Computerized Detection and Classification of Malignant and Benign Microcalcifications on Full Field Digital Mammograms

Lubomir Hadjiiski, Peter Filev, Heang-Ping Chan, Jun Ge, Berkman Sahiner, Mark A. Helvie, and Marilyn A. Roubidoux

University of Michigan, Department of Radiology, Ann Arbor, MI 48109
lhadjisk@umich.edu

Abstract. The purpose of the study is to develop an automated system for detecting microcalcifications within a predefined region of interest (ROI), and classifying the clusters as malignant and benign on full-filled digital mammograms (FFDM). Our system consists of two stages. In the first stage, a detection program is used to detect cluster candidates within the ROI. A rule-based identification method is designed to differentiate the true and false clusters. In the second stage, morphological and texture features are extracted from the selected clusters and a classifier is trained to classify malignant and benign clusters. In this study, a data set of 247 ROIs (63 malignant and 184 benign) containing biopsy-proven calcification clusters were used. An MQSA radiologist identified 117 corresponding clusters on the CC and MLO pairs of mammograms. Leave-one-case-out resampling was used for feature selection and classification. Two MQSA radiologists evaluated the two view pairs. The detection program correctly detected 100% (247/247) of the clusters of interest with 0.14 (35/247) FPs/ROI. The identification program correctly selected 99.2% (245/247) of the index clusters. In the classification stage an average of 4 features was selected from the training subsets. The most frequently selected features included 3 morphological and 1 texture features. The classifier achieved a test Az of 0.73 for classifying the 247 clusters as malignant or benign. For the 117 pairs of matched CC and MLO views the test Az was 0.77. The partial area index above a sensitivity of 0.9, $Az^{(0.9)}$, was 0.21. In comparison, the two experienced MQSA radiologists achieved Az of 0.76 and 0.73, respectively, for the 117 CC and MLO view pairs. The partial area index $Az^{(0.9)}$ was 0.27 and 0.12, respectively. Our classification system can detect the microcalcifications within the specified ROI on mammogram with high sensitivity and satisfactory specificity, and classify them with an accuracy comparable to that of an experienced radiologist.

Keywords: Full field digital mammograms, CAD, microcalcifications, classification.

1 Background

Characterization of microcalcification clusters as malignant and benign is a difficult task because of the large overlap between the features of malignant and benign

E.A. Krupinski (Ed.): IWDM 2008, LNCS 5116, pp. 336–342, 2008.

microcalcifications. We are developing computer-aided diagnosis (CAD) systems to assist radiologists in classification of breast lesions. In our previous system for analysis of microcalcifications, the individual microcalcifications were manually identified to reduce the chance that false calcifications were mixed with true microcalcification in training of the CAD system. There were previous studies investigating the effect of semi-automated detection of microcalcifications on the accuracy of classifying malignant and benign microcalcification [1],[2]. These studies were performed using screen-film mammograms. In this study, our goal was to develop an automated system for detecting individual microcalcifications and clusters on ROIs and classifying the clusters as malignant and benign on full-field digital mammograms (FFDM). The performance of the CAD system was compared to that of experienced radiologists for the same data set.

2 Methods

Our system consists of two stages. In the first stage, an automated program is used to detect cluster candidates within a predefined ROI (Fig. 1) on FFDM [3]. The detection program was tuned to perform well in a local region. The microcalcification detection process is summarized as follows. The algorithm involves three key steps: preprocessing/filtering, segmentation, and classification. In the preprocessing/filtering step two filters are applied to the extracted region inside the breast boundary. The first is a signal enhancement filter which enhances any potential microcalcification on the image. The second is a signal suppression filter whose main function is to smooth and remove noise from the image. The filtered images resulting from application of the two filters are subtracted, yielding a difference image. In the segmentation stage, candidate signal sites surpassing a global gray level threshold are identified. This threshold is updated iteratively until the potential number of sites identified is within an input range. Rule-based classification is used to exclude some large or high-contrast artifacts. A clustering criterion is applied to the signals eliminating isolated noise points. Lastly a trained convolution neural network is used to distinguish false positives from true microcalcifications.

The detected cluster candidates within the local region may include true positives and false positives. A cluster candidate is categorized as a true positive if it overlaps with the cluster identified by the radiologist. Any detected clusters that fail to achieve overlap are categorized as false positives.

