Classification of breast tumours using morphological characteristics extracted from MR images

Ana Catarina Pelicano[1], Nuno A. M. Araújo[2], Raquel C. Conceição[1]

[1]Instituto de Biofísica e Engenharia Biomédica, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisbon, Portugal; [2]Centro de Física Teórica e Computacional, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016, Lisbon, Portugal

Abstract

In this work, we present a framework for automatic tumor segmentation of Magnetic Resonance (MR) images of the breast, as well as the study of three classifiers – Linear Discriminant Analysis (LDA), Support Vector Machines (SVMs) and *k*-Nearest Neighbours (*k*NN) – to differentiate between malignant and benign lesions. The dataset in this study comprises 24 lesions: 12 malignant and 12 benign. The 3D Region Growing algorithm was adopted to perform automatic segmentation of lesions and 16 morphological features were considered for lesion classification. The Mann-Whitney U-test was employed for feature selection, and the classifiers' performance evaluated with accuracy, sensitivity, specificity, F1-score and Matthew's Correlation Coefficient (MCC) metrics. *k*NN (with *k*=6 and Chebychev distance) outperformed the other classifiers with an accuracy, sensitivity and specificity of 87.5%, 83.3% and 91.7%, respectively.

Introduction

Breast Cancer

Over 2.2 million new cases – Most common types of cancer worldwide in 2020[1]

Over 684.000 deaths – Fifth deadliest type of cancer worldwide in 2020[1]

Manual Segmentation – Gold standard procedure[2]

- ✗ Time consuming;
- ✗ Prone to human error;
- ✗ High rates of false positives;
- ✗ Inter-observer variability.

Breast Tumour Automatic Segmentation Algorithms[3]

- ✓ Fast data processing algorithms;
- ✓ Reduces subjectivity on results.

Breast Tumours Morphological Characteristics[4]

- Benign: Round shape; Smooth Margins
- Malignant: Irregular shape; Irregular Margins

Main goal:

Differentiate between malignant and benign lesions using morphological features extracted from MR images.

Segmentation Methodology and Results

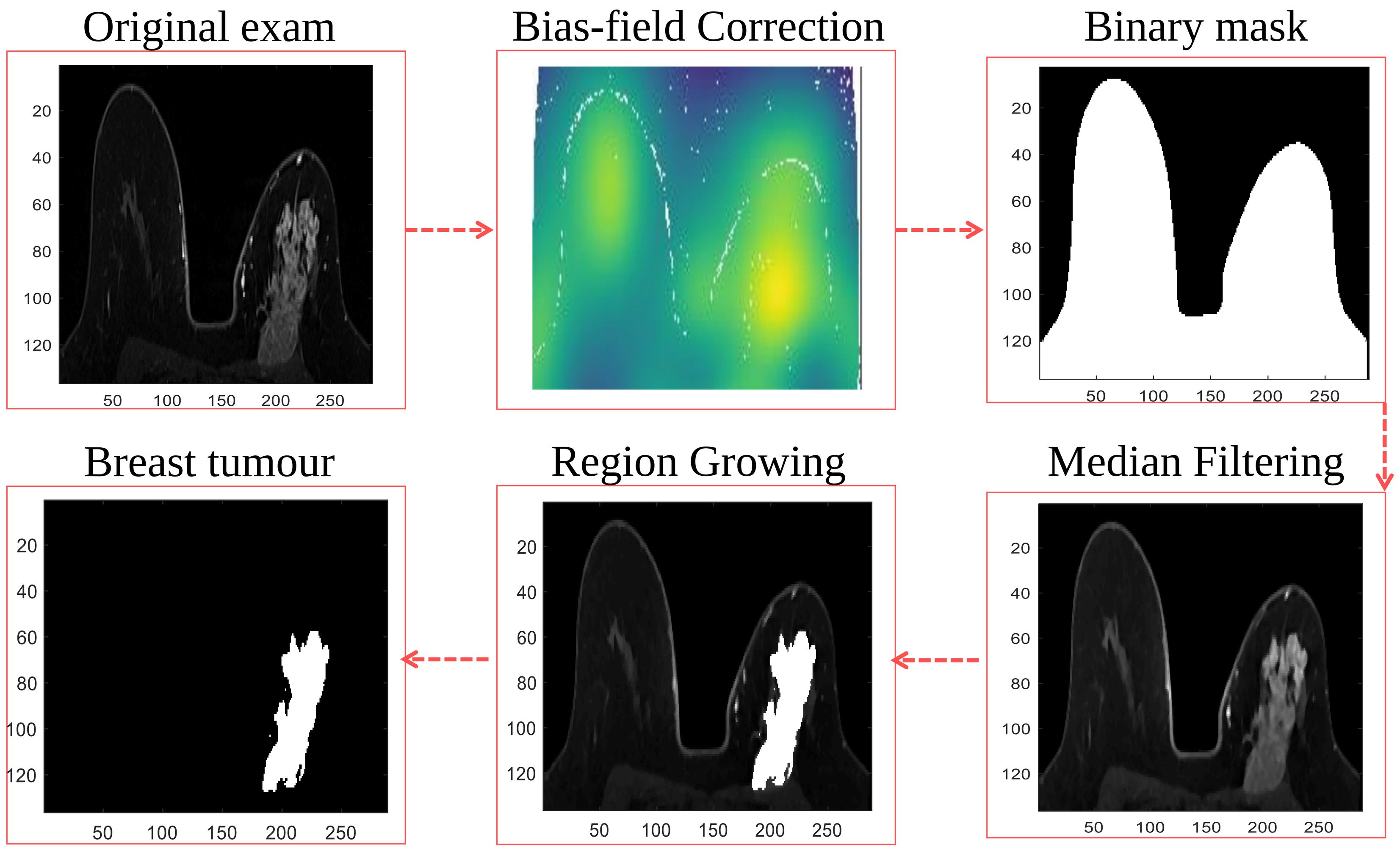

Figure 1. Breast tumour segmentation pipeline.[5]

