

Computer-aided Diagnosis in Breast MRI: Do Adjunct Features Derived from T_2 -weighted Images Improve Classification of Breast Masses?

Willem van Aalst¹, Thorsten Twellmann¹, Hans Buurman²,
Frans A. Gerritsen^{1,2}, Bart M. ter Haar Romeny¹

¹Dept. of Biomedical Engineering, Eindhoven University of Technology, NL

²Philips Medical Systems, Healthcare Informatics - Advanced Development, Best, NL

t.twellmann@tue.nl

Abstract. In the field of dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) of breast cancer, current research efforts in computer-aided diagnosis (CADx) are mainly focused on the temporal series of T_1 -weighted images acquired during uptake of a contrast agent, processing morphological and kinetic information. Although static T_2 -weighted images are usually part of DCE-MRI protocols, they are seldom used in CADx systems. The aim of this work was to evaluate to what extent T_2 -weighted images provide complementary information to a CADx system, improving its performance for the task of discriminating benign breast masses from life-threatening carcinomas. In a preliminary study considering 64 masses, inclusion of lesion features derived from T_2 -weighted images increased the classification performance from $A_z=0.94$ to $A_z=0.99$.

1 Introduction

In recent years, dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) has evolved to a powerful tool for detection, diagnosis and management of breast cancer [1]. A temporal series of 3D T_1 -weighted MR images is acquired during the uptake of a contrast agent based on gadolinium chelates. This enables radiologists to localize suspicious tissue areas and to further assess their likelihood of malignancy by means of their morphologic and kinetic characteristics. Static T_2 -weighted images provide complementary information for discriminating benign fibroadenoma from malignant cancers, as both expose similar enhancement patterns in the DCE T_1 -weighted images. Furthermore, several secondary signs such as edemas, which frequently accompany pathological tissue growth, are accentuated in T_2 -weighted images [2].

The role of T_2 -weighted images for the general discrimination of benign and malignant findings was investigated in several studies, e.g. [3, 4]. In the majority of studies, suspicious masses were initially located by means of the DCE T_1 -weighted images. Subsequently, the brightness of the corresponding region in the T_2 -weighted image was visually rated as hypo-, iso- or hyperintense to a

reference region, e.g. the tissue surrounding the mass or air. In summary, the different studies reached different conclusions allowing no clear statement about the added value of T_2 -weighted images for the discrimination of benign and malignant breast masses.

Computer-aided diagnosis (CADx) systems supporting radiologists in detecting and assessing suspicious tissue regions in DCE-MRI data are expected to improve the reliability of clinical decisions and, therewith, to potentially decrease the number of unnecessary biopsies. (Semi-)automatic discrimination of benign and malignant masses in breast MRI was investigated by several authors [5, 6, 7]. The proposed approaches have in common that they solely concentrate on processing of kinetic and morphologic features as depicted in the DCE T_1 -weighted images, which are, without doubt, the most important source of information in breast MRI. However, the inclusion of the frequently acquired T_2 -weighted images might be beneficial in the context of CADx.

This work investigates for the first time whether adjunct features derived from static T_2 -weighted images improve the performance of computer-aided classification of benign and malignant breast masses based on supervised pattern recognition techniques.

2 Methods

In bilateral DCE-MRI studies¹ of 50 patients, $N = 64$ breast masses (28 benign and 36 malignant) were segmented in the T_1 -weighted images with a semi-automatic segmentation algorithm [8]. Additionally, binary masks separating breasts from thoraces were computed by applying a sequence of smoothing, thresholding and morphological operators. To adjust for the different voxel size in the T_1 - and T_2 -weighted images, the latter were resampled by tri-linear interpolation. Sufficient spatial alignment between the dynamic T_1 -weighted series and between T_1 - and T_2 -weighted images was verified by visual inspection.

2.1 T_1 - and T_2 -weighted Lesion Features

A total of 22 morphological and kinetic T_1 -weighted features described in [5, 6, 7] were evaluated for all binary lesion segmentations. Due to the limited space we have to refer the reader to the corresponding articles for more information about these features. In addition to the T_1 -weighted features, the following T_2 -weighted lesion features were computed: Let I_{L_i} be the set of grey values in the T_2 -weighted image corresponding to the voxels as marked by a binary segmentation of lesion i , I_{B_i} the set of grey values in the left or right breast (depending on the location of the lesion i) excluding the lesion and $I_{S_i^t}$ the set of grey values in a t

¹ Dynamic T_1 -weighted images: TR=4.4-6.9ms, TE=2.1-3.4ms, FA=10°, Δt =60-90s, 6-7 dynamics, fat-suppressed, transversal orientation. Static T_2 -weighted image: TR=6.7-16.9s, TE=120ms, FA=90°, transversal orientation.

mm thick shell enclosing the segmentation of lesion i . The *normalized intensity mean* and *variance* features are defined as

$$f_{\mu}^{T_2}(\Gamma_{L_i}, \Gamma_{R_i}) = \frac{\mu_{\Gamma_{L_i}}}{\mu_{\Gamma_{R_i}}} \quad \text{and} \quad f_{\sigma^2}^{T_2}(L_i, R_i) = \frac{\sigma_{\Gamma_{L_i}}^2}{[\mu_{\Gamma_{R_i}}]^2},$$

with μ and σ^2 being the intensity mean and variance of the lesion set Γ_{L_i} or one of the reference regions $\Gamma_{R_i} \in \{\Gamma_{B_i}, \Gamma_{S_i^t}\}$. By normalizing with respect to the mean intensity of a reference region, we take into account the qualitative nature of MR images and resemble the clinical practice of rating lesions as hypo-, iso- or hyperintense with respect to some reference region.

The *normalized α -percentile* is defined as

$$f_{perc_{\alpha}}^{T_2}(\Gamma_{L_i}, \Gamma_{B_i}, \alpha) = \frac{H_{\Gamma_{L_i}}^{\alpha}}{\mu_{\Gamma_{B_i}}},$$

with $H_{\Gamma_{L_i}}^{\alpha}$ being the intensity threshold in the cumulative relative frequency histogram H of lesion Γ_L for which the cumulative relative frequency exceeds α percent. A second normalized α -percentile is computed for the enclosing shell, reflecting potential bright regions (caused e.g. edema) in the tissue surrounding the lesion:

$$f_{perc_{\alpha}}^{T_2}(\Gamma_{S_i^t}, \Gamma_{B_i}, \alpha) = \frac{H_{\Gamma_{L_i}}^{\alpha}}{\mu_{\Gamma_{B_i}}}.$$

Both α -percentiles were normalized with respect to the mean breast intensity.

The aforementioned features were evaluated for the sets Γ_{L_i} containing voxels as marked by the binary segmentation of lesion i and for the subsets $\Gamma_{L_i}^* \subset \Gamma_{L_i}$ of lesion voxels as marked after applying a morphological erosion operator to the binary lesion mask. The free parameters t and α were varied in reasonable ranges ($t \in \{2, 4, \dots, 20\}$ in *mm*, $\alpha \in \{0, 2, \dots, 20, 80, 82, \dots, 100\}$ in %), considering each setting as a new feature.

The sharpness of the lesion margin in the T_2 -weighted image is quantified by evaluating the *mean* and *variance of the margin gradient* as described for DCE T_1 -weighted images in [5] in the static T_2 -weighted images. These features reflect the mean and variance of gradient magnitudes computed with Gaussian derivatives of reasonable scale for all voxels in a three voxels thick shell centered on the lesion segmentation's surface.

2.2 Feature Selection and Classification

Due to the limited number of lesions, feature selection is a mandatory step to avoid overfitting of the data. For the final classification of lesions we selected two sets of features, each containing five features in total. A linear classification function was estimated by Fisher linear discriminant analysis (FDA) [9].

To determine the baseline performance of a CADx system based on T_1 -weighted information only, a set of three T_1 -weighted features was selected by

exhaustive search and subsequently extended to a size of five features by forward-backward feature selection. At each iteration, the performance of the FDA after including a new candidate feature was evaluated in a leave-one-lesion-out scheme. For all permutations of the N lesions, the FDA was trained with data of $N - 1$ lesions and subsequently applied to the omitted unseen lesion. The final performance was quantified by estimating the area-under-the-ROC-curve (A_z) computed for the FDA output [10].

The same feature selection process was repeated a second time, this time selecting candidate features from the pooled set of T_1 - and T_2 -weighted features.

3 Results

Figure 1 shows the ROC curves for the FDA evaluating the selected set of five DCE T_1 -weighted features (dashed line) and the mixed set of DCE T_1 - and static T_2 -weighted features (solid lines). The non-parametric estimates of the area-under-the-ROC curve are $A_z = 0.94$ and $A_z = 0.99$, respectively. Though inclusion of the T_2 -weighted features leads to a considerable improvement of $\Delta A_z = 0.05$, a statistical test for the significance of the difference between the A_z values of the two correlated ROC curves [11] failed with $p = 0.064$ (CI=95%).

4 Discussion

After feature selection from the pooled set, the final set consists of two DCE T_1 -weighted and three T_2 -weighted features. The DCE T_1 -weighted features are *circularity* and *variance of the radial gradient histogram* as described in [5], reflecting the lesion shape and the sharpness of its margin, respectively. The

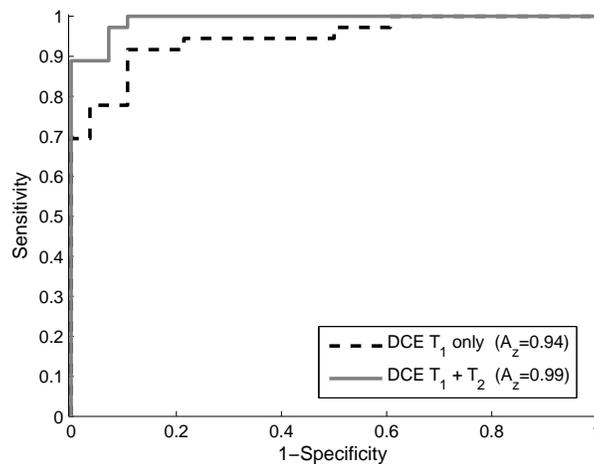


Fig. 1. Results

three T_2 -weighted features are the normalized α -percentile for the eroded lesion segmentation ($\alpha = 20\%$) and the normalized α -percentile for the lesion shell with ($\alpha = 92\%, t = 2\text{mm}$) and ($\alpha = 98\%, t = 20\text{mm}$). The latter features resemble clinical practice of rating the intensities within the lesion and within the enclosing shell of non-lesion tissue.

For the considered set of lesions, inclusion of T_2 -weighted features improved the performance of the classification of breast masses by $\Delta A_z = 0.05$ to $A_z = 0.99$, yet a statistical test for significance failed. In consideration of the low power of the test, we nevertheless believe that inclusion of information derived from the T_2 -weighted images, which are usually part of today's breast MRI protocols, provides valuable complementary information to CADx systems and we will reinvestigate this issue in a future larger study.

Acknowledgement. This work was conducted in cooperation with the University of Chicago; the authors are grateful to Dr. Gillian Newstead (UoC) and Ursula Kose (PMS) for their support.

References

1. Kuhl CK. The current status of breast MR imaging (Part I). *Radiology*. 2007;244(2):356–78.
2. Fischer DR, Wurdinger S, Boettcher J, et al. Further signs in the evaluation of magnetic resonance mammography: A retrospective study. *Invest Radiol*. 2005;40(7):430–5.
3. Kuhl CK, Klaschik S, Mielcarek P, et al. Do T2-weighted pulse sequences help with the differential diagnosis of enhancing lesions in dynamic breast MRI? *J Magn Reson Imaging*. 1999;9(2):187–96.
4. Kelcz F. Quantitative assessment of T2 imaging information in differential diagnosis of enhancing breast lesions. *Eur Radiol*. 2006;16 (Suppl 5):E51–E53.
5. Gilhuijs KGA, Giger ML, Bick U. Computerized analysis of breast lesions in three dimensions using dynamic magnetic-resonance imaging. *Med Phys*. 1998;25:1647–54.
6. Chen W, Giger ML, Lan L, et al. Computerized interpretation of breast MRI: Investigation of enhancement-variance dynamics. *Med Phys*. 2004;31(5):1076–82.
7. Arbash-Meinell L, Stolpen AH, Berbaum KS, et al. Breast MRI lesion classification: Improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system. *J Magn Reson Imaging*. 2007;25:89–95.
8. Bülow T, Meinell LA, Wiemker R, et al. Segmentation of suspicious lesions in dynamic contrast-enhanced breast MR images. In: *Proc. of SPIE Medical Imaging 2006*; 2006.
9. Duda RO, Hart PE. *Pattern Classification and Scene Analysis*. Wiley; 1973.
10. Swets JA, Pickett RM. *Evaluation of Diagnostic Systems*. Academic Press; 1982.
11. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristics curves: A nonparametric approach. *Biometrics*. 1988;44:837–45.