Medical Physics
The International Journal of Medical Physics Research and Practice

Use of volumetric features for temporal comparison of mass lesions in full field digital mammograms

Jelena Bozek, Michiel Kallenberg, Mislav Grgic, and Nico Karssemeijer

Citation: Medical Physics **41**, 021902 (2014); doi: 10.1118/1.4860956
View online: http://dx.doi.org/10.1118/1.4860956
View Table of Contents: http://scitation.aip.org/content/aapm/journal/medphys/41/2?ver=pdfcov
Published by the American Association of Physicists in Medicine

Use of volumetric features for temporal comparison of mass lesions in full field digital mammograms

Jelena Bozek[a)]
Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, HR-10000 Zagreb, Croatia

Michiel Kallenberg
Department of Radiology, Radboud University Nijmegen Medical Centre, Geert Grooteplein Zuid 18, 6525 GA Nijmegen, The Netherlands

Mislav Grgic
Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, HR-10000 Zagreb, Croatia

Nico Karssemeijer
Department of Radiology, Radboud University Nijmegen Medical Centre, Geert Grooteplein Zuid 18, 6525 GA Nijmegen, The Netherlands

(Received 26 June 2013; revised 18 December 2013; accepted for publication 18 December 2013; published 10 January 2014)

Purpose: Temporal comparison of lesions might improve classification between benign and malignant lesions in full-field digital mammograms (FFDM). The authors compare the use of volumetric features for lesion classification, which are computed from dense tissue thickness maps, to the use of mammographic lesion area. Use of dense tissue thickness maps for lesion characterization is advantageous, since it results in lesion features that are invariant to acquisition parameters.
Methods: The dataset used in the analysis consisted of 60 temporal mammogram pairs comprising 120 mediolateral oblique or craniocaudal views with a total of 65 lesions, of which 41 were benign and 24 malignant. The authors analyzed the performance of four volumetric features, area, and four other commonly used features obtained from temporal mammogram pairs, current mammograms, and prior mammograms. The authors evaluated the individual performance of all features and of different feature sets. The authors used linear discriminant analysis with leave-one-out cross validation to classify different feature sets.
Results: Volumetric features from temporal mammogram pairs achieved the best individual performance, as measured by the area under the receiver operating characteristic curve (A_z value). Volume change ($A_z = 0.88$) achieved higher A_z value than projected lesion area change ($A_z = 0.78$) in the temporal comparison of lesions. Best performance was achieved with a set that consisted of a set of features extracted from the current exam combined with four volumetric features representing changes with respect to the prior mammogram ($A_z = 0.90$). This was significantly better ($p = 0.005$) than the performance obtained using features from the current exam only ($A_z = 0.77$).
Conclusions: Volumetric features from temporal mammogram pairs combined with features from the single exam significantly improve discrimination of benign and malignant lesions in FFDM mammograms compared to using only single exam features. In the comparison with prior mammograms, use of volumetric change may lead to better performance than use of lesion area change.
© *2014 American Association of Physicists in Medicine*. [http://dx.doi.org/10.1118/1.4860956]

Key words: volumetric features, dense tissue thickness, temporal comparison, digital mammography, CAD

1. INTRODUCTION

Computer-aided diagnosis (CAD) systems with temporal analysis of mammograms have the potential to help radiologists in the interpretation of benign and malignant masses[1] and to significantly improve radiologists' accuracy.[2]

Mass lesions are defined with wide range of characteristics such as density (fat containing, low density, isodense, high density), margins (circumscribed, microlobular, obscured, indistinct, spiculated), and shape (round, oval, lobular, irregular).[3] Depending on those characteristics lesions can be classified as benign or malignant. Round lesions of low density with circumscribed and well-defined margins usually present a benign change. Presence of a fatty halo which is a fine radiolucent line surrounding a mass in some cases might as well be a benign sign.[4] Malignant lesions are usually irregularly shaped with a spiculated and indistinct boundary. Change in the appearance of a lesion and its growth between two mammographic screening examinations might be an indication of malignancy. On the other hand, benign lesions usually remain stable in appearance and size. Thus, changes in the appearance and size of lesions might give valuable information for their classification and temporal features may be very useful for the development of CAD systems.

Few studies that use temporal information for classifying lesions have been reported. Hadjiiski *et al.*[5] used run

length statistics (RLS) texture features, spiculation features, and morphological features to discriminate between benign and malignant lesion in digitized screen-film mammograms. Their feature space consisted of the current RLS features, the difference RLS features, the current and prior spiculation features, and the current and prior mass sizes. The information on the prior mammograms significantly ($p = 0.015$) increased the accuracy for classification of the masses from A_z value (area under the receiver operating characteristic curve) of 0.82 when only single exam features were used to A_z value of 0.88 when temporal features were used.