Classification Methodology and Results

1 3D Morphological features selection

Features:

- Volume;
- Surface area;
- Compactness;
- Radius of gyration;
- Inertia tensor;
- Number of voxels;
- Largest axis;
- NRL mean;
- Sphericity;
- NRL entropy;
- NRL ratio;
- Roughness.

→ **Mann-Whitney U-Test:**

H_0: $M_{benign} = M_{malignant}$

H_1: $M_{benign} \neq M_{malignant}$

α = 1%

$if\ p - value\ < \alpha$:

$\therefore$ ***reject*** $\boldsymbol{H_0}$

– statistically signifcant diference between malignant and benign breast tumours.

→ **Selected Features:**

- Volume;
- Surface area;
- Compactness;
- Radius of gyration;
- Inertia Tensor;
- Number of voxels;
- Largest axis;
- NRL entropy.

Figure 2. Schematics of morphological features selection method.[6,7]

2 Classification Algorithms

Linear Discriminant Analysis (LDA)[8]

Data Pre-Processing: Standardization (zero mean and unit variance) → LDA → Cross Validation: Leave-One-Out

Support Vector Machines (SVMs)[9]

Data Pre-Processing: Standardization (zero mean and unit variance) → SVMs: Kernels – linear and RBF → Cross Validation: Leave-One-Out

***K*-Nearest Neighbour (kNN)[10]**

Data Pre-Processing: Standardization (zero mean and unit variance) → *k*NN: Distances – Cityblock, Hamming, Chebychev, Euclidean. *k* = 1,2,…11 → Cross Validation: Leave-One-Out

Figure 3. Schematics of the adopted classification methodology.

3 Results

Table 1. Optimal diagnostic performance of LDA, SVMs and *k*NN classifiers.

	Parameters	Accuracy	Sensitivity	Specificity	F1-score	MCC
LDA	——	0.5833	0.4167	0.7500	0.5000	0.1768
SVMs	Kernel – RBF	0.8333	0.7500	0.9167	0.8182	0.6761
*k*NN	Distance – Chebychev, *k* = 6	**0.8750**	**0.8333**	**0.9167**	**0.8696**	**0.7526**

Conclusions

- The developed segmentation framework is suitable to segment tumours with a varying level of heterogeneity regarding voxel intensity.
- From the initially considered 16 morphological features, only 8 were found statistically significant for a significance level of 0.01.
- The *k*NN classifier (with *k*=6 and Chebychev distance) outperformed LDA and SVMs classifiers.
- Future work will include morphologic and texture features to study and optimise the performance of the classifiers.

References

[1] International Agency for Research on Cancer, "World Health Organization, GLOBOCAN 2020: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2020 – Cancer Fact Sheets" [Online]. Available: https://gco.iarc.fr/today/fact-sheets-cancers Accessed 06/16/2021.**[2]** Cheng HD et al. (2003), "Computer-aided detection and classification of micro calcification in mammograms: a survey," **[3]** Thakran S et al. (2018) "Automatic outer and inner breast tissue segmentation using multi-parametric MRI images of breast tumor patients". **[4]** Ikeda DM et al (2001) "Development, standardization, and testing of a lexicon for reporting contrast-enhanced breast magnetic resonance imaging studies". **[5]** Gonzalez R and Woods R. Digital Image Processing. New Jersey, Prentice Hall, 2002. **[6]** Nie K et al. (2008) "Quantitative analysis of lesion morphology and texture features for diagnostic prediction in breast MRI". **[7]** Fusco R et al. (2018) "Use of Quantitative Morphological and Functional Features for Assessment of Axillary Lymph Node in Breast Dynamic Contrast-Enhanced Magnetic Resonance Imaging". **[8]** Hair JF et al. Multivariate Data Analysis. New Jersey, Prentice Hall; 2006. **[9]** Cortes C, Vapnik V. (1995) "Support-vector networks" **[10]** Flach P. Machine Learning, The Art and Science of Algorithms that Make Sense of Data. Cambridge University Press; 2012

Acknowledgements:

This research was funded by Fundação para a Ciência e Tecnologia (FCT) under the fellowship UI/BD/150762/2020, FCT/MET (PIDDAC) under the Strategic Program UIDB/00645/2020 and under Contracts no. PTDC/FIS-MAC/28146/2017 (LISBOA-01-0145-FEDER-028146), UIDB/00618/2020, and UIDP/00618/2020. The authors would like to acknowledge the study with references CES/44/2019/ME (19/09/2019) and CES/34/2020/ME (06/11/2020) in Hospital da Luz Lisboa and the technicians for the MRI acquisitions.

Ciências ULisboa · IBEB · FCT Fundação para a Ciência e a Tecnologia · CFTC Centro de Física Teórica e Computacional

Poster Session
Encontro Ciência 2021