SVM CLASSIFICATION OF BREAST TUMORS ON ULTRASOUND IMAGES USING MORPHOLOGICAL FEATURES

I. Domingues*, A. V. Alvarenga**, J. S. Cardoso* and W. C. A. Pereira***

*INESC TEC (formerly INESC Porto)/Faculdade de Engenharia, Universidade do Porto, Portugal
** Laboratory of Ultrasound, National Institute of Metrology, Standardization, and Industrial Quality (Inmetro), Rio de Janeiro, RJ, Brasil
***Programa de Engenharia Biomédica - COPPE/UFRJ, Rio de Janeiro, RJ, Brasil

e-mail: inesdomingues@gmail.com

Abstract: This work aims at investigating morphological features in distinguishing malignant and benign breast tumours on ultrasound images. Support Vector Machines were applied as the classification methodology. Performances were assessed with accuracy, sensitivity and specificity.

As previously seen, the most relevant individual feature is the normalized residual value, calculated from the convex polygon technique. When combined, normalized residual value, morphological-closing ratio and overlap ratio achieved an accuracy of 87% in distinguishing malignant and benign breast.

The methodology developed here can be used to analyse other databases.

***Keywords*: Breast cancer, ultrasound, classification, Support Vector Machines**

Introduction

Approximately 230,000 new cases of invasive breast cancer and 40,000 breast cancer deaths were expected to occur among US women in 2011 [1]. According to the Brazilian National Cancer Institute (INCa), breast cancer is one of the main causes of women death in Brazil [2].

At present, there is no known method to prevent breast cancer but early detection and diagnosis increase the chance of cure [3]. Therefore, screening is recommended by all medical community [4], [5]. Mammography and ultrasound (US) are the main imaging techniques for breast cancer. While mammography is the adopted screening method, US is important to diagnosis complement and has been proved to diminish both false positives and false negatives from screening mammography. Annual ultrasound screening may detect small, node-negative breast cancers that are not seen on mammography [6].

Up to date, to confirm whether a patient has breast cancer, biopsy has to be performed. However, biopsy is one kind of surgical procedure that can cause psychological and physical consequences on patients. To avoid unnecessary biopsies, researchers have investigated computer-aided diagnosis (CAD) systems that offer more objective evidences and stable high diagnostic rates [7].

In this study, we investigate the potentiality of seven previously proposed morphological features in assessing breast tumour on US images. Support vector machines are applied to evaluate the individual and combined capability of features in distinguishing breast tumours between malignant and benign. The features performance is assessed using accuracy, sensitivity and specificity as figures of merit. Moreover, based on their performances, we also endeavour to infer which features formulation characteristics influence their performances, and which features are capable of quantifying tumours contours.

State of the art

Seven morphological features were investigated in [8] to distinguish malignant and benign breast tumours on US images. Linear discriminant analysis was applied to sets of up to five features. The most relevant individual features were the normalized residual value (nrv) and overlap ratio (RS), both calculated from the convex polygon technique, and the circularity (C).When nrv and C were taken together with roughness (R), calculated from normalized radial length (NRL), a performance slightly over 83% in distinguishing malignant and benign breast tumours was achieved.

The current work is mainly based on the above described work. So, this literature review will only focus papers published in 2010 or after. For an earlier literature survey, consult [8].

Wan *et al*. [9] formulated the problem of choosing discriminative features as a decomposition of the computerized feature matrix into a low-rank principal matrix and a sparse error matrix. By identifying and selecting essential features, the low-rank matrix based feature selection method can improve the classification outcomes. Training datasets comprise 92 benign cases and 172 malignant cases, and the test datasets have 21 benign cases and 36 malignant cases.

Liao's group [10] established a new set of features for differentiating benign from malignant breast lesions. Sonograms of 321 pathologically proven breast cases are analysed. The discrimination capability of the extracted features are evaluated using the support vector machines (SVM) in comparison with the results obtained from artificial neural networks (ANN) and K-nearest neighbour (KNN) classifier.

Gómez *et at*. [11] proposed a CAD system for breast US. A differential evolution technique was used to optimize the structure of a radial basis function (RBF) neural network. The dataset consisted of 641 breast US images: 228 carcinomas and 413 benign masses.

Shi and his team [12] also developed a CAD system. This time, the classifier used was fuzzy SVM. Experimental results were achieved with a dataset of 87 cases (36 malignant solid masses and 51 benign ones).

Literature results are summarized in Table 1.

Table 1: Comparison of some breast US classification systems.