A rule based identification method is designed to exclude some false clusters in the ROI based on the cluster size. In the second stage, morphological and texture features are extracted from the selected clusters and a classifier is trained to classify the clusters as malignant and benign.

Our computerized classification system uses raw FFDM mammograms as input. The individual microcalcifications are first segmented by the system from the mammographic background. Five morphological features describing the size, the density, and the shape of the individual microcalcifications are extracted. The mean and the variation of each of these features for the individual microcalcifications within a cluster are calculated. These cluster features in combination with the number of microcalcifications in a cluster form the morphological feature space.

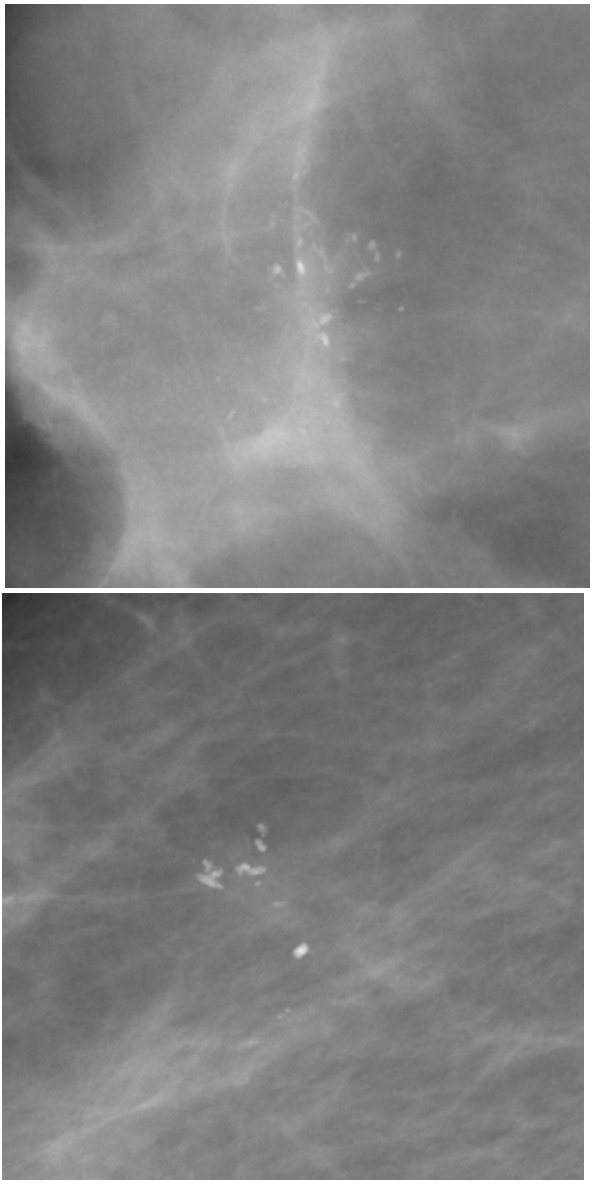

Fig. 1. Malignant (a) and benign (b) microcalcification clusters on FFDM. Visibility rating for both clusters is 3 on a 10-point scale (1: very obvious, 10: very subtle). The ROIs sizes are 35 mm ×35 mm.

For texture feature extraction, a region of interest (ROI) containing the cluster of microcalcifications is identified on the mammogram. Background correction is applied to the ROI to reduce the intensity variation in the breast tissue areas. Texture features including the mean, entropy, contrast, and angular second moment are

extracted from the gray level dependence difference statistics of the background-corrected ROIs in four directions [4].