Timp *et al.*[6] designed two kinds of temporal features: difference features and similarity features. Difference features indicated the relative change in feature values between prior and current views. Similarity features measured whether two regions were comparable in appearance. The classification performance significantly ($p = 0.005$) increased when using temporal features compared to the single exam classification. They obtained an average A_z value of 0.74 without temporal features and 0.77 with the use of temporal features. They used digitized screen-film mammograms in their analysis.

In this study, we performed temporal comparison of lesions in full-field digital mammograms (FFDM). We extracted temporal features that characterize change in the lesion between two mammographic examinations. Our proposed method has some important advantages. First, the analysis is performed using only FFDM mammograms. Other studies on temporal comparison used digitized mammograms. Second, we introduce volumetric change of a lesion which is determined using dense tissue thickness maps. We developed four volumetric features that contain information about the size of the lesion and that are robust considering the segmentation of the lesion and the surrounding tissue. Moreover, volumetric features overcome the limitations introduced when measuring size of a lesion using only area, which is a two-dimensional projection of the three-dimensional object. To our knowledge, no other studies that include volumetric information in the temporal analysis of digital mammograms have been reported.

2. MATERIALS AND METHODS

2.A. Dataset

Digital mammograms for this study were collected from the Foundation of Population Screening Mid-West, The Netherlands, where they were acquired with a Hologic Selenia FFDM system. The dataset comprised 60 cases with mammogram pairs consisting of craniocaudal (CC) and mediolateral oblique (MLO) views taken in two subsequent screening rounds. The most recent mammogram we denote as current mammogram, and mammogram taken in the previous examination as prior mammogram. Average period between two subsequent exams is 2 years.

All mammograms used in the study have a lesion that has been proven as benign or malignant by biopsy performed on the current mammogram. For this study, we only selected cases in which a lesion was visible both in the current and in the prior mammograms. The number of benign cases was 37 and the number of malignant cases was 23. There was a total of 65 lesions, of which 41 were benign and 24 malignant, i.e., all cases had one lesion visible, except four benign cases and one malignant case which had two lesions visible in both prior and current mammograms. Lesions in the current mammograms are manually matched with the ones in the prior mammograms.

In this study under the term lesion we consider masses, architectural distortions, and focal asymmetries. We included only lesions that are projected within the breast area, i.e., not overlapping with the pectoral muscle. We did not include lesions larger than 4 cm in the diameter since for those lesions our automated segmentation routine was not designed. The average size of lesions in current mammograms was 1.58 cm^2 and the smallest and the biggest lesions were 0.27 and 5.44 cm^2, respectively. All FFDM mammograms were downsampled to a resolution of 200 μm using bilinear interpolation.

2.B. Feature computation

The center location of each region that contained a lesion was computed as the geometrical center of the contour delineated by the radiologist. The computed center of a lesion was used as a seed point for automated segmentation. The segmentation method was based on the region boundary information and gray level distribution in a region of interest around the lesion. The best contour was selected using an optimization technique known as dynamic programming. The method is explained in detail in Ref. 7.

For each current and prior segmented region, nine different features were calculated to characterize the lesion. The summary of the features is presented in Table I and feature computation is explained in detail in Subsections 2.B.1–2.B.5.

2.B.1. Spiculation feature

Since malignant lesions have stellate pattern with lines radiating from the center of the lesion, we computed a feature

TABLE I. Summary of used features.

Feature	Description
$\overline{f1}$	Mean value of concentration of spicules (feature $f1$) inside segmented lesion
Contrast	Absolute contrast between the lesion and the surrounding area
Fourier descriptor	Shape feature based on Fourier descriptors
Area	Area of segmented lesion
Compactness	Shape feature measuring efficiency of contour to contain a given area
Vol	Volume of segmented lesion
VolDilRing	Volume normalized with a dilated ring (width 2 mm) around the lesion
VolDilCont2	Volume normalized with a dilated contour that is 2 mm from the lesion
VolDilCont3	Volume normalized with a dilated contour that is 3 mm from the lesion

for detection of spiculations.[8] The spiculation feature that detects stellate patterns of straight lines was derived from the map of pixel orientations which was computed using directional second order Gaussian derivatives. The computed feature $f1$ was the normalized measure of the fraction of pixels with a line orientation toward the center of the lesion. The feature used in our study was $\overline{f1}$ which is the mean value of $f1$ inside the segmented lesion.

2.B.2. Contrast feature

One of the commonly used features in CAD systems is contrast.[9] In our analysis, we used contrast computed as the difference between mean intensity of the lesion and mean intensity of the surrounding area.