Paper	Az	Accuracy (%)	Sensitivity (%)	Specificity (%)
[8]	0.86	84	83	85
[9]	Not specified	86	78	91
[10]	Not specified	87	75	96
[11]	0.97	90	90	91
[12]	0.96	94	92	96

Materials and Methods

In this retrospective study, 246 breast tumour US images were acquired from at the Cancer National Institute (Brazil, Rio de Janeiro) using a 7.5MHz linear array B-mode (Sonoline—Sienna® Siemens) equipped with a 40mm US probe (axial and lateral beam resolutions of 0.45mm and 0.49mm, respectively). The present study was carried out considering INCa's diagnosis routine. Histopathological diagnosis concluded that 177 tumours were malignant and 69 benign. Database examples are shown in Figure 1.

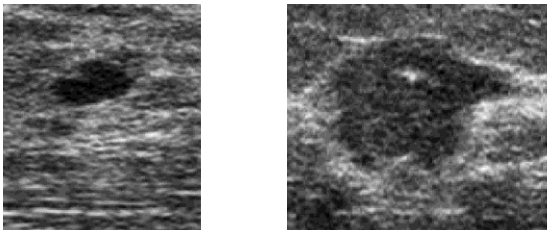
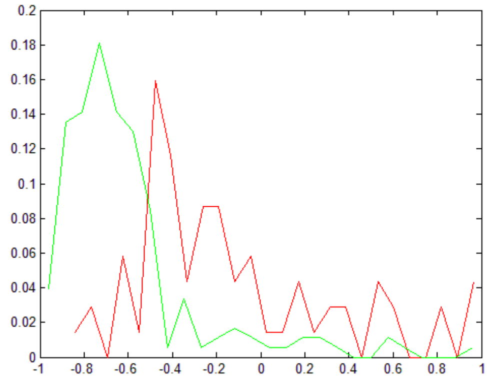


Figure 1: Database examples. Left: benign case; right: malign case.

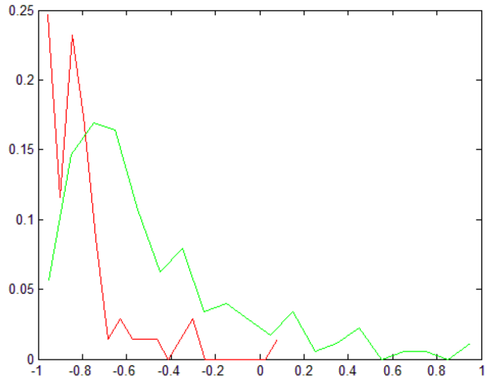
A radiologist determined a rectangular region of interest (ROI) including the tumour and its neighbouring area. Each ROI was segmented using the semi-automatic contour procedure (SAC), based on morphological operators [8].

The seven morphological features extracted were: area ratio (RA), circularity (C), morphological-closing ratio (mShape), normalized radial length standard deviation (D_{NRL}), normalized residual value (nrv), overlap ratio (RS), roughness (R).

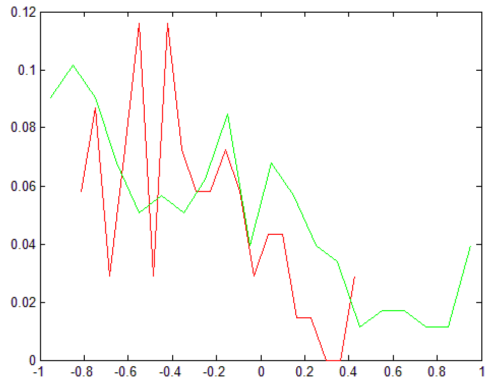
All features were normalized to the interval [-1, 1]. Histograms can be seen in Figure 2.