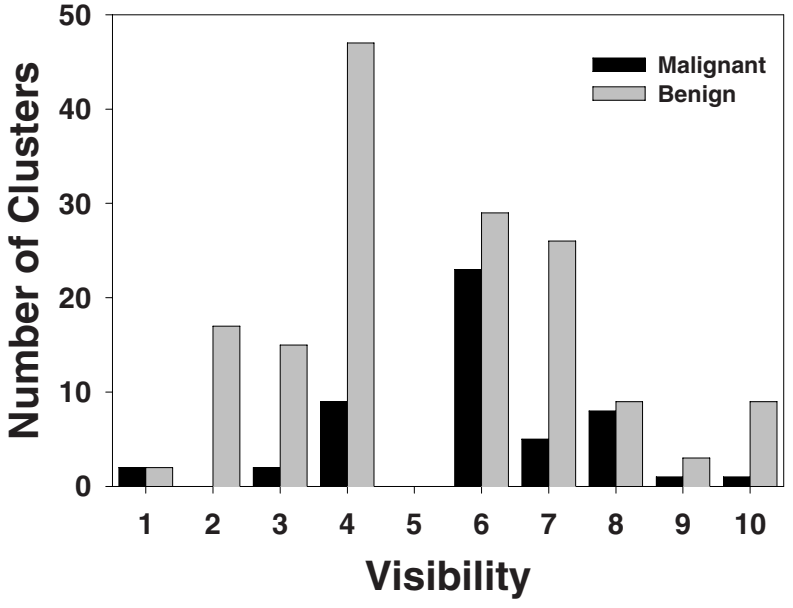

Fig. 2. The distribution of the visibility ranking of the microcalcification clusters in the data set of 117 CC and MLO views. Each view was ranked individually by an experienced radiologist (1: very obvious, 10: very subtle).

In this study, 247 full field digital mammograms from 111 patients (30 malignant and 81 benign cases) containing biopsy-proven calcification clusters were used. An MQSA radiologist identified the corresponding clusters on the different views. A 35mm × 35mm ROI containing the cluster was extracted from each mammogram resulting in 247 ROIs (63 malignant and 184 benign). From the 247 ROIs, 117 pairs of matched CC and MLO views (30 malignant and 87 benign) were identified. The distribution of the visibility ranking of the microcalcification clusters in the data set of 117 CC and MLO views is presented in Fig. 2. The ranking was performed by an experienced radiologist on a 10-point scale (1: very obvious, 10: very subtle). The average visibility ratings for the malignant clusters were 5.7±1.8 and 5.6±1.7 for benign clusters.

A leave-one-case-out resampling scheme is used to train and test the linear discriminant classifier. The most effective features from the combined morphological and texture feature space are identified using stepwise feature selection with simplex optimization in each training cycle. The performance of the classifier is evaluated by receiver operating characteristic (ROC) analysis.

Two experienced MQSA radiologists evaluated the 117 CC and MLO view pairs in an observer study. The CC- and MLO-view pairs of each microcalcification cluster

were presented to the radiologist side-by-side on the workstation. The FFDM images of $100\mu m \times 100\mu m$ pixel size were displayed. The radiologist evaluated the displayed two-view pairs and provided an assessment of the likelihood of malignancy (LM) on a 100-point scale (1=benign, 100=high likelihood of malignancy).

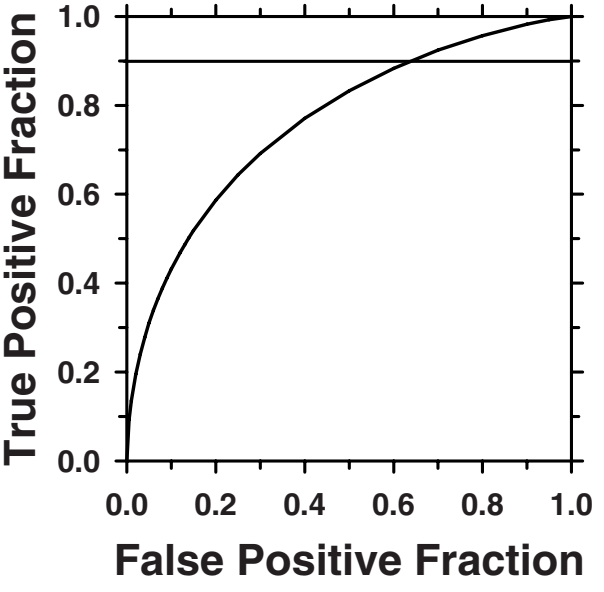

Fig. 3. The test ROC curve for the computer classifier score for the 117 CC and MLO pairs: $A_z = 0.77$, $Az^{(0.9)} = 0.21$

