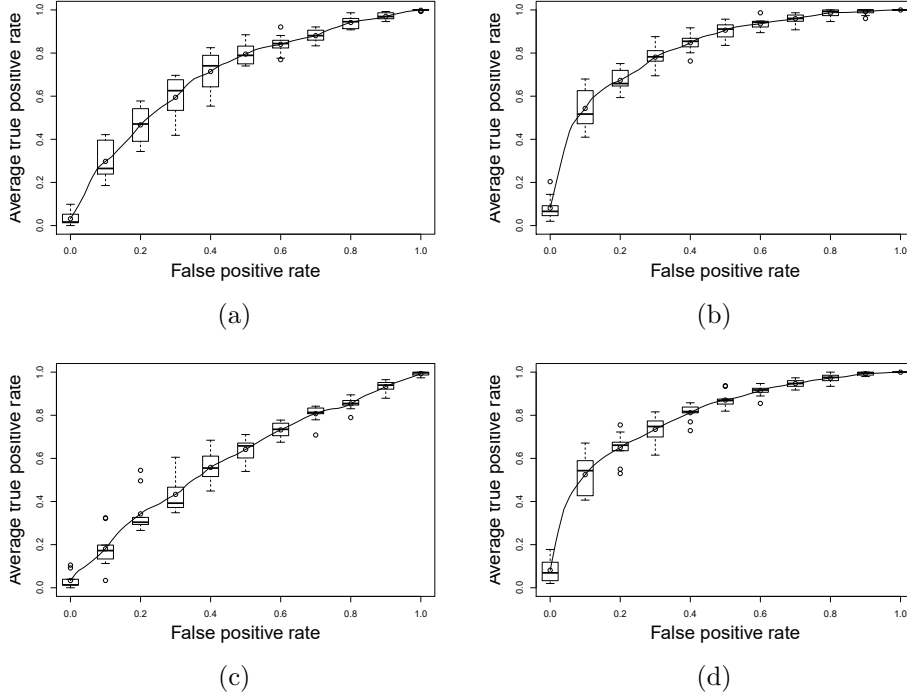
Fig. 5 *ROC curves for experiment A: (a) Only ICA features; (b) ICA + age; (c) ICA + density; (d) ICA + age + density. False positive rate is plotted in x axis and average true positive rate is plotted in y axis.*

Fig. 6 provides ROC curves obtained for experiment B with the same arrangement shown in graphs (a-d). Tab. III provides a summary of the AUC numeric values obtained both on experiment A and experiment B, and its corresponding confidence intervals for a significance level of 95%.

Direct observation of the ROC graphs for the same input configuration in experiments A and B allows to visually assess that there are differences in performance. Nishikawa [20] associated this behavior to the composition of the data sets. Furthermore, we can also observe a performance dependence on the normalized age (see Figs. 5 or 6), being this dependence statistically meaningful in both experiments for a significance level of 95% (see Tab. III). The inclusion of the density parameter worsens the overall system performance in both experiments, though this difference is significative only in experiment A. This result suggests that the use of this parameter, which characterizes the global breast tissue, can provide information not relevant to the scale ROI. On the other hand, comparing the configurations *ICA+age* and *ICA+age+density* in our experiment, we observe a slightly different behavior regarding the usefulness of the density parameter in the diagnosis of ROIs. Experiment B achieves higher AUC by including the density parameter in the input configuration *ICA+age*, though this improvement was not statistically significant.
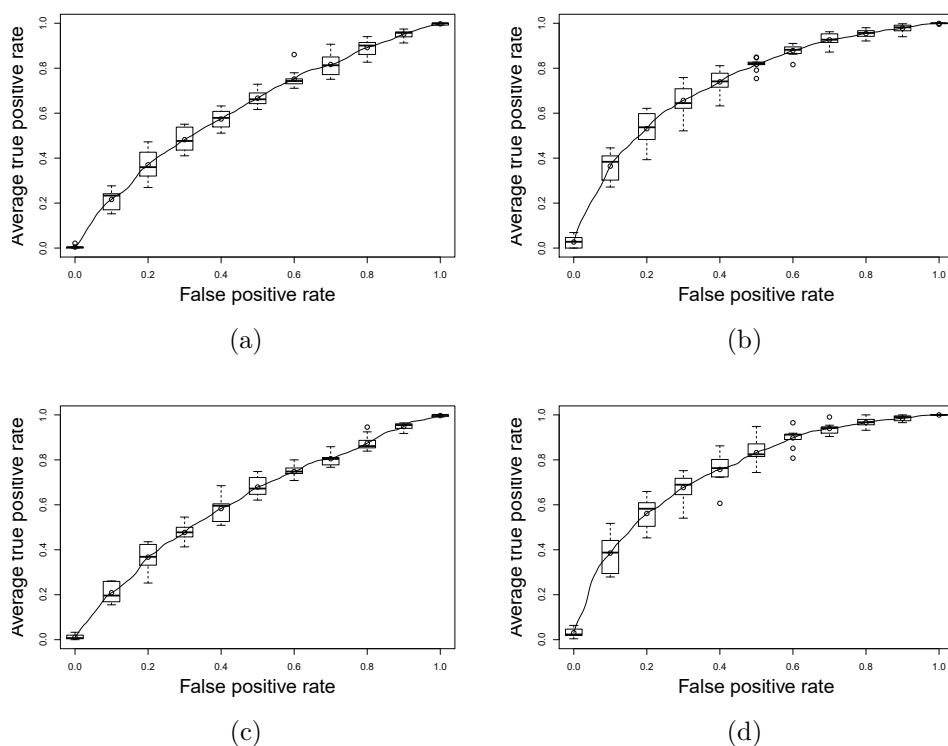
Fig. 6 *ROC curves for Experiment B: (a) Only ICA features; (b) ICA + age; (c) ICA + density; (d) ICA + age + density. False positive rate is plotted in x axis and average true positive rate is plotted in y axis.*

The accuracy $(P(\widehat{Y} = Y))$ is estimated as $\frac{TP+TN}{P+N}$ and gives the likelihood of a correct diagnosis with the implemented system. The accuracy analysis performed on our system provides the results presented in Tab. III. The values shown indicate the average probability for each configuration analyzed. This study shows lower performance in experiment B, presenting an overall behavior similar to previous ROC analysis.

As we stated in the introduction, considering [32, 28, 29], we can only reliably compare our results with those presented in [32]. Here, the authors selected the cases provided only by one scanner, the Lumysis, for the purpose of data consistency. They obtained an accuracy of 56% with a single classifier trained with all mass prototypes, and an accuracy of 72% with the ensemble system of specialist classifiers. As can be seen in Tab. III, we have obtained the best performance for the experiment A, with an AUC about 0.8 and an accuracy about 77%, while for the experiment B was obtained an AUC about 0.7 and an accuracy around 69%. As can be seen, our results (for both experiments) are better than those presented in [32] where a single classifier is used. Besides, if we compare with their results for an ensemble of specialist classifiers, our results are better for experiment A, and

40

Configuration	Experiment A		Experiment B	
	AUC	95% CI	AUC	95% CI
ICA	0.706	[0.679, 0.732]	0.624	[0.605, 0.642]
ICA+age	0.825	[0.805, 0.846]	0.739	[0.719, 0.759]
ICA+density	0.602	[0.576, 0.628]	0.620	[0.599, 0.641]
ICA+age+den	0.805	[0.788, 0.822]	0.755	[0.731, 0.780]
	ACC[%]	95% CI	ACC[%]	CI 95% CI
ICA	70.5	[68.5, 72.6]	61.3	[59.5, 63.0]
ICA+age	77.6	[76.1, 79.0]	69.6	[67.4, 71.7]
ICA+density	63.3	[62.4, 64.2]	61.0	[59.1, 63.0]
ICA+age+den	75.9	[74.4, 77.4]	70.5	[68.5, 72.6]

Tab. III *ROC area results and Accuracy results. Summary results obtained for both experiments over the test subset. AUC values provide the average area under ROC curve, and ACC represents the average true positive fraction, both with the given confidence interval (CI) for a significance level of 95%.*

only a little worse for experiment B. Moreover, it is important to remark that our method is more robust since we consider all cases of the DDSM database, regardless of the source (Lumisys 200 Laser, Howtek, DBA M2100 ImageClear, Howtek MultiRAD 850 [13]).

5. Conclusions

This work presents a method for the diagnosis of masses provided in mammograms of the DDSM. The masses are classified into malignant and benign, considering image-extracted features complemented by non-image parameters: age of the patient and breast density of the mammogram. The method has been tested with a concrete subset of the DDSM, using two different procedures to select the prototypes in the training stage. The results obtained in the experiments suggest that ICA can be considered as a serious technique for feature extraction when designing a CADx system for the diagnosis of masses. Though comparisons can not be made, as we discussed in Section 1, these results are not far from those of other authors, and can be improved when features extracted by ICA are combined with features not directly obtained from images, such as the age. In fact, the system performance increases significantly when normalized age is included, though this increase is not observed with breast density. Therefore, more research effort should be conducted to determine which non-image features are suitable for the task. On the other hand, we have observed a significant variation linked to the prototype selection criteria chosen, what can be considered as a drawback of the method. In this case, it would be interesting to study more deeply this dependence, in order to determine the procedure that better retains the prediction capacity of the classifier. Finally, to reduce the number of false positives per image and improve the rates

of correct diagnosis, we intend to use techniques for feature selection as genetic algorithms (GA). In summary, a lot of work should be done to have a completely reliable diagnosis system, but these results can be considered quite promising.

Acknowledgments

References

[1] AMERICAN COLLEGE OF RADIOLOGY. *ACR BI-RADS$^{\circledR}$ – Mammography*. Reston, VA: American College of Radiology, 2003.

[2] ANGELINI E., CAMPANINI R., IAMPIERI E., LANCONELLI N., MASOTTI M., ROFFILLI M. Testing the performances of different image representations for mass classification in digital mammograms. *International Journal of Modern Physics C [Computational Physics and Physical Computation]*. 2006, 17(1), pp. 113–131, doi: `10.1142/S0129183106009199`.

[3] BIRDWELL R., IKEDA D., O'SHAUGHNESSY K., SICKLES E. Mammographic Characteristics of 115 Missed Cancers Later Detected with Screening Mammography and the Potential Utility of Computer-aided Detection. *Radiology*. 2001, 219(1), pp. 192–202, doi: `10.1148/radiology.219.1.r01ap16192`.

[4] BISHOP C.M. *Pattern Recognition and Machine Learning*. Springer, 2006.

[5] BRADSKI G., KAEHLER A. *Learning OpenCV: Computer Vision with the OpenCV Library*. 1st ed. O'Reilly Media, 2008.

[6] BRZAKOVIC D., LUO X., BRZAKOVIC P. An approach to automated detection of tumors in mammograms. *IEEE Transactions on Medical Imaging*. 1990, 9, pp. 233–241, doi: `10.1109/42.57760`.

[7] CHRISTOYIANNI I., KOUTRAS A., DERMATAS E., KOKKINAKIS G. Computer aided diagnosis of breast cancer in digitized mammograms. *Computerized medical imaging and graphics*. 2002, 26(5), pp. 309–319, doi: `10.1016/S0895-6111(02)00031-9`.

[8] ELTER M., WITTENBERG T. The prediction of breast cancer biopsy outcomes using two CAD approaches that both emphasize an intelligible decision process. *Medical Physics*. 2007, 34(11), doi: `10.1118/1.2786864`.

[9] FAWCETT T. An introduction to ROC analysis. *Pattern Recognition Letters*. 2006, 27, pp. 861–874, doi: `10.1016/j.patrec.2005.10.010`.

[10] FOGARTY J., BAKER R.S., HUDSON S.E. Case studies in the use of ROC curve analysis for sensor-based estimates in human computer interaction. In: *Proceedings of Graphics Interface 2005*, Victoria, British Columbia: Canadian Human-Computer Communications Society, 2005, pp. 129–136. Available also from: `http://dl.acm.org/citation.cfm?id=1089508.1089530`,

[11] GONZALEZ R.C., WOODS R.E. *Digital image processing*. 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 2008.

REFERENCES

[12] HALL F., STORELLA J., SILVERSTONE D., WYSHAK G. Nonpalpable breast lesions: recommendations for biopsy based on suspicion of carcinoma at mammography. *Radiology.* 1988, 167(2), pp. 353–358, doi: `10.1148/radiology.167.2.3282256`.

[13] HEATH M., BOWYER K., KOPANS D., MOORE R., KEGELMEYER P. The digital database for screening mammography. In: MJ YAFFE, ed. *Proceedings of the 5th International Workshop on Digital Mammography*, 2001, pp. 212–218,

[14] HORSCH A., HAPFELMEIER A., ELTER M. Needs assessment for next generation computer-aided mammography reference image databases and evaluation studies. *International Journal of Computer Assisted Radiology and Surgery.* 2011, 6(6), pp. 749–767, doi: `10.1007/s11548-011-0553-9`.

[15] HYVÄRINEN A., OJA E. Independent components analysis: Algorithms and applications. *Neural Networks,* 2000, 13, pp. 411–430, doi: `10.1016/S0893-6080(00) 00026-5`.

[16] HYVÄRINEN A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks.* 1999, 10(3), pp. 626–634, doi: `10.1109/72.761722`.

[17] JEMAL A., BRAY F., CENTER M., FERLAY J., WARD E., FORMAN D. Global cancer statistics. *CA: A Cancer Journal for Clinicians.* 2011, 61(2), pp. 69–90, doi: `10.3322/caac.20107`.

[18] KIM S., YOON S. Mass Lesions Classification in Digital Mammography using Optimal Subset of BI-RADS and Gray Level Features. In: *6th International Special Topic Conference on ITAB*, Tokyo: IEEE, 2007, doi: `10.1109/ITAB.2007.4407354`.

[19] KOPANS D. *Breast imaging.* Baltimore, MD : Lippincott Williams & Wilkins, 2007.

[20] NISHIKAWA R.M. Current status and future directions of computer-aided diagnosis in mammography. *Computerized Medical Imaging and Graphics.* 2007, 31, pp. 224–235, doi: `10.1016/j.compmedimag.2007.02.009`.

[21] OLIVER A., FREIXENET J., MARTÍ J., PÉREZ E., PONT J., DENTON E.R., ZWIGGELAAR R. A review of automatic mass detection and segmentation in mammographic images. *Medical Image Analysis.* 2010, 14, pp. 87–110, doi: `10. 1016/j.media.2009.12.005`.

[22] PEART O. *Mammography and Breast Imaging. Just the facts.* 1st ed. McGraw-Hill, 2005.

[23] RANGAYYAN R.M., AYRES F.J., DESAUTELS J.L. A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. *Journal of the Franklin Institute.* 2007, 344, pp. 312–348, doi: `10.1016/j.jfranklin.2006.09.003`.

[24] RIEDMILLER H., BRAUN H. A direct adaptive method for faster backpropagation learning. The RPROP algorithm. *IEEE International Conference on Neural Networks.* 1993, 1, pp. 586–591, doi: `10.1109/ICNN.1993.298623`.

[25] SIEGEL R., WARD E., BRAWLEY O., JEMAL A. Cancer statistics, 2011. The impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA: A Cancer Journal for Clinicians.* 2011, 61(4), pp. 212–236, doi: `10. 3322/caac.20121`.

[26] SING T., SANDER O., BEERENWINKEL N., LENGAUER T. ROCR: Visualizing classifier performance in R. *Bioinformatics.* 2005, 21(20), pp. 3940–3941, doi: `10. 1093/bioinformatics/bti623`.

[27] Beowulf: A Parallel Workstation For Scientific Computation. 24th International Conference on Parallel Processing. 1995.

REFERENCES

[28] TAHMASBI A., SAKI F., SHOKOUHI S.B. Classification of benign and malignant masses based on Zernike moments. *Comp. in Bio. and Med.* 2011, 41(8), pp. 726–735, doi: `10.1016/j.compbiomed.2011.06.009`.

[29] VERMA B., MCLEOD P., KLEVANSKY A. Classification of benign and malignant patterns in digital mammograms for the diagnosis of breast cancer. *Expert Syst. Appl.* 2010, 37(4), pp. 3344–3351, doi: `10.1016/j.eswa.2009.10.016`.

[30] YOON S., KIM S. AdaBoost-based multiple SVM-RFE for classification of mammograms in DDSM. *BMC Medical Informatics and Decision Making.* 2009, 9(Suppl 1)(S1), pp. 10, doi: `10.1186/1472-6947-9-S1-S1`.

[31] ZELL A., MACHE N., HUEBNER R., MAMIER G., VOGT M., SCHMALZL M., HERRMANN K. SNNS (Stuttgart Neural Network Simulator). *Neural Network Simulation Environments.* 1994, 254, pp. 165–186, doi: `10.1007/978-1-4615-2736-7_9`.

[32] ZHANG Y., TOMURO N., FURST J., RAICU D.S. Using BI-RADS Descriptors and Ensemble Learning for Classifying Masses in Mammograms. *Lecture Notes in Computer Science.* 2010, 5853, pp. 69–76, doi: `10.1007/978-3-642-11769-5_7`.

[33] ZHENG L., CHAN A. An artificial intelligent algorithm for tumor detection in screening mammogram. *IEEE Transactions on Medical Imaging.* 2001, 20(17), pp. 559–567, doi: `10.1109/42.932741`.