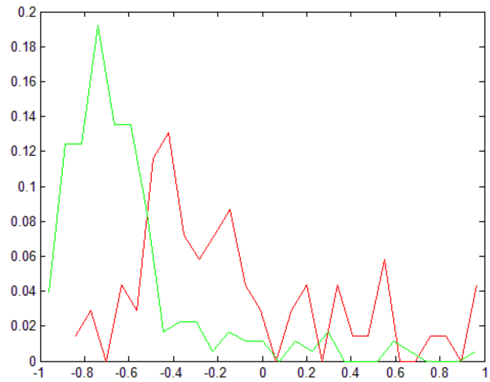
RA



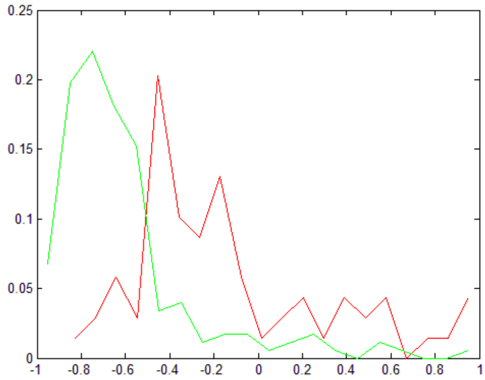
C



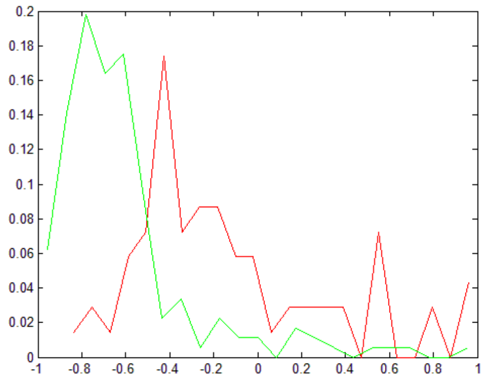
mShape



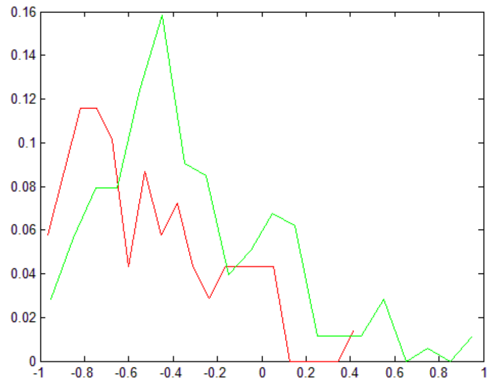
D_{NRL}



nrv



RS



R

Figure 2: Features histogram plots (red are malign examples and green are benign examples) for the database.

Features were used as input to a SVM classifier. Four kernels, linear, polynomial of degree 3, RBF and sigmoid, were tested. Since RBF was the one with best behaviour, only results with this kernel are presented. A grid search was performed over Cost = $2^{-5}, \ldots, 2^{15}$ and $\eta = 2^{-15}, \ldots, 2^{3}$ and each model parameterization was optimized by two-fold cross-validation inside the training set. Cost is a penalty factor for each point misclassified whereas η controls the fitting of the kernel to the data. Additionally, the leave-one-out re-sampling method was applied to assure the reliability and effectiveness of the SVM, considering the number of available samples.

Feature selection was made with a classifier dependent method called Sequential Forward Selection (SFS). SFS, starting with an empty feature set, selects the best single feature and then adds that feature to the feature set. Since it is a strategy that makes local decisions, it cannot be expected to find a globally optimal solution [13]. Best combination in each stage was selected by minimizing the error (number of misclassified samples normalized by the total samples number).

Results

We have first analysed each feature individually. Results are shown in Table 2.

Table 2: SVM individual performance of each of the seven features sorted by Accuracy value.

Feature	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**
nrv	79	84	67
C	77	93	36
D_{NRL}	74	95	19
mShape	72	100	0
RS	72	100	0
R	72	100	0
RA	69	94	3

We have also tested the features combination given in [8], Table 3.

Table 3: SVM performance of the best set of features in [8].

Feature	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**
nrv	79	84	67
nrv+R	79	84	65
nrv+R+C	81	85	71
R+C+D_{NRL}	78	85	59

The order selected by SFS was: nrv, mShape, C, RA, RS, R, and D_{NRL}. Evolution of Accuracy, Sensitivity and Specificity is given in Figure 3.

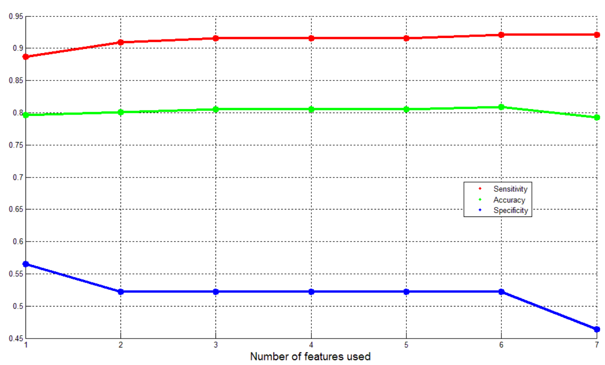


Figure 3: SFS results (accuracy, sensitivity and specificity) in relation to the selected number of features.

The above results indicate that the best combination might be the one composed by the 2 features: nrv and mShape, leading to an accuracy of 87%, sensitivity of

92% and specificity of 77%. A representation of the surface generated is given in Figure 4.

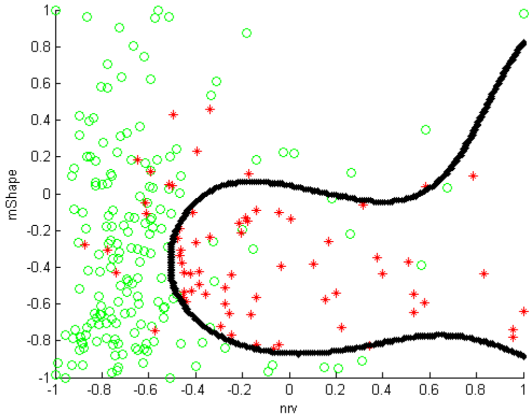


Figure 4: SVM discriminant line example using nrv and mShape. Malign examples in red; benign examples in green and hyperplane in black (Cost = 256; η = 0.5).

Note that the shown hyperplane corresponds to one iteration of the leave-one-out procedure, and thus does not generate (by itself) the presented results.

Some other operating points were generated (by changing Cost and η values). In Figure 5, these results are plotted together with some LDA operating points [8].

Discussion

Breast cancer can be treated most effectively when detected in its early stage. Sonography has become an important adjunct to mammography in breast cancer detection and has been especially useful in distinguishing cysts from solid tumours [14]. This operator-dependent method entails real-time image detection and analysis and requires extensive training and experience in identifying and differentiating between benign and malignant [15].

In this paper, we have built upon a previous work on breast masses classification on US images. Namely, classification was made with Support Vector Machines and feature selection with Sequential Forward Selection.

By observing Table 2, the first conclusion is that the method always provides a high sensitivity having as consequence a low specificity. It means we give priority to classifying correctly the malign cases (we assume it is less harmful to say it is cancer when it is benign than to say it is benign when in fact it is cancer).

Table 3 indicates that mass characterization is not summarized in just one feature. Some features may complete each other and thus the performance increases when they are used together.

For feature selection, while in [8] an exhaustive search was performed, here only a sub-optimum methodology was used. The selection of nrv as the best feature is consistent with the results in [8], however the remaining features do not follow the same order. This demonstrates not only the differences in the classifiers used but also in the feature selection method. Moreover, Alvarenga's goal in [8] was to study which parameters would be more discriminative for the specific dataset, while here the focus is on the development of a classifier.

The second feature to be selected with SFS was mShape, which was not considered as an important feature by LDA in [8]. However, using mutual information as feature selection, mShape was ranked in third place, after nrv and C [16]. D_{NRL} was consistently among the least important features in all of the three studies.

In summary, leave-one-out SVM classification using only nrv and mShape as features reached an accuracy of 87%, sensitivity of 92% and specificity of 77%.

The graphics in Figure 5 suffers from the same shortcomings as ROC curves. Namely, when plotting the performance for all values of the parameters, it is not decided which parameters should be used in practice [17]. It can be seen, however, that for most of the operating points, SVM leads to a higher performance than LDA, shown by the black dots closer to the top-right of Figure 5.

Conclusion

This work indicates that morphological features may be useful to assist physicians in the diagnostic process. However, it is important to bear in mind that the present work evaluated just morphological features. An investigation with both, morphological and texture image characteristics features will be carried out in the future. Another concern is the fact that classes are not equally represented (unbalanced dataset). The effect of this bias will be studied and tests in wider databases will be pursued.

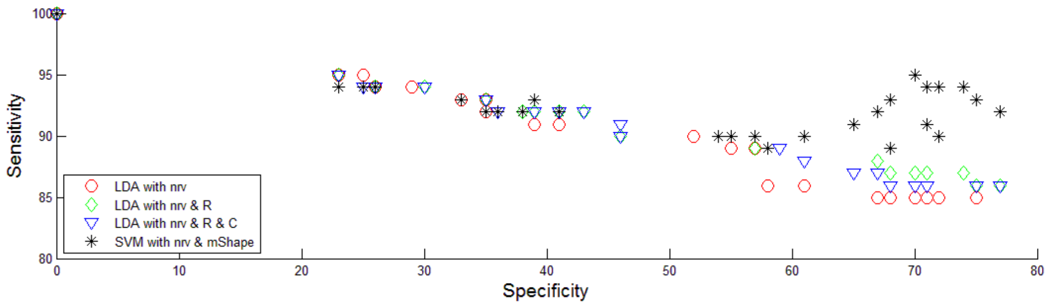


Figure 5: Performance of SVM and LDA [8] for different parameter combinations. Red circles: LDA with nrv; green diamonds: LDA with nrv and R; blue triangles: LDA with nrv, R and C; black dots: SVM with nrv and mShape.

Acknowledgement

This work was partially supported by the Brazilian agencies CNPq, CAPES, FAPERJ and by Fundação para a Ciência e Tecnologia (FCT) - Portugal through projects PTDC/SAU-ENB/114951/2009 and SFRH/BD/70713/2010.

References

[1] C. DeSantis, R. Siegel, P. Bandi, and A. Jemal, "Breast cancer statistics," *CA: A cancer journal for clinicians*, vol. 61, no. 6, pp. 409–418, 2011.

[2] INCa, "Estimativa 2012 – Incidência de câncer no Brasil – Síntese de resultados e comentários – Câncer de mama feminino, Ministério da Saúde," 2011. [Online]. Available: http://www.inca.gov.br/estimativa/2012/index.asp?ID=5.

[3] B. Lairenjam and S. Wasan, "Neural Network with Classification Based on Multiple Association Rule for Classifying Mammographic Data," *Intelligent Data Engineering and Automated Learning (IDEAL)*, 2009.

[4] "Fact sheet No. 297: Cancer," *World Health Organization*, 2009.

[5] "Health statistics: atlas on mortality in the European Union," *Eurostat*, 2009.

[6] W. A. Berg, Z. Zhang, D. Lehrer, R. A. Jong, E. D. Pisano, R. G. Barr, M. Böhm-Vélez, M. C. Mahoney, W. P. Evans, L. H. Larsen, M. J. Morton, E. B. Mendelson, D. M. Farria, J. B. Cormack, H. S. Marques, A. Adams, N. M. Yeh, and G. Gabrielli, "Detection of Breast Cancer With Addition of Annual Screening Ultrasound or a Single Screening MRI to Mammography in Women With Elevated Breast Cancer Risk," *JAMA*, vol. 307, no. 13, pp. 1394–1404, 2012.

[7] D. R. Chen, R. F. Chang, C. J. Chen, M. F. Ho, S. J. Kuo, S. T. Chen, S. J. Hung, and W. K. Moon, "Classification of breast ultrasound images using fractal feature," *Clinical Imaging*, vol. 29, no. 4, pp. 235–245, 2005.

[8] A. V. Alvarenga, A. F. C. Infantosi, W. C. A. Pereira, and C. M. Azevedo, "Assessing the performance of morphological parameters in distinguishing breast tumors on ultrasound images," *Medical engineering & physics*, vol. 32, no. 1, pp. 49–56, 2010.

[9] T. Wan, R. Liao, and Z. Qin, "A robust feature selection approach using low rank matrices for breast tumors in ultrasonic images," in *18th IEEE International Conference on Image Processing*, 2011.

[10] R. Liao, T. Wan, and Z. Qin, "Classification of Benign and Malignant Breast Tumors in Ultrasound Images Based on Multiple Sonographic and Textural Features," in *International Conference on Intelligent Human-Machine Systems and Cybernetics*, 2011, vol. 1, pp. 71–74.

[11] W. Gomez, W. C. de A. Pereira, A. F. C. Infantosi, and A. Diaz-Perez, "Computerized diagnosis of breast lesions on ultrasonography," in *XXII Congresso Brasileiro de Engenharia Biomedica (CBEB)*, 2010.

[12] X. Shi, H. D. Cheng, L. Hu, W. Ju, and J. Tian, "Detection and classification of masses in breast ultrasound images," *Digital signal processing*, vol. 20, no. 3, pp. 824–836, 2010.

[13] M. R. N. Kousarrizi, F. Seiti, and M. Teshnehlab, "An Experimental Comparative Study on Thyroid Disease Diagnosis Based on Feature Subset Selection and classification," *International Journal of Electrical & Computer Sciences (IJECS-IJENS)*, vol. 12, no. 1, pp. 13–19, 2012.

[14] R. Chang, W. Wu, W. K. Wu, and D. Chen, "Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors," *Breast Cancer Research and Treatment*, vol. 89, no. 2, pp. 179–185, 2005.

[15] M. J. G. Calas, R. M. V. R. Almeida, B. Gutfilen, and W. C. de A. Pereira, "Interobserver concordance in the BI-RADS classification of breast ultrasound exams," *Clinics*, vol. 67, no. 2, 2012.

[16] W. C. de A. Pereira, A. V. Alvarenga, A. F. C. Infantosi, L. Macrini, and C. E. Pedreira, "A non-linear morphometric feature selection approach for breast tumor contour from ultrasonic images," *Computers in Biology and Medicine*, vol. 40, no. 11–12, pp. 912–918, 2010.

[17] S. Bengio, J. Mariethoz, and M. Keller, "The Expected Performance Curve," in *Proceedings of the ICML 2005 workshop on ROC Analysis in Machine Learning*, Bonn, Germany, 2005.