3 Results

The microcalcification detection program correctly detected 100% (247/247) of the index clusters within the ROIs with an average of 0.14 (35/247) FPs/ROI. The rule based identification program correctly selected 99.2% (245/247) of the index clusters. In two of the ROIs the identification program selected a different microcalcification cluster, which was in close proximity to the index cluster. Both ROIs contained benign clusters. In the classification stage an average of 4 features was selected from the training subsets. The most frequently selected features included 3 morphological features and 1 texture feature. The microcalcification classifier achieved a test Az of 0.73 for classifying the 247 clusters as malignant or benign. For the 117 two-view pairs the test Az was 0.77 (Fig. 3). The partial area index above a sensitivity of 0.9, $Az^{(0.9)}$, was 0.21. In comparison, the two experienced MQSA radiologists achieved an Az of 0.76 and 0.73, respectively, for the 117 two-view pairs (Fig. 4). The partial area index $Az^{(0.9)}$ was 0.27 and 0.12 respectively.

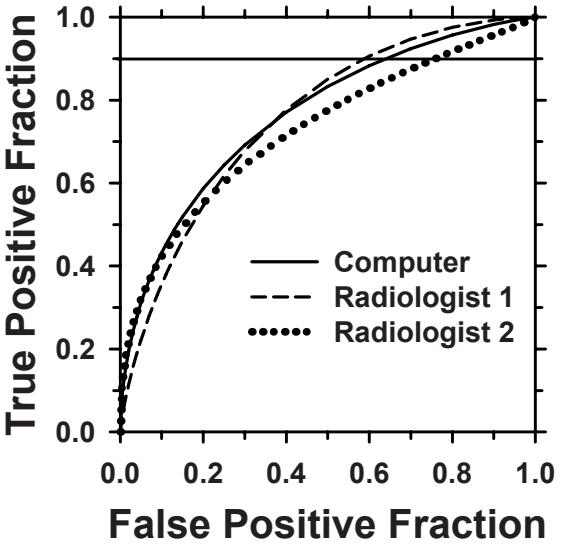

Fig. 4. The test ROC curve for the computer classifier score, Radiologist 1 and Radiologist 2 LM scores for the 117 CC and MLO pairs: computer classifier A_z = 0.77, $Az^{(0.9)}$ = 0.21; Radiologist1 A_z = 0.76, $Az^{(0.9)}$ = 0.27; Radiologist2 A_z = 0.73, $Az^{(0.9)}$ = 0.12

4 Discussion

Our detection/classification system can detect the clusters within the specified mammogram ROI with high sensitivity and satisfactory specificity. One of the two index clusters that the identification program could not select correctly was very subtle.Fig. 4 shows the comparison of the ROC curves for the computer classifier and the two experienced MQSA radiologists. The computer classifier performed comparable to that of the Radiologist 1 and it showed slight improvement when compared with Radiologist 2 for both the Az and the partial area index $Az^{(0.9)}$; however, the difference was not statistically significant ($p>0.4$). Any of the differences in Az obtained by the comparisons of the classifier and the Radiologists 1, classifier and the Radiologists 2 as well as the Radiologists1 and the Radiologists 2 did not reach statistical significance. The detection/classification system was able to classify the microcalcification clusters with an accuracy comparable to that of an experienced radiologist.

Acknowledgements

This work is supported by USAMRMC grant DAMD17-02-1-0214 and USPHS grant CA95153.

References

1. Jiang, Y., Nishikawa, R.M., Papaioannou, J.: Dependence of computer classification of clustered microcalcifications on the correct detection of microcalcifications. Medical Physics 28, 1949–1957 (2001)
2. Salfity, M.F., Nishikawa, R.M., Jiang, Y., Papaioannou, J.: The use of a priori information in the detection of mammographic microcalcifications to improve their classification. Medical Physics 30, 823–831 (2003)
3. Ge, J., Hadjiiski, L.M., Sahiner, B., Wei, J., Helvie, M.A., Zhou, C., Chan, H.P.: Computer-aided detection system for clustered microcalcifications: comparison of performance on full-field digital mammograms and digitized screen-film mammograms. Physics in Medicine And Biology (4), 981–1000 (2007)
4. Chan, H.P., Sahiner, B., Lam, K.L., Petrick, N., Helvie, M.A., Goodsitt, M.M., Adler, D.D.: Computerized analysis of mammographic microcalcifications in morphological and texture feature space. Medical Physics 25, 2007–2019 (1998)