Computer Vision and Learning Approaches to Medical Imaging

Dave Tahmoush and Hanan Samet

Abstract

A technique that radiologists use to diagnose breast cancer involves first finding suspicious sites in the mammograms and then comparing the left and right breasts to reduce the number of false positives. The symmetry of the human body is utilized to increase the accuracy of the diagnosis through visual registration of the mammograms. We emulate this technique by combining both computer vision and learning approaches, thus capturing the diagnosis of the radiologist. First, the suspicious sites are found and ranked using a filtering method, and then the top candidates are sorted into a point set. Second, the point sets of the left and right breast mammograms are compared by learning a spatial symmetry method that utilizes clustering to bypass direct registration. Clustered comparisons that are space-defined are found to compare favorably against clustered comparisons that are data-defined in this application. The space-defined clustered comparison process determines the presence of cancer 97% of the time, outperforming commercial systems, thus making it a strong classifier which should significantly improve computer-aided breast cancer detection systems.

1 Introduction

Contextual and spatial comparisons can be combined to determine image similarity, which has been often utilized for content-based image retrieval (CBIR) from image databases [12, 7, 10]. Medical image databases have also used image similarity, from rule-based systems for chest radiographs [23] to anatomical structure matching for 3D MR images [13] to learning techniques [11]. However, the focus is often on the non-cancerous structures, while it is the cancerous structures that are of principle interest. In this paper we apply image similarity concepts to the problem of detecting breast cancer in mammograms.

Breast cancer remains a leading cause of cancer deaths among women in many parts of the world. In the United States alone, over forty thousand women die of the disease each year [1]. Mammography is currently the most effective method for early detection of breast cancer [19]. For two-thirds of the women whose initial diagnosis of their mammogram is negative but who actually have breast cancer, the cancer is evident upon a second diagnosis of their mammogram [19]. Computer-aided detection (CAD) of mammograms could be used to avoid these missed diagnoses, and has been shown to increase the number of cancers detected by more than nineteen percent [9]. Measuring asymmetry, which consists of a comparison of the left and right breast images [8], is a technique that could be used to improve the accuracy of CAD. An automated prescreening system only classifies a mammogram as either normal or suspicious, while CAD picks out specific points as cancerous [4]. One of the most challenging problems with prescreening is the lack of sensitive algorithms for the detection of asymmetry [3].

The majority of work on CAD analysis of mammograms has focused on determining the contextual similarity to cancer, finding abnormalities in a local area of a single image [14, 21]. The primary methods used range from filters to wavelets to learning methods, but a detailed discussion of various imaging techniques is beyond the scope of this paper. Problems arise in using filter methods [14] because of the range of sizes and morphologies for breast cancer, as well as the difficulty in differentiating cancerous from non-cancerous structures. The size range problem has been addressed by using multi-scale models [21]. Similar issues affect wavelet methods, although their use has led to reported good results [17] with the size range issue being improved through the use of a wavelet pyramid [18]. Learning techniques have included support vector machines [5] and neural networks [17].

Detecting breast cancer in mammograms is challenging because the cancerous structures have many features in common with normal breast tissue. This means that a high number of false positives or false negatives are possible. Asymmetry can be used to help reduce the number of false positives so that true positives are more obvious. Previous work utilizing asymmetry has used wavelets or structural clues to detect asymmetry with correct results as often as 77% of the time [8, 20]. Additional work has focused on bilateral or temporal subtraction, which is the attempt to subtract one breast image from the other [25, 28]. This approach is hampered by the necessity of exact registration and natural asymmetry of the breasts. Bilateral subtraction is good because it does try to utilize the multiple images taken with the same machine by the same technician and analyzed using the same process in an effort to reduce the systematic differences that can be introduced. We believe that developing ways to better utilize asymmetry is consistent with a philosophy of trying to use methods that

Figure 1: Mammograms of left and right breasts with cancerous area outlined. The similarity of texture between cancerous and normal tissue makes asymmetry an important tool in cancer detection.

can capture measures deemed important by doctors thereby building upon their knowledge base, instead of trying to supplant it. However, measuring asymmetry involves registration and comparing multiple images, and thus it is a more complicated process.

Registration is the matching of points, pixels, or structures in one image to another image. Registration of mammograms is difficult because mammograms are projections of compressed three-dimensional structures. Primary sources of misregistration are differences in positioning and compression, which manifests itself in visually different images. The problem is more complex because the breast is elastic and subject to compression. Additional sources of difficulties include lack of clearly defined landmarks and normal variations between breasts. Strictly speaking, precise mammogram registration is intractable. However, an approximate solution is possible [24]. Warping techniques have been used [22], as well as statistical models [27] or mutual information as a basis for registration [26]. In this paper we learn image comparison models based upon clustering that encapsulate an approximate registration and use them to compare the mammograms of the left and right breasts.

Comparing multiple mammograms using learning techniques has been shown to be effective in CBIR [7]. Neural nets can be an effective supervised learning technique [2]. Our application lends itself well to supervised learning because the data set has already been screened for cancer and thus classified by expert radiologists. However, care must be taken since the expert classification is known not to be perfect [19].

The rest of this paper is organized as follows. Section 2

presents our method for measuring similarity between the mammogram images. Section 3 discusses the evaluation of the performance of the measure. Section 4 compares the results with other work, while Section 5 discusses the conclusions that can be drawn from this work.

2 Measuring Similarity

Our work utilizes filtering followed by spatial symmetry analysis to determine an overall measure of similarity by combining the contextual similarity of the filtering with the spatial similarity of the analysis. This can be a useful measure for prescreening mammograms since only an overall determination is required. We believe that many of the techniques described here can also be adapted for use in CAD analysis. A secondary goal of our work is to determine the importance of similarity or asymmetry in the computer analysis of mammograms. Figure 1 shows why spatial asymmetry is important in finding cancers in mammograms since we see that the texture and appearance of cancer are both very similar to the texture and appearance of normal tissue in the breast.
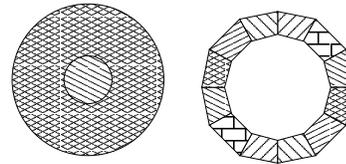

Figure 2: Relative brightness and spiculation detection filters: the left filter calculates the percent of the pixels in the outer ring that are less bright than the least bright of the pixels in the inner disk. The filter highlights the areas of the image that have bright cores, a characteristic of spiculated lesions. Multiple sizes are used because the sizes of the cancers vary. The right filter is used to detect spiculation, or radiating lines out of the core.

Our analysis starts with filtering to find the contextually similar suspicious points that could be cancers in the mammograms. The filter highlights the areas of the image that have bright cores, a characteristic of spiculated lesions. The filter calculates the percent of the pixels in the outer ring that are less bright than the least bright of the pixels in the inner disk to produce a suspiciousness value, and an example is given in Figure 2. This suspiciousness value represents the degree to which the surrounding region of a point radially decreases in intensity, and is done over several sizes. This is focusing on the bright central core of the cancer and ignoring the radiating lines of spiculation. A second filter can be used to detect the radiating lines of spiculation, as shown in Figure 2 on the right, but a combined filter that detects both the cores and the spiculation should improve the performance, especially if the relative weighting of the

measurements is learned on an appropriate data set.

A further improvement might be possible by first transforming the data before filtering, such as applying wavelet analysis to the images before simply thresholding or applying the filter. This has been successfully attempted previously [8] with good results. However, an optimal solution would first combine all of the various filtering and transform methods which create meaningful suspicious points, and then learn an effective analysis from them. This is similar to the effective combination of weak classifiers into a single strong classifier through ensemble learning methods like boosting, which has been successfully used before in tumor classification [6].

Points with a high suspiciousness value have a higher chance of corresponding to an occurrence of cancer. The centroid of each local maxima in the filtered image is initially marked as a candidate detection site with its suspiciousness value. This collection of sites is then sorted in decreasing order of suspicion. All suspicious sites that are closer than 5mm from a more suspicious site are removed. This yields a set of potential detection sites that can be analyzed for asymmetry. This technique was advanced by Heath [14]. Although it may not be the optimal choice of either filtering or ranking, the spatial analysis that we used can be applied to any technique that can rank the suspiciousness of areas. The number of points returned by the filtering step is one of the variables that is learned to optimize the analysis. Alternatively, we can also make use of a threshold on the suspiciousness value instead of taking the top few. However, we chose to take the top few in order to to try to be insensitive to image processing choices that might bias the analysis.

The analysis for similarity or asymmetry that we used performs a comparison of clusters of suspicious points in order to avoid an exact registration. We experimented with three models, two that are built using a space-defined clustering with pre-set cluster registration, and one that is built using a data-defined clustering with no pre-set registration. A space-defined clustered comparison can be seen in Figure 3, where the points are assigned to clusters that are defined by large tracts of space and have a set spatial relationship between each other. The clusters have a pre-set registration with their sister cluster in the other image. For simplicity, the clusters are assumed to be non-overlapping and space-filling, but this is not required. Additionally, the sister clusters are assumed to contain the same areas in the images of the left and right breasts out of symmetry. This reduces the number of parameters and increases the ability of the model to be generalized to a larger data set, based on the assumption that there is no important anatomical differences between the left and right breasts and that breast cancer is equally as likely to be in the left or right breast.

In a space-defined clustered comparison, a cluster is assigned all of the suspicious points in the space that the
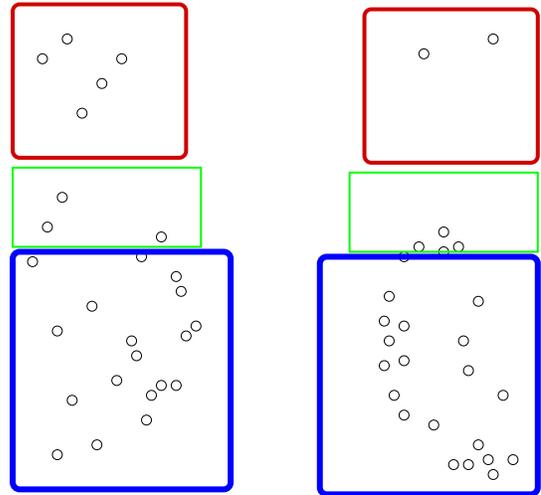

Figure 3: Space-Defined Clustered Comparison. The suspicious points are the small circles, with the points on the left coming from the image of the left breast and the points on the right coming from the image of the right breast. The clusters are the large boxy shapes containing the points, and sister clusters have the same size and border. The points are assigned to clusters that are defined by large tracts of space and have a set spatial relationship between each other. The clusters have a pre-set registration with their sister cluster in the other image.
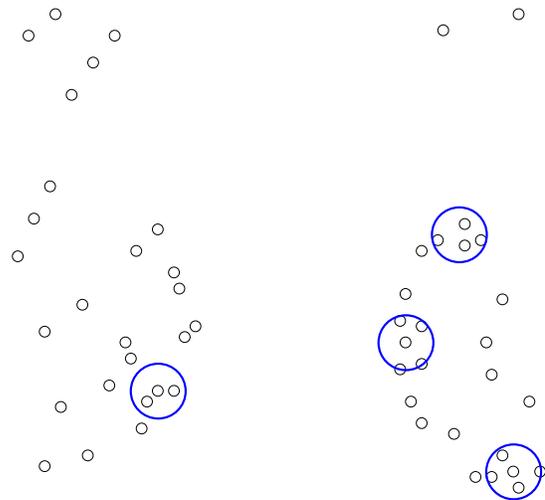

Figure 4: Data-Defined Clustered Comparison. The suspicious points are the small circles, with the points on the left coming from the image of the left breast and the points on the right coming from the image of the right breast. The clusters are the larger circles. This method searches for small clumps of suspicious points and then assigns a cluster there, comparing the number of clusters in the two images.
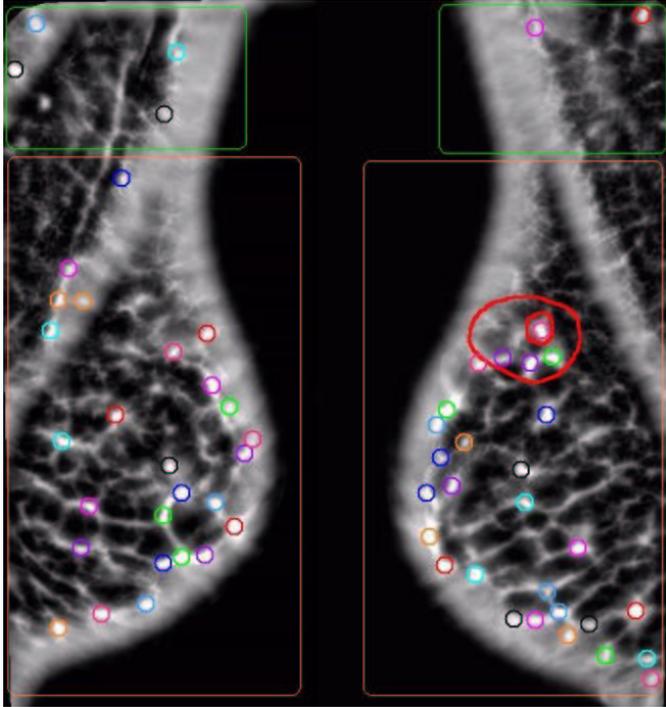
3

Figure 5: The two breasts are shown, with the suspicious points indicated by circles. The two hand-drawn circles (one inside the other) in the right breast are the radiologist's diagnosis of cancer. The asymmetry is demonstrated by the presence of considerably fewer suspicious points in the matching area in the left breast – that is, the distribution of suspicious points changes slightly from one breast to the other when there is cancer. Note that there are circles within the hand-drawn circles, showing that the filtering does find the cancer and that the suspicious points do tend to cluster around the cancer. This was the motivation for the data-defined clustered comparison method. The large boxy outlines are like the clusters used in the two-cluster space-defined model.

cluster spans. Exact registration of the suspicious points is avoided by using the clusters for the comparisons as they are registered with their sister cluster. A comparison of the number of suspicious points contained within the cluster is done with the sister cluster, and the absolute value of the differences between sister clusters are combined and compared against an optimized threshold. A variation on this approach could learn a threshold for each pair of sister clusters, and which has the advantage of being able to emphasize the importance of some areas in the breast over others. This could be used to distinguish noisy areas in the breasts where many spurious suspicious points are found from important areas where even small variations are indicative of malignant cancer. This technique of learning important areas in images using clustered comparisons can be thought of as an image enhancement technique.

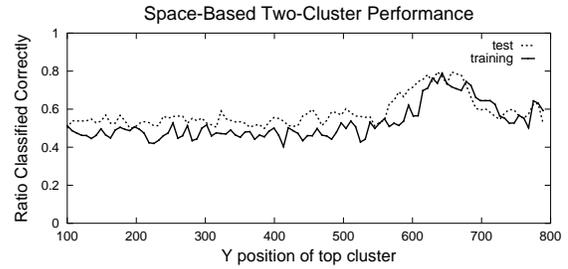. We test two space-defined clustered comparisons. The

Figure 6: The comparison of the analysis on the training and test data for the two cluster space-defined clustered comparison. The Y-axis is the ratio of the number of correctly classified mammograms to the total number, while the X-axis is the position of the bottom of the top cluster. This data corresponds to the situation when cancerous and non-cancerous cases equally weighted. Since the two cluster analysis worked well, a three cluster analysis was tried and is shown in Figure 7.

simplest model assigns the suspicious points to one of two clusters and then compares the clusters in the left and right breasts, as shown in Figure 5. The second model was an extension of the first, using three cluster areas instead of two and shown in Figure 3. These models were motivated by the observation that the cancer would change the distribution of the suspicious points, leading to a different clustering and thus an indication of cancer. An improvement to the method would be to adaptively determine the optimal number of clusters through a split-and-merge type methodology [16].

The third model that we tried does not set the number of clusters arbitrarily, but instead learns the number of clusters from the data. These data-defined clustered comparisons search for small clumps of suspicious points and then assign a cluster there, as shown in Figure 4. The maximum distance between points and the minimum points needed to define a cluster are learned on a training set. The clusters were also defined to be centered on a suspicious point because we believed that small clumps of suspicious points tended to form around the central cancer. This assumption may be incorrect, and having cluster centers free from that constraint may improve the performance. Exact registration is avoided again by registering the clusters instead of the image or the suspicious points. Comparing the number of clusters in the right image versus the number of clusters in the left image is a first cut at registering the clusters since a difference in the numbers of clusters implies some clusters cannot be registered. Improving the cluster registration may improve the performance of the method. This data-defined clustered comparison was motivated by the data, where we observed a small cluster of suspicious points at a cancer site as shown in Figure 5.

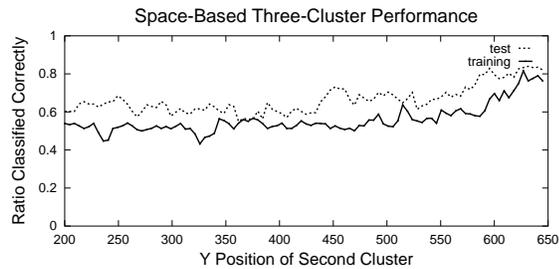For diagnosing breast cancer, the importance of correct

4

Figure 7: The comparison of the analysis on the training and test data for the three cluster space-defined clustered comparison. The Y-axis is the ratio of the number correctly classified mammograms to the total number, while the X-axis is the position of the top of the middle cluster. For this plot, the bottom of the top cluster is set at 645, the high end of the range. Thus the performance improves as the second cluster gets smaller, pushing up on the boundary at the high end. Since a small area is given immense importance by the learning process, we can come to the conclusion that it is a very important area in the diagnosis of breast cancer. Note that the three-cluster comparison does provide improved results over the two-cluster comparison.
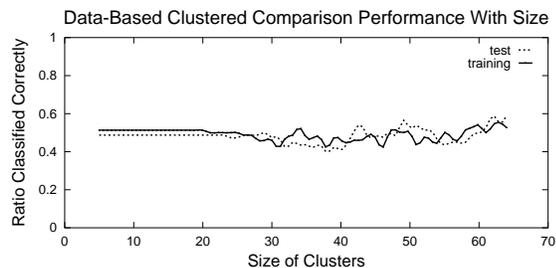

Figure 8: The comparison of the analysis on the training and test data for the data-defined clustered comparison. The Y-axis is the ratio of the number correctly classified mammograms to the total number, while the X-axis is the maximum distance allowed from the center of the cluster, or the size. This plot shows that the analysis is poor across the range of sizes. The space-defined clustering vastly outperforms the more conventional data-defined clustering
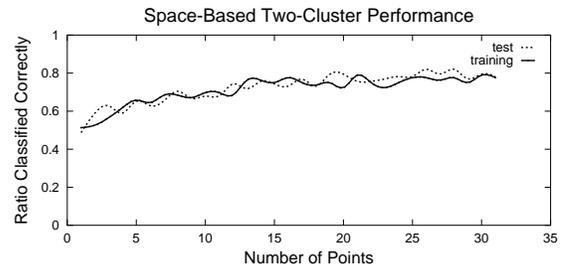

Figure 9: The comparison of the analysis on the training and test data for the two cluster space-defined clustered comparison. The Y-axis is the ratio of the number correctly classified mammograms to the total number, while the X-axis is the number of suspicious points used in the clustered comparison. Note that points do not get fed into the learning algorithm, just the cluster differences, so there is no effect from the learning based on the number of data points. Only one or maybe a few points should be associated with the cancer, and usually that is within the top eight points, so the comparison performs better with more points that are not cancer related.

classification of the cancerous cases is much more important than the non-cancerous cases. To reflect this, the associated weighting of the cancerous cases was varied, and we evaluate the performance of various weightings.

3 Evaluation

The filtering and the clustered comparisons were applied to the mediolateral oblique (MLO) mammogram views of both the left and right breast of patients that were diagnosed with cancer and patients that were diagnosed as normal, or free from cancer. The analysis was performed over test and training data sets, with cases that were roughly split between normal mammograms and mammograms with malignant spiculated lesions from the Digital Database for Screening Mammography [15]. The focus was on one type of breast cancer which creates spiculated lesions in the breasts. Spiculated lesions are defined as breast cancers with central areas that are usually irregular and with ill-defined borders. Their sizes vary from a few millimeters to several centimeters in diameter and they are very difficult cancers to detect [18].

The training set had 39 non-cancerous cases and 37 cancerous cases, while the test set had 38 non-cancerous cases and 40 cancerous cases. The data is roughly spread across the density of the breasts and the subtlety of the cancer. The breast density and subtlety were specified by an expert radiologist. The subtlety of the cancer shows how difficult it is to determine that there is cancer. The training data set was used to determine optimal parameters for classification. These cases indicated that a difference in the clusters of one
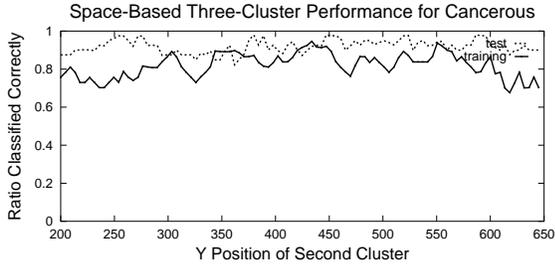
Figure 10: The comparison of the analysis on the training and test data for the three cluster space-defined clustered comparison for only cancerous cases. The Y-axis is the ratio of the number correctly classified cancerous mammograms to the total number of cancerous mammograms, while the X-axis is the position of the top of the middle cluster. For this plot, the bottom of the top cluster is set at 645, the high end of the range. The general flatness of the data suggest that adding the third cluster is not that important for the diagnosis of cancer.

or more suspicious points indicated cancer. The mammogram in Figure 5 shows how the spatial distribution of suspicious points is changed by the presence of a cancer.

4 Results

Our results are good on all cases of the test set, correctly classifying 80% on the test set and 79% on the training set for the two-cluster model as shown in Figure 6. The three-cluster model achieved correct classification 84% and 81% of the time, respectively, as shown in Figure 7, while the small-clusters model results as shown in figure 8 were not much better than a simple Naive Bayes approach. The results are summarized in Table 1. However, it is much more important to correctly classify the cancerous cases, and by heavily weighting the importance of the cancerous cases, we correