

Matching mammographic regions in mediolateral oblique and cranio caudal views: A probabilistic approach

Maurice Samulski and Nico Karssemeijer

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

ABSTRACT

Most of the current CAD systems detect suspicious mass regions independently in single views. In this paper we present a method to match corresponding regions in mediolateral oblique (MLO) and craniocaudal (CC) mammographic views of the breast. For every possible combination of mass regions in the MLO view and CC view, a number of features are computed, such as the difference in distance of a region to the nipple, a texture similarity measure, the gray scale correlation and the likelihood of malignancy of both regions computed by single-view analysis. In previous research, Linear Discriminant Analysis was used to discriminate between correct and incorrect links. In this paper we investigate if the performance can be improved by employing a statistical method in which four classes are distinguished. These four classes are defined by the combinations of view (MLO/CC) and pathology (TP/FP) labels. We use distance-weighted k-Nearest Neighbor density estimation to estimate the likelihood of a region combination. Next, a correspondence score is calculated as the likelihood that the region combination is a TP-TP link. The method was tested on 412 cases with a malignant lesion visible in at least one of the views. In 82.4% of the cases a correct link could be established between the TP detections in both views. In future work, we will use the framework presented here to develop a context dependent region matching scheme, which takes the number and likelihood of possible alternatives into account. It is expected that more accurate determination of matching probabilities will lead to improved CAD performance.

Keywords: Methods: classification and classifier design, multiple view, region matching, Modalities: mammography, Diagnostic task: detection

1. INTRODUCTION

One of the major challenges in computer-aided detection (CAD) of mammographic masses is the reduction of false positives while sensitivity is maintained. Although many studies report that CAD systems improve the radiologists' accuracy, its effectiveness is not undisputed. In particular, CAD systems that give a high number of false positive markings results in radiologists not to have sufficient confidence in CAD results.¹

Screening usually consists of two-view mammography, i.e., a mediolateral oblique (MLO) and a cranio caudal (CC) film is obtained from both breasts. Using both views in screening improves the chance of detecting abnormalities, mainly due to additional information from the cranio caudal that allows the lesion to be seen more easily on the mediolateral oblique view. Furthermore, it reduces the number of false positives by offering an additional perspective in which superimposition of normal breast structures in the MLO view simulating a suspect lesion can be recognized as such.

One of the most important steps in multi view CAD techniques is to match corresponding regions in the available views. Few studies have been devoted to the investigation of methods for finding corresponding regions in different mammographic views. Paqueralt *et al.*² developed a two view approach, by calculating a correspondence score for each possible combination of segmented structures. Combining this correspondence score with the single view detection score resulted in a significant improvement of their detection results. Highnam *et al.*³

Copyright 2008 Society of Photo-Optical Instrumentation Engineers. This paper was (will be) published in SPIE Medical Imaging and is made available as an electronic reprint (preprint) with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

used a compression model to determine a curve in the mediolateral oblique mammogram which corresponds to potential positions of a point in the cranio caudal mammogram. Recently, Wu *et al.*⁴ developed a CAD system that incorporates information from two-view mammograms and bilateral mammograms where the corresponding region of interest on the contralateral mammogram is identified using a regional registration technique. Qian *et al.*⁵ designed an ipsilateral multiview CAD system where a region of interest in one view is matched with a region of interest in the other view based on their projection distance and analyzed for corresponding shape and characteristic features.

In a previous project, van Engeland *et al.*⁶ presented a method to match suspicious regions segmented by a single-view CAD system. For all possible region combinations in the MLO and CC view, a feature vector was calculated containing a number of features that describe the similarity between both regions and the likelihood of malignancy of both regions computed by single-view analysis, whereupon a correspondence measure was being determined using a Linear Discriminant Analysis (LDA) classifier. Finally, for every region in one view, the region in the other view with the highest correspondence score was selected as the corresponding candidate. Using the obtained correspondences in a multi-view CAD scheme resulted in an significant improvement of the lesion based detection performance.⁷ However, in the case based evaluation there was no improvement. The main reason is that links between a TP region in one view and a FP region in another view tend to lead to more suspicious ratings of the false positives and less suspicious ratings of the true positives, severely degrading the case based detection performance. Even a small number of TP-FP links can negate the improvement of the lesion based performance. Another reason that the improvement of the lesion based performance is not seen in the case based performance is that the lesions of which the mass likelihood was increased after applying the two-view CAD scheme, were already very suspicious in the other view.

In this work we aim to reduce the number of incorrect links and achieve an optimal balance between finding the correct TP-TP links and discarding the TP-FP links. The primary idea is to distinguish four classes rather than two and use a statistical approach. After introducing the region linking method, we will present some preliminary results and discuss future work.

2. MATERIALS AND METHODS

2.1 Dataset

The digitized mammograms that were used in this study have been obtained from the Dutch breast cancer screening program. In this study, 412 cases with both a MLO and CC view available were used, 41 prior and 371 diagnostic mammograms. A diagnostic mammogram is taken after a sign or symptom of breast cancer has been found, and a prior mammogram is the screening mammogram taken before breast cancer was diagnosed. In each case there was a malignant lesion visible in at least one of the views. Approximately one half (213 cases) was digitized with a Lumisys 85 digitizer, and the other half (199 cases) was digitized with a Canon CFS300 digitizer. All mammograms were digitized at a pixel resolution of 50 μm and averaged down to a resolution of 200 μm , maintaining a gray value depth of 12 bits.

We will introduce first our single view detection scheme which we will briefly explain in section 2.2. In section 2.2 we will present our region matching procedure.

2.2 Single view detection scheme

To each image in the dataset a CAD scheme was applied and consists of the following steps (Figure 1):

- Segmentation of the mammogram into breast tissue, pectoral muscle (if image is a MLO view), and background area.
- Initial detection step resulting in an image representing the likelihood of malignancy and a number of suspect image locations (local maxima in the likelihood image).
- Region segmentation, by dynamic programming, using the suspicious locations as seed points.
- Final classification step to improve the prediction of malignancy using region features.

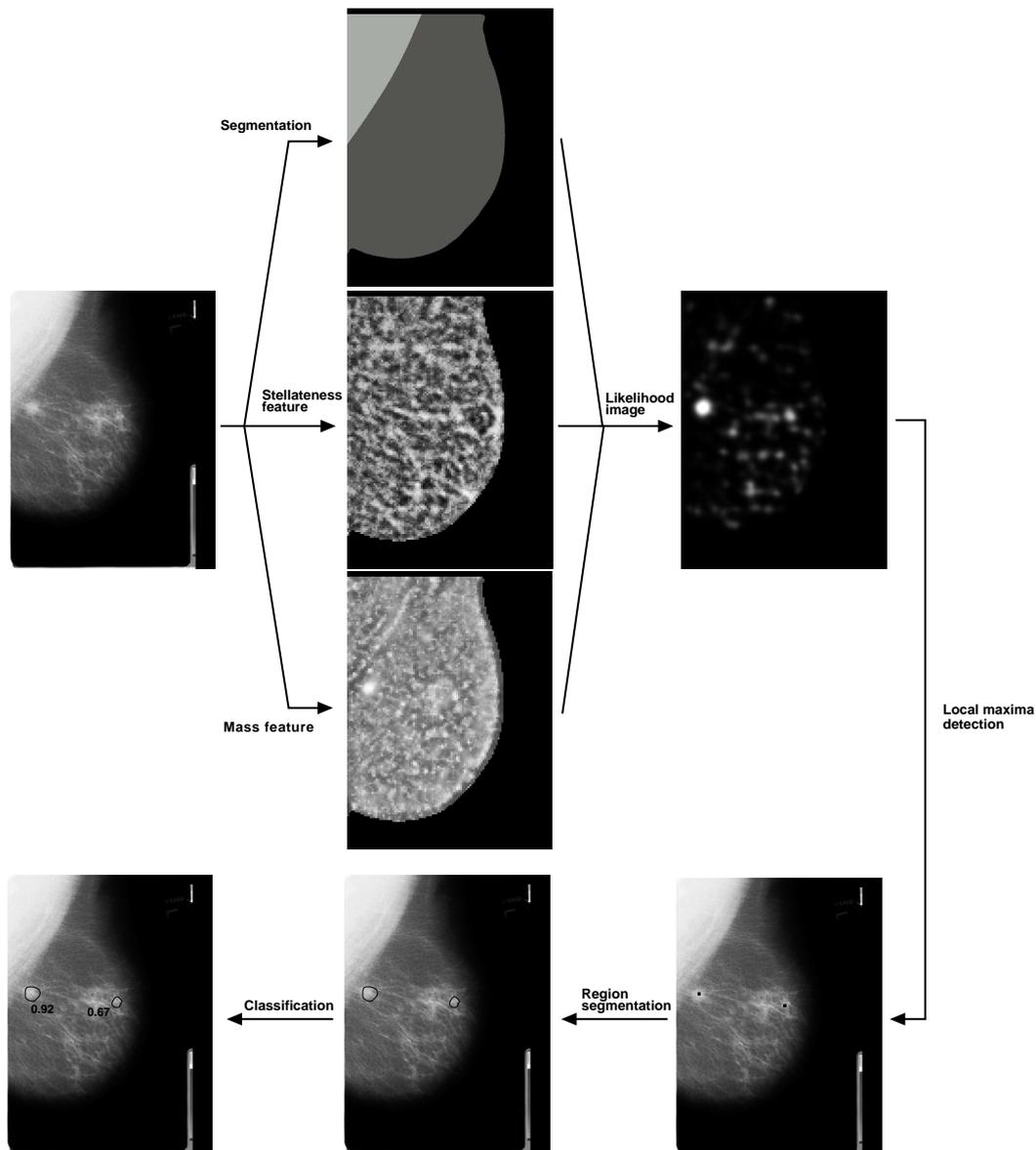


Figure 1. Schematic overview of the single view CAD scheme employed in this paper. First the mammogram is segmented into breast tissue, background tissue and the pectoral muscle. We then calculate at each location two stellateness features for the detection of spiculation and two mass features for the detection of a focal mass. A neural network classifier combines these features into a likelihood of a mass at that location, resulting in a likelihood image. The most suspicious locations on the likelihood image (bright spots) are selected and used as seed points for the region segmentation. After that, features are calculated for each segmented region. Finally a second classifier combines these features into a malignancy score that represents the likelihood that the region is malignant.

These steps will be described in more detail in the following paragraphs.

Segmentation of the mammogram The first step of our CAD scheme is the segmentation of an image into breast tissue and background, using a skin line detection algorithm. Additionally, it finds the edge of the pectoralis muscle if the image is a MLO view.⁸ After these steps, a thickness equalization algorithm is applied to enhance the periphery of the breast.⁹ A similar algorithm is used to equalize background intensity in the pectoralis muscle, to avoid problems with detection of masses located on or near the pectoral boundary.

Initial mass detection step In this step we use a multi-scale technique for the detection of stellate patterns, based on a statistical analysis of a map of the texture orientation in mammographic images through the use of operators sensitive to stellate patterns. The same algorithm is used to identify patterns of radial gradient vectors, rather than radial spiculations. For each pixel inside the breast area this results in a small number of features calculated that represent presence of a mass and the presence of spiculation.¹⁰ A neural network classifies each pixel using these features and assigns a level of suspiciousness to it. The neural network is trained using pixels sampled inside and outside of a representative series of malignant masses. The result is an image in which pixel values represent the likelihood that a malignant mass or architectural distortion is present. This likelihood image is then slightly smoothed and a local maxima detection is performed. A local maximum is detected when the likelihood is above a certain threshold and no other nearby locations have a higher likelihood value. This results in a number of suspicious locations. Finally an algorithm searches for local maxima that are located closer than 8 mm together and remove multiple candidate locations to avoid multiple suspicious locations on the same lesion.

Region segmentation Each of the detected local maxima in the previous step are used as seed points for region segmentation, based on dynamic programming.¹¹

Final classification For each segmented region, 81 features are calculated related to lesion size, roughness of the boundary, linear texture, location of the region, contour smoothness, contrast, and other image characteristics. A second neural network combines these features into a malignancy score that represents the likelihood that the region is malignant.

Multi view scheme For every region in one view an annular search area is defined in the other view based on the distance to the nipple, as this is a quite reliable landmark on a mammogram for correlating lesions in MLO and CC views. This reference point is also used by radiologists and remains fairly constant. To define the search area width, we used an annotated database containing 424 cases with a mass lesion that is visible in both the MLO and CC view. For varying width of the search area, the percentage of lesions in the corresponding view that is within this search area is shown in Figure 2. A corresponding lesion is within the search area if the difference in radial distance to the nipple is less than half the search area width. The nipple location was estimated by a simple procedure that assumes that the nipple is the point on the skin contour with the largest distance to the chest or pectoral muscle. If we set the search area width to 48 mm, almost all corresponding lesions are within this search area. For large breasts the search area is only a fraction of the breast area but for smaller breasts the search area covered almost the whole breast area. Figure 3 shows that there is a small correlation (correlation coefficient 0.17) between breast area and the absolute difference in radial distance between the nipple and the lesion in one view and the radial distance between the nipple and the corresponding lesion in the other view. Based on this, we set the width of the search area from 40 to 48 mm depending on the breast area in order to reduce the number of false positive candidate regions.

2.3 Correspondence features

The following paragraphs describe features we use that are invariant to compression and positioning and have high correlation between the values in the MLO and CC view.

2.3.1 Radial distance to the nipple

The radial distance between the lesion and the automatically estimated nipple position remains fairly constant between views (Pearson’s linear correlation coefficient was 0.89 for our annotated database containing 424 cases). The distance feature is defined as follows:

$$distance = \frac{|d_{MLO} - d_{CC}|}{w} \quad (1)$$

where d_{MLO} is the radial distance between the lesion and the nipple in the MLO view, d_{CC} is the radial distance between the lesion and the nipple in the CC view and w is the search area width.

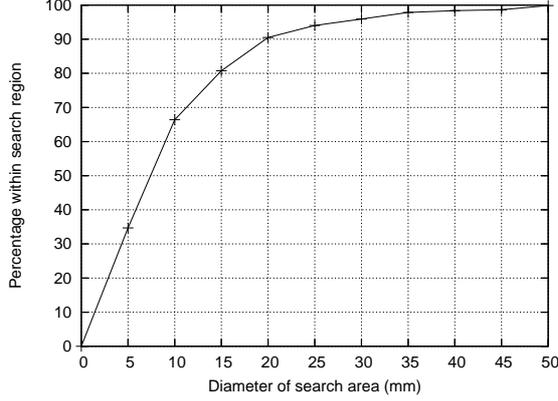


Figure 2. Corresponding lesions within the annular search region for varying search area size.

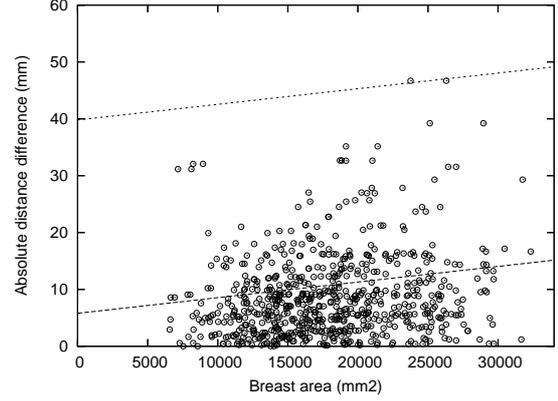


Figure 3. Correlation between breast area and the absolute difference in distance between the nipple and the lesion in one view and the distance between the nipple and the corresponding lesion in the other view. Based on this we vary the width of the search area between 40 to 48 mm.

2.3.2 Gray scale correlation

We compute a gray scale correlation feature used by Timp *et al.*¹² and Sanjay-Gopal *et al.*¹³ for the registration of lesions in temporal mammograms. It is based on the pixel correlation between a region of interest in the current mammogram and a candidate region in the prior mammogram. This region of interest consists all the pixels inside the contour of the region which represents the underlying mass lesion and a band of pixels around the contour which represents the surrounding tissue. This is illustrated in Figure 4.

The problem of finding corresponding regions in MLO and CC views is different, since the breast is an elastically deformable soft-tissue structure that is compressed to different extents and in different directions for the two views. To compensate for these effects a modified measure is deduced. First a polar coordinate transformation is applied to the regions using the center of mass of the regions as the center.

Then Pearson's correlation measure in polar space is calculated, allowing also a rotation ϕ of the CC region with respect to the MLO region. The maximum gray scale correlation over all angles is then used as a feature:

$$gray_scale_corr = \frac{\sum_{x,y} (g_{mlo}(x,y) - \overline{g_{mlo}})(g_{cc}(x,y) - \overline{g_{cc}})}{\sqrt{(\sum_{x,y} (g_{mlo}(x,y) - \overline{g_{mlo}})^2)(\sum_{x,y} (g_{cc}(x,y) - \overline{g_{cc}})^2)}} \quad (2)$$

$$polar_corr = \max(gray_scale_corr(\phi)). \quad (3)$$

2.3.3 Entropy

The Shannon's entropy is a measure of the average information carried in a pattern, which is widely used to quantify the smoothness of image texture. Tourassi *et al.*¹⁴ used it for content based retrieval and detection of masses in screening mammograms. We use the observation that Shannon's entropy will be relatively low in homogeneous patterns and increases in inhomogeneous regions to match corresponding regions. We use the absolute difference between the entropy of the polar representation of the region of interest and the candidate region in the other view as a feature.

2.3.4 Histogram correlation

The histogram correlation⁶ between the region in the MLO view and the CC view is determined by using two templates. The first template contains pixels inside the region and the second template contains pixels in the band outside the contour as shown in Figure 4. A lookup table of gray level values is constructed from the cumulative distribution functions of both the MLO region and candidate region. This lookup table is then used

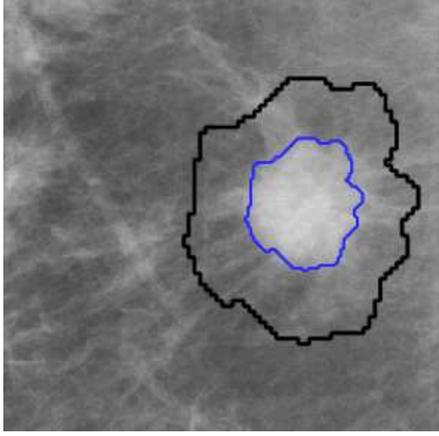


Figure 4. Illustration of the inner and outer template used for the computation of the gray scale correlation and histogram correlation features.

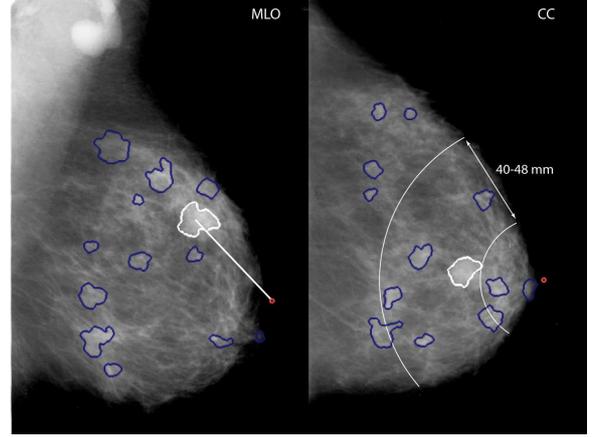


Figure 5. Schematic of the search area in the CC view, based on the lesion to nipple distance in the MLO view. The light dot indicates the estimated nipple location.

to approximately map the pixel values of the MLO region to pixel values of the candidate region to correct to some degree for differences in exposure. After the application of the lookup table, the gray value histograms for both regions are obtained. The histogram correlation is calculated as follows:

$$HC = 1 - \frac{1}{2} \sum_g \left| \frac{H_{MLO}[g]}{T_{MLO}} - \frac{H_{CC}[g]}{T_{CC}} \right| \quad (4)$$

where g is the gray level value, H_{MLO} is the gray level histogram of the region in the MLO view, H_{CC} is the histogram of the candidate region in the CC view, and T_{MLO} and T_{CC} are the total number of counts in the MLO, respectively, the CC histogram. The histogram correlations of the inner and outer template are combined into one feature as follows:

$$histogram_correlation = \frac{HC_{MLO} + HC_{CC}}{2}. \quad (5)$$

2.3.5 Mass likelihood

The output of the second neural network classifier in our single-view CAD scheme represents the mass likelihood of a region. From this mass likelihood we derive three features: the mass likelihood of the region in the MLO view, the mass likelihood of the candidate region in the CC view and the absolute difference between the mass likelihood of both views.

2.3.6 Compactness difference

Compactness represents the roughness of an objects boundary relative to its area. This feature is included in the single-view CAD scheme because benign masses often have a round or oval shape compared to a more irregular shape of malignant masses. We use the difference between the compactness of the region in the MLO view and the CC view as correlation feature.

2.3.7 Linear texture difference

The linear texture feature that originates from the single-view CAD system represents presence of linear structures inside the segmented region, as normal breast tissue often has different texture characteristics than tumor tissue. Again we use the difference of this feature between the region in the MLO and CC view as correlation feature.

2.4 Classification of region combinations

We can distinguish four classes of region combinations. The first class contains links between TP regions in both views. The second class and third class represent TP-FP combinations and FP-TP combinations, in the MLO and CC views, respectively. The fourth class deals with FP-FP combinations, which primarily include links between normal breast structures for which no ground truth is available. This approach is markedly different from previous research⁶ of our group, where a LDA classifier was trained only on TP-TP region pairs (correct combinations) and TP-FP region pairs (incorrect combinations). The resulting LDA classifier output, referred to as correspondence score, was used to select the most likely region combination. The region was linked if the correspondence score exceeded a fixed threshold. When linked, a number of features that describe the resemblance between the best corresponding regions and their likelihood of malignancy was used to train a new two-view classifier. Instead, we propose a statistical approach by using distance-weighted k-nearest-neighbor density estimation and use the matching probabilities in the two-view classifier. We want to differentiate between the TP-FP class and FP-TP class, considering that linking a false positive lesion in one view to a true positive lesion in the other view has a minor effect on the case based performance as this slightly increases the likelihood of malignancy of the false positive lesion. This is relatively a small fraction of the false positives since most mammograms in screening are normal. However, if a true positive lesion is linked to a false positive lesion, the likelihood of malignancy of the true positive lesion decreases, resulting in a decrease in sensitivity of the CAD system.

2.5 Distance-weighted k-nearest-neighbor

The k-Nearest Neighbor (kNN) technique is a straightforward yet effective method for density estimation and has a long history in the pattern classification field. A well-known estimate of posterior probabilities can be determined as follows:¹⁵ we use a fixed value of K neighbors, and use the training data to find an appropriate value for volume V . To do this, a small sphere is centered on a new sample \vec{x} at which we want to estimate the unconditional density $p(\vec{x})$, and grow the sphere until it contains exactly K neighbors, irrespective of their class. The posterior probability of class C given a point \vec{x} is given by

$$p(C|\vec{x}) = \frac{p(\vec{x}|C)p(C)}{p(\vec{x})} = \frac{\frac{N_C}{N} \frac{K_C}{N_C V}}{\frac{K}{NV}} = \frac{K_C}{K}$$

where K is the number of neighbors, K_C is the number of neighbors from class C , V is the volume of the sphere, N_C is the number of points in the training dataset from class C , and N is the total points in the training set. We selected K by employing 10-fold cross-validation on the training set.

This estimation procedure treats each neighbor equally. We use a variant of the typical kNN algorithm proposed by Dudani,¹⁶ the distance-weighted k-nearest neighbor method, suggesting that training samples closest to the test sample should be given greater weight than more distant training samples. This is especially more true in posterior probability estimation than in classification, because a larger K is needed for posterior probability estimation as it lessens the effect of discretization. We have defined the neighbor's weight to be the inverse of its distance to the query point.

3. EVALUATION

To evaluate the performance of our statistical framework, we will compare it to the region linking method used in a previous study⁶ by calculating a correspondence score. The correspondence score is defined as the linear combination:

$$\text{Correspondence score} = L_{TPTP} - \alpha L_{TPFP}$$

where L_{TPTP} is the likelihood that the region combination is a TP-TP, L_{TPFP} is the likelihood that the region combination is a TP-FP, and α is a parameter that can be used to tune between the number of TP-TP links that will be established and the number of TP-FP links. When a threshold is put on whether to link two regions or not, increasing α will reduce the number of TP-FP links, but increase the number of missed TP-TP links. We will set α to 0, for the comparison with the linking method from our previous study, where the correspondence score was the output of the LDA classifier. The LDA classifier is trained using only the region combinations,

Table 1. For every view containing a lesion we tested whether this lesion was correctly linked with the lesion in the corresponding view (TP-TP) or linked to a false positive region (TP-FP) using the LDA classifier from previous research and the kNN classifier used in this study.

Method	TP-TP	TP-FP
2-class LDA classifier	694 (80.8%)	164 (19.2%)
4-class kNN classifier	707 (82.4%)	151 (17.6%)

where the region in the MLO and/or CC view is a lesion. The kNN classifier is trained on FP-FP combinations as well.

The classifiers are tested using 2-fold cross-validation, resulting in a correspondence score for every region combination:

$$\begin{bmatrix} C_{0,0} & C_{0,1} & \cdots & C_{0,CC} \\ C_{1,0} & C_{1,1} & & C_{1,CC} \\ \vdots & & \ddots & \\ C_{MLO,0} & C_{MLO,1} & & C_{MLO,CC} \end{bmatrix}$$

Based on this correspondence score matrix, the best corresponding link is selected, that is, the one with the highest correspondence score. For each view we only take, at most, eight candidate regions into account that have the highest mass likelihood. As evaluation measure we count the TP detections that were correctly linked to the TP detection in the other view, and the number of TP regions linked with a FP region in the other view. FP-FP combinations were not taking into account.

4. RESULTS

Results in Table 1 show the number of correct TP-TP links and the number of incorrect TP-FP links using the LDA classifier from previous research⁷ and the kNN classifier used in this study. Some improvement was noted in classifying corresponding links but did not reach statistical significance using the McNemar’s chi-square test ($p \leq 0.13$).

Additionally, we can apply a threshold to the correspondence score, such that we only establish a link when the score of the best corresponding link exceeds a certain value. The reduction of TP-FP links is important as previous research⁷ showed that links between TP and FP regions severely degrade detection performance. When using a fixed threshold of 0.5, the number of TP-FP combinations decreased from 151 to 107, at the cost of TP-TP combinations found which decreased from 707 to 643 for the kNN classifier. Using the same threshold on the correspondence score of the LDA classifier results in a decrease of TP-FP combinations from 164 to 152 at a cost of 14 unlinked TP-TP combinations. This is shown in Table 2. If the threshold is chosen such that the percentage of correct TP-TP combinations is 70%, the number of TP-FP combinations decreases with 35 in favor of the kNN classifier. This decrease is significant using the two tailed Fisher’s exact test ($p \leq 0.0008$).

Changing the α to 0.3 in the correspondence score formula (Eq. 3) for the kNN classifier while maintaining a fixed threshold of 0.5 on the correspondence score, the number of TP-TP combinations decreases from 643 to 587, and the number of TP-FP combinations decreases from 107 to 88.

5. CONCLUSIONS AND FUTURE WORK

We have developed a two view region matching method to link mammographic regions in MLO and CC views. This method uses the distance-weighted kNN technique to discriminate region links into the four possible categories: TP-TP, TP-FP, FP-TP, and FP-FP. The correspondence score was set to the likelihood that the region combination is a TP-TP link. For every view that contained a lesion we tested whether this lesion was correctly linked with the lesion in the corresponding view. For 82.4% of the TP regions, a correct link could be established, as is shown in Table 1. The maximum performance that could have been achieved was 92.4%, because not all cases contained an annotated TP region in both views. Additionally, we can apply a threshold to the correspondence score, such that we only establish a link when the score of the best corresponding link exceeds a

Table 2. For every view containing a lesion we linked the lesion with the best corresponding lesion in the other view, only if the correspondence score is above the threshold. Then we tested whether the lesion was correctly linked with the lesion in the corresponding view (TP-TP) or linked to a false positive region (TP-FP). The region combinations with a correspondence score less than the threshold are shown in the third column. The first two rows of the table show the performance when using a fixed threshold of 0.5. The next two rows show the results when the threshold is chosen such that we achieve 70% correct TP-TP combinations.

Method	Linked TP-TP	Linked TP-FP	Unlinked
2-class LDA classifier with threshold 0.5	680	152	26
4-class kNN classifier with threshold 0.5	643	107	108
2-class LDA classifier with threshold, TP-TP = 70%	604	89	165
4-class kNN classifier with threshold, TP-TP = 70%	604	54	200

certain value. This reduces the amount of TP-FP combinations considerably, especially in 7.6% of the TP regions where there was no annotated TP region in the other view. Changing the α in the correspondence score formula can be used to further reduce the number of false TP-FP combinations, but at the cost of TP-TP combinations found. Previous research⁷ showed that links between TP and FP regions severely degrade detection performance. We expect that the decrease in TP-TP combinations has a less negative effect on the detection performance of the two-view classifier whereas the regions are independently analyzed by the single-view CAD system. If the threshold is chosen such that the percentage of correct TP-TP combinations is 70%, the number of TP-FP combinations decreases significantly when using the 4-class kNN classifier (Fisher’s exact test, $p \leq 0.0008$).

Several reasons could be given as causes of incorrect links. The most important cause is the occurrence of two regions with the same distance to the nipple and a near similar feature vector. Also incorrect segmentations such as multiple overlapping regions in the same lesion in the first stages of our CAD scheme is a common cause for incorrect links. The linking performance without using a threshold is better using the 4-class kNN classifier in comparison to the 2-class LDA classifier, but the improvement did not reach statistical significance using the McNemar’s chi-square test ($p \leq 0.13$). However, the real performance gain is expected in the next stage of our CAD scheme where we will combine the information from our single view CAD system with correspondence information and the four class probabilities from our linking method.