2.B.3. Fourier descriptor

Fourier descriptor is a shape feature based on the Fourier transform of the object boundary sequence. Irregular contours result in larger values and round contours result in smaller values. The feature computed in our study was based on normalized Fourier descriptors defined in Ref. 10.

2.B.4. Area and compactness

Area is a commonly used feature for representing the size of a lesion in CAD systems.[9] The computation of area of a lesion was straightforward and it was obtained from the segmented region. Compactness or perimeter-to-area ratio is a simple measure that expresses efficiency of a contour to contain certain area. Compactness was computed using the expression

$$C = \frac{P^2}{A}, \tag{1}$$

where P is the perimeter and A the area of the segmented lesion.

2.B.5. Volumetric features

As it was shown in our previous work,[11] volume might be a better feature than area for representing size of a lesion and for distinguishing between benign and malignant lesions. Volume as a feature is more robust compared to area. Thus, if lesions are embedded in fatty tissue, the volume estimates will not be affected even if a lesion is oversegmented since the dense tissue thickness around the lesion will be zero. In order to compute the volume of a lesion, we first determined the dense tissue thickness map. The thickness of dense tissue was estimated using method proposed previously.[12] The method is based on a physical model of image acquisition and assumes that the breast is composed of two types of tissue, dense fibroglandular tissue and fatty tissue. For each type of tissue, the effective attenuation coefficient was computed as a function of the anode and filter material, tube voltage and breast thickness. The thickness of dense tissue at location $\mathbf{r}$ was then computed as

$$h_d(\mathbf{r}) = -\frac{1}{\mu_{d,\text{eff}} - \mu_{f,\text{eff}}} \ln \frac{\bar{g}(\mathbf{r})}{\bar{g}_f}, \tag{2}$$

where $\mu_{f,\text{eff}}$ and $\mu_{d,\text{eff}}$ are effective attenuation coefficients for fatty and dense tissue, respectively, $\bar{g}(\mathbf{r})$ is pixel value of tissue at location $\mathbf{r}$, and $\bar{g}_f$ is the mean value of fatty tissue determined in a reference region in the mammogram.

Volume of the lesion was computed by multiplying obtained dense tissue thickness $h_d(\mathbf{r})$ at location $\mathbf{r}$ in the lesion region S with the size A_0 of a pixel as

$$V = A_0 \cdot \sum_S h_d(\mathbf{r}), \tag{3}$$

where the pixel size A_0 used in our study is 0.02×0.02 cm.

In order to correct the volume of a lesion for overlapping dense tissue, we normalized lesion volume and thereby obtained three additional volumetric features. We defined three regions around the lesion as depicted in Fig. 1 which were used for normalizing the volume of the lesion.

The first region R_1 formed a dilated ring or a band around the lesion and contained all pixels within the distance of 2 mm from the border of the lesion. The second region R_2 was a dilated contour that is 2 mm from the segmented lesion border. The third region R_3 was also a dilated contour, but unlike the contour that defines the region R_2, the contour in the region R_3 was 3 mm from the segmented lesion border.

The three normalized volumetric features were computed as

$$V_i = A_0 \cdot \sum_S (h_d(\mathbf{r}) - \overline{h}_{d,R_i}), \tag{4}$$

where $i \in [1, 2, 3]$, A_0 is the pixel size, $h_d(\mathbf{r})$ is dense tissue thickness at location $\mathbf{r}$ in the region S, and $\overline{h}_{d,R_i}$ is the average dense tissue thickness in the observed region.

2.C. Temporal features

We designed temporal features to characterize temporal change and to discriminate between benign lesions, which generally do not change much in the interval between screenings, and malignant lesions, which usually grow between two mammographic examinations. Each temporal feature was computed as relative difference

$$f_t = \frac{f_c - f_p}{f_p}, \tag{5}$$

where f_c and f_p are features from current and prior mammogram, respectively. We will refer to those features as temporal feature, current feature, and prior feature in the remainder of

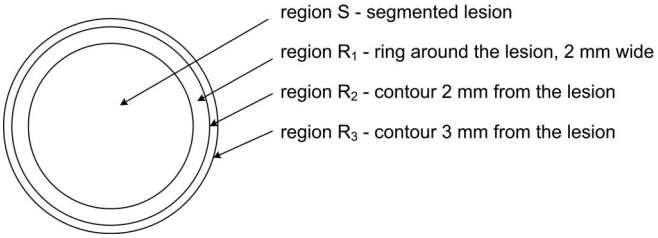


FIG. 1. Scheme of regions defined for normalizing volume.

this paper. All features were computed per lesion by averaging over views in which the lesion is visible, i.e., averaging over CC and MLO views.

2.D. Experiments

We analyzed the individual discriminative power of nine features computed from current and prior mammograms as well as their temporal change. Each feature was used to compute receiver operating characteristic (ROC) curve and the value of the area under the ROC curve (A_z value) using the ROCR package.[13]

We assessed the use of the temporal information compared to using only single exam information. We combined current and temporal features in different feature sets and analyzed the improvement in the classification performance of single exam features when used together with temporal features. Linear discriminant analysis with leave-one-out cross validation was used for the classification. Classification performance was analyzed using ROC curves and A_z values. The significance of differences between ROC curves was assessed by bootstrapping with 1000 stratified replications using the pROC package[14] and the criterion for significance was adjusted using the correction for multiple comparisons according to Bonferroni-Holm.[15]

3. RESULTS

3.A. Feature values

Figures 2–4 present boxplots for the prior, current, and temporal features normalized to zero mean and unit standard deviation. The bottom and top of the box are the first and third quartiles, and the band inside the box presents the median. The ends of the whiskers represent the lowest datum still within 1.5 interquartile range of the lower quartile and the highest datum still within 1.5 interquartile range of the upper quartile. All points that are outside the whisker are outliers.

The boxplots are depicted for benign and malignant lesions and show the difference between the two classes and the tendency of feature values to change in time.

From the boxplots with temporal features that present the change of a lesion during a time period it can be noticed that the median is higher for almost all temporal features derived for malignant lesions compared to those for benign lesions which indicates that malignant lesions change more in time. In most cases, benign lesions do not change. On average, malignant lesions are larger and have higher values of most features compared to the lesions one screening round earlier.

Figure 5 presents an example of a segmented malignant mass in a mammogram pair with the corresponding values of volume and area of the lesion. It can be noticed that area and volume of the lesion both increase between the two exams. When comparing the measurements in CC and MLO views, there is more agreement between the computed volumes than in the area measurements.

3.B. Individual feature performance

In order to evaluate individual performance of each temporal feature, current feature, and prior feature, we performed

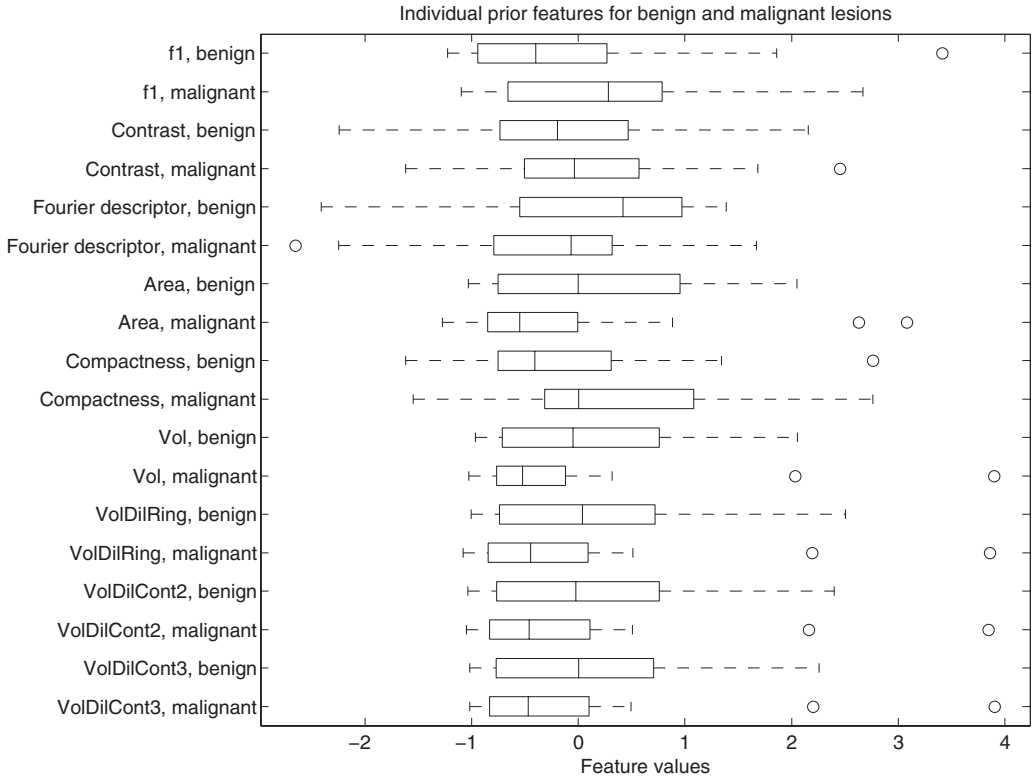


FIG. 2. Boxplot of normalized feature values for benign and malignant lesions computed from prior exams.

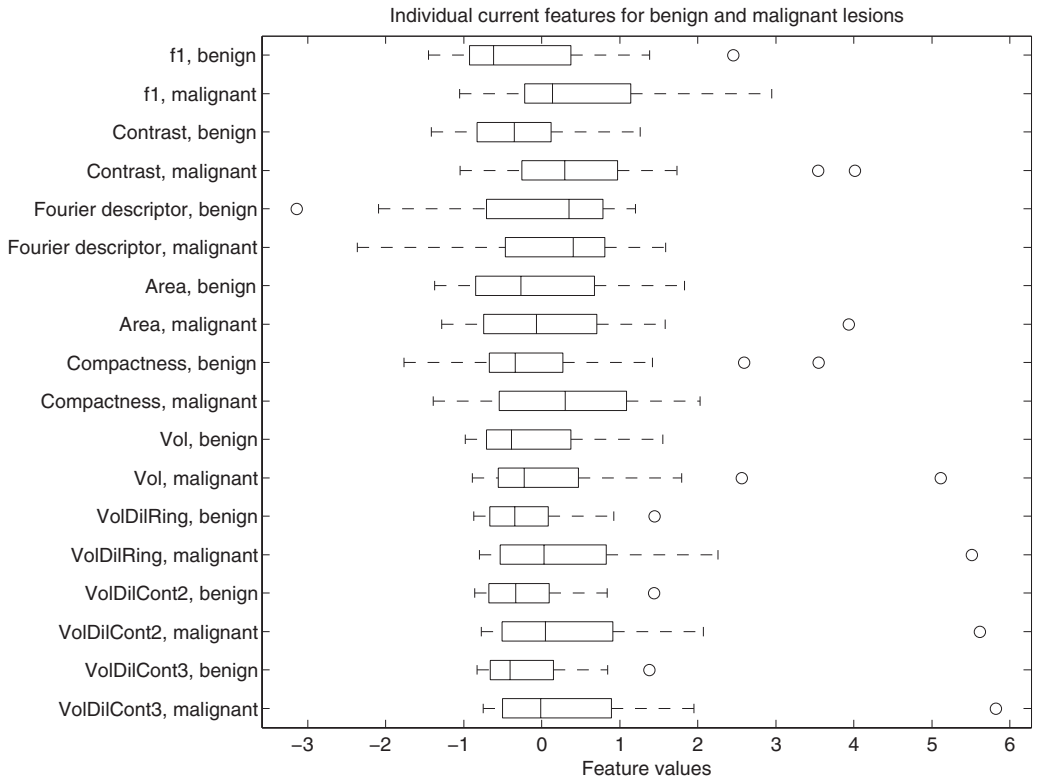


FIG. 3. Boxplot of normalized feature values for benign and malignant lesions computed from current exams.

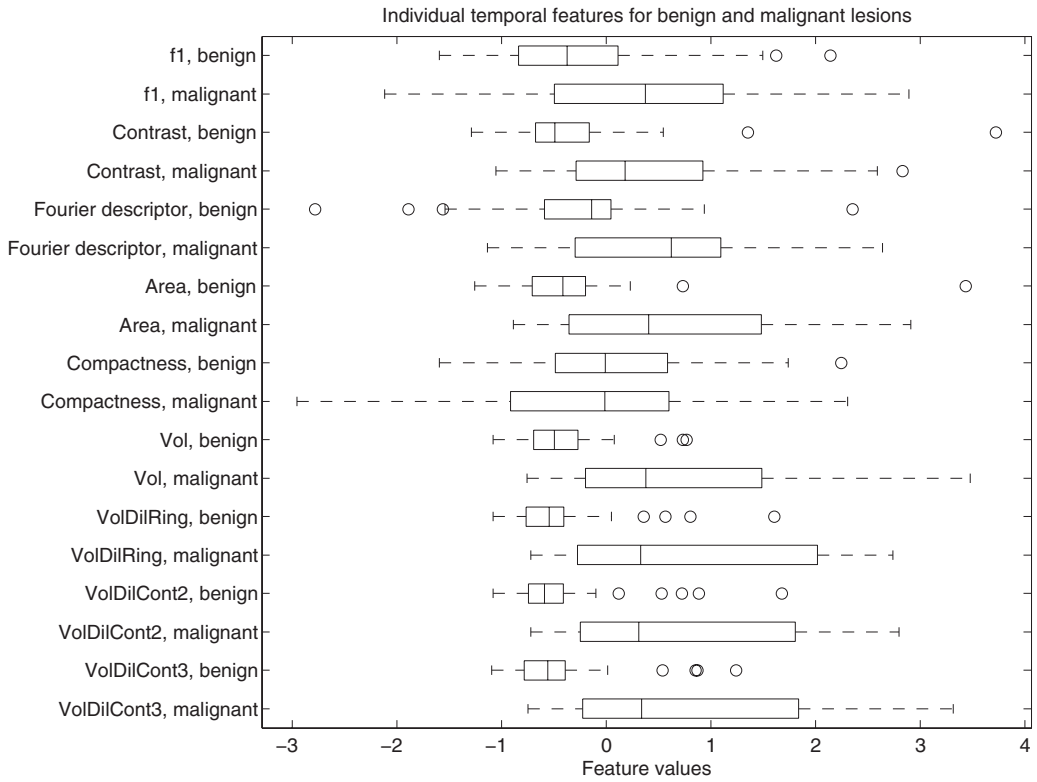
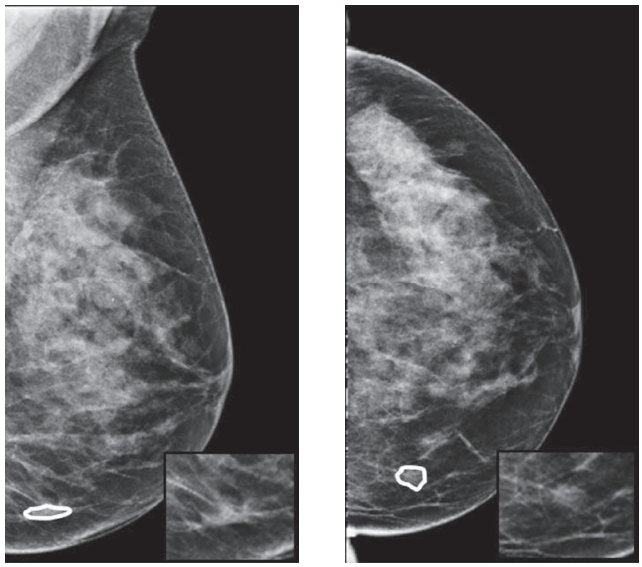


FIG. 4. Boxplot of normalized temporal feature values for benign and malignant lesions computed as difference between current and prior exams.



(a) A=0.34cm^2; V=0.26cm^3 (b) A = 0.77cm^2; V = 0.21cm^3

(c) A = 1.32cm^2; V = 1.18 cm^3 (d) A = 1.65cm^2; V = 1.20cm^3

FIG. 5. Example of segmented malignant lesion and corresponding values of area and volume in a temporal left mammogram pair and magnified lesion region in the lower right corner: (a) prior left mammogram in MLO view; (b) prior left mammogram in CC view; (c) current left mammogram in MLO views; (d) current left mammogram in CC view.

ROC analysis. Obtained A_z values with the corresponding 95% confidence intervals (CI) are presented in Table II.

The best individual performances were achieved by volumetric temporal features with their A_z values of above 0.84. The temporal change of feature VolDilRing, which is volume normalized with a dilated ring around the lesion, achieved performance of $A_z = 0.88$. The worst performance among temporal features was achieved by the compactness feature with A_z value of 0.62.

When observing features from the current mammogram the best performance was achieved by the contrast feature with an A_z value of 0.74. Very poor performance was obtained using Fourier descriptor and compactness, with A_z values of 0.55.

For the features from the prior mammogram, the best performance was obtained by using the compactness feature with A_z value of 0.69 and the worst performance was obtained by using the contrast feature with A_z value of 0.53.

TABLE II. Value of area under the ROC curve (A_z) for the individual temporal, current, and prior features with 95% CI. The bootstrapped CI is based on 1000 replications.

	A_z value (95% CI)		
Feature	Temporal	Current	Prior
$\overline{f1}$	0.66 (0.52–0.79)	0.73 (0.59–0.84)	0.67 (0.53–0.80)
Contrast	0.78 (0.65–0.90)	0.74 (0.61–0.86)	0.53 (0.39–0.70)
Fourier descriptor	0.74 (0.60–0.86)	0.55 (0.40–0.70)	0.63 (0.49–0.76)
Area	0.78 (0.64–0.89)	0.56 (0.42–0.70)	0.63 (0.48–0.76)
Compactness	0.62 (0.47–0.77)	0.55 (0.40–0.71)	0.69 (0.54–0.83)
Vol	0.84 (0.72–0.93)	0.57 (0.42–0.71)	0.63 (0.48–0.76)
VolDilRing	0.88 (0.77–0.95)	0.63 (0.48–0.77)	0.63 (0.49–0.77)
VolDilCont2	0.87 (0.77–0.94)	0.64 (0.50–0.78)	0.62 (0.47–0.75)
VolDilCont3	0.87 (0.77–0.95)	0.64 (0.50–0.78)	0.62 (0.48–0.76)

3.C. Combination of current and temporal features

Features from the current exam were combined with temporal features and the obtained results are presented in Table III. The combination of all nine features from the current exam (current feature set) achieved a performance of $A_z = 0.77$. When adding all nine temporal features to the current feature set the achieved classification performance increased to $A_z = 0.86$. Adding only one temporal feature, namely, temporal area change, to the current feature set increased the performance to $A_z = 0.85$. The best performance was achieved by combining all four temporal volumetric features with the current feature set as presented in Fig. 6. The obtained performance was $A_z = 0.90$.

After applying the Bonferroni-Holm correction for multiple comparisons, the results showed that the combination of all four temporal volumetric features with the current feature

TABLE III. Combination of current and temporal features and their performance expressed with A_z value and 95% CI. The bootstrapped CI is based on 1000 replications. Last column includes statistical difference expressed by p-value between the observed feature set and set with only current features.

Feature set	A_z value (95% CI)	P-value
All current features	0.77 (0.65–0.88)	...
All current features and all nine temporal features	0.86 (0.73–0.95)	0.141
All current features and temporal area feature	0.85 (0.74–0.94)	0.058
All current features and temporal area and contrast feature	0.85 (0.75–0.93)	0.140
All current features and four temporal volumetric features	0.90 (0.82–0.97)	0.005

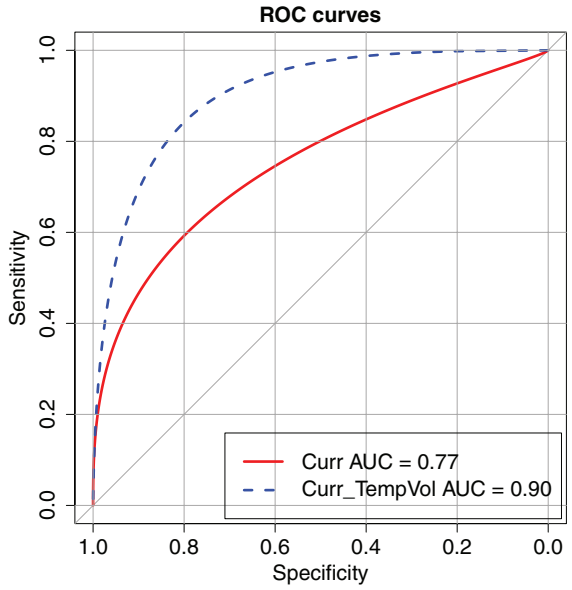


FIG. 6. Performance of the current feature set (Curr) and set with the combination of current feature set and four temporal volumetric features (Curr_TempVol).

set performed significantly better ($p = 0.005$) compared to the performance of single exam features.

4. DISCUSSION

As has been shown previously,[11] volume as a feature characterizes lesion size independently of mammographic projections and leads to more consistent size estimates than area based measures when comparing views. The correlation coefficient between area based lesion size in CC and MLO views was 0.70, while the lesion size computed from volume estimates lead to a correlation of 0.83. These results suggested that volume might be a more effective feature than area for the analysis of the temporal change of lesion size.

In this study, we compared the performance of lesion volume change to that of lesion area change in the task of classifying benign and malignant breast lesions. We hypothesized that temporal volumetric features are more efficient in discriminating benign and malignant lesions compared to temporal area change. Moreover, we hypothesized that inclusion of temporal volumetric features in a set of features obtained only from the current exam would improve classification performance. These hypotheses are supported by our results.

We computed volume and three normalized volumetric features using dense tissue thickness maps, which were computed from full field digital mammograms archived in raw format. By normalizing volume, we tried to eliminate the influence of dense tissue overlapping the lesions of interest. The fact that the normalized volume features had better individual performance than the uncorrected volume (Table II) indicates that normalization was helpful. The temporal change of normalized VolDilRing feature seemed to be the most relevant one, since volume feature is not expected to work well when there is overlapping dense tissue. The normalized VolDilCont2 and VolDilCont3 features are variations on VolDilRing but may be less stable because there are less pixels involved in computing the normalization.

Some prior features seem to be better (in the terms of larger A_z) compared to the current features (Table II). However, this result should be interpreted with caution, since there is a bias introduced by the selection of masses visible on priors: only masses that looked suspicious on the current mammograms are included, which is why the cases were recalled. Thus, malignant lesions are smaller than benign on the prior mammogram which might have influenced the results.

By adding the temporal volumetric features to features computed from the current exam, a significant increase in classification performance could be obtained. Thus, we may conclude that volumetric change of a lesion between two exams is a valuable feature for discriminating benign and malignant lesions in FFDM.

The limitation of this study is relatively small size of our sample. Whether there is an advantage of using volume change in the investigated task should be investigated with more data in a follow-up study. Such a study should also include lesions that were not visible on the prior mammogram, where a measure of volume change may be defined by assuming that at the previous screening the lesion was smaller than some threshold value and was therefore not visible.

In this study, we have shown that processing of dense tissue thickness maps as opposed to processing of the original intensity images is a viable approach. Results indicate that this approach may lead to superior results of computer aided diagnosis systems. An important advantage of the method we presented is that by converting mammograms to dense tissue thickness maps all subsequent analysis is made invariant to the x-ray machine and acquisition parameter settings used when the mammograms were acquired. This may make CAD algorithms more robust in practice.

ACKNOWLEDGMENTS

The authors are grateful to the Foundation of Population Screening Mid-West, The Netherlands, for giving access to the mammograms used in this study.

[a]Electronic mail: jelena.bozek@fer.hr

[1]S. Timp, C. Varela, and N. Karssemeijer, "Computer-aided diagnosis with temporal analysis to improve radiologists' interpretation of mammographic mass lesions," IEEE Trans. Inf. Technol. Biomed. **14**, 803–808 (2010).
[2]L. Hadjiiski, B. Sahiner, M. A. Helvie, H.-P. Chan, M. A. Roubidoux, C. Paramagul, C. Blane, N. Petrick, J. Bailey, K. Klein, M. Foster, S. K. Patterson, D. Adler, A. V. Nees, and J. Shen, "Breast masses: Computer-aided diagnosis with serial mammograms," Radiology **240**, 343–356 (2006).
[3]*American College of Radiology (ACR): ACR Breast Imaging Reporting and Data System, Breast Imaging Atlas*, 4th ed. (American College of Radiology, Reston, VA, 2003).
[4]E. S. de Paredes, *Atlas of Mammography*, 3rd ed. (Lippincott, Philadelphia, PA, 2007).
[5]L. Hadjiiski, B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. Gurcan, "Analysis of temporal changes of mammographic features: Computer-aided classification of malignant and benign breast masses," Med. Phys. **28**, 2309–2317 (2001).
[6]S. Timp, C. Varela, and N. Karssemeijer, "Temporal change analysis for characterization of mass lesions in mammography," IEEE Trans. Med. Imaging **26**, 945–953 (2007).

[7]S. Timp and N. Karssemeijer, "A new 2D segmentation method based on dynamic programming applied to computer aided detection in mammography," Med. Phys. **31**, 958–971 (2004).
[8]N. Karssemeijer and G. M. te Brake, "Detection of stellate distortions in mammograms," IEEE Trans. Med. Imaging **15**, 611–619 (1996).
[9]R. Hupse and N. Karssemeijer, "Use of normal tissue context in computer-aided detection of masses in mammograms," IEEE Trans. Med. Imaging **28**, 2033–2041 (2009).
[10]L. Shen, R. M. Rangayyan, and J. E. L. Desautels, "Application of shape analysis to mammographic calcifications," IEEE Trans. Med. Imaging **13**, 263–274 (1994).
[11]J. Bozek, M. Kallenberg, M. Grgic, and N. Karssemeijer, "Comparison of lesion size using area and volume in full field digital mammograms," in *Proceedings of the 11th International Workshop on Breast Imaging, IWDM 2012*, LNCS Vol. 7361, edited by A. D. A. Maidment, P. R. Bakic, and S. Gavenonis (Springer-Verlag, Philadelphia, PA, 2012), pp. 96–103.
[12]S. van Engeland, P. R. Snoeren, H. Huisman, C. Boetes, and N. Karssemeijer, "Volumetric breast density estimation from full-field digital mammograms," IEEE Trans. Med. Imaging **25**, 273–282 (2006).
[13]T. Sing, O. Sander, N. Beerenwinkel, and T. Lengauer, "ROCR: Visualizing classifier performance in R," Bioinformatics **21**, 3940–3941 (2005).
[14]X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, and M. Mueller, "pROC: An open-source package for R and S+ to analyze and compare ROC curves," BMC Bioinf. **12**, 77 (2011).
[15]S. Holm, "A simple sequential rejective multiple test procedure," Scand. J. Stat. **6**, 65–70 (1979).