Another application of this method we recently implemented is linking CAD regions on a mammographic workstation where a radiologist points at a region of interest in one view and the CAD system presents the corresponding region in the other view. The effect of presenting CAD results in both views to the radiologists needs to be investigated further. Future work includes developing a context dependent region matching scheme, which will also take the number and likelihood of possible alternatives into account to improve the case based sensitivity of our single view detection system. To reduce the number of incorrect links that are caused by similar regions with almost the same distance to the nipple, we will investigate additional features to improve linking.

ACKNOWLEDGMENTS

This work has been funded by the Netherlands Organization for Scientific Research (NWO) as a “hefboom” grant, Pr. Nr. 641.000.403.

REFERENCES

1. K. Moberg, N. Bjurstam, B. Wilczek, L. Rostgard, E. Egge, and C. Muren, “Computed assisted detection of interval breast cancers,” *European Journal of Radiology* **39**(2), pp. 104–10, 2001.
2. S. Paquerault, N. Petrick, H. Chan, B. Sahiner, and M. Helvie, “Improvement of computerized mass detection on mammograms: Fusion of two-view information,” *Medical Physics* **29**(2), pp. 238–247, 2002.
3. R. Highnam, Y. Kita, M. Brady, B. Shepstone, and R. English, “Determining correspondence between views,” in *4th International Workshop on Digital Mammography*, N. Karssemeijer, M. Thijssen, J. Hendriks, and L. van Erning, eds., pp. 111–118, 1998.

4. Y. Wu, J. Wei, L. Hadjiiski, B. Sahiner, C. Zhou, J. Ge, J. Shi, Y. Zhang, and H. Chan, "Bilateral analysis based false positive reduction for computer-aided mass detection," *Medical Physics* **34**(8), pp. 3334–3344, 2007.
5. W. Qian, D. Song, M. Lei, R. Sankar, and E. Eikman, "Computer-aided mass detection based on ipsilateral multiview mammograms.," *Academic Radiology* **14**(5), pp. 530–538, 2007.
6. S. van Engeland, S. Timp, and N. Karssemeijer, "Finding corresponding regions of interest in mediolateral oblique and craniocaudal mammographic views," *Medical Physics* **33**, pp. 3203–3212, September 2006.
7. S. van Engeland and N. Karssemeijer, "Combining two mammographic projections in a computer aided mass detection method.," *Medical Physics* **34**(3), pp. 898–905, 2007.
8. N. Karssemeijer, "Automated classification of parenchymal patterns in mammograms," *Physics in Medicine and Biology* **43**(2), pp. 365–378, 1998.
9. P. Snoeren and N. Karssemeijer, "Thickness correction of mammographic images by anisotropic filtering and interpolation of dense tissue," in *Medical Imaging 2005: Image Processing. Edited by Fitzpatrick, J. Michael; Reinhardt, Joseph M. Proceedings of the SPIE, Volume 5747, pp. 1521-1527 (2005).*, J. M. Fitzpatrick and J. M. Reinhardt, eds., pp. 1521–1527, April 2005.
10. N. Karssemeijer and G. te Brake, "Detection of stellate distortions in mammograms," *IEEE Transactions on Medical Imaging* **15**(5), pp. 611–619, 1996.
11. S. Timp and N. Karssemeijer, "A new 2d segmentation method based on dynamic programming applied to computer aided detection in mammography," *Medical Physics* **31**(5), pp. 958–971, 2004.
12. S. Timp, S. van Engeland, and N. Karssemeijer, "A regional registration method to find corresponding mass lesions in temporal mammogram pairs," *Medical Physics* **32**, pp. 2629–2638, Augustus 2005.
13. S. Sanjay-Gopal, H. P. Chan, T. Wilson, M. Helvie, N. Petrick, and B. Sahiner, "A regional registration technique for automated interval change analysis of breast lesions on mammograms.," *Medical Physics* **26**(12), pp. 2669–2679, 1999.
14. G. Tourassi, B. Harrawood, S. Singh, J. Lo, and C. Floyd, "Evaluation of information-theoretic similarity measures for content-based retrieval and detection of masses in mammograms.," *Medical Physics* **34**(1), pp. 140–50, 2007.
15. C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.
16. S. Dudani, "The distance-weighted k-nearest-neighbor rule," *IEEE Transactions on Systems, Man, and Cybernetics* **6**, pp. 325–327, April 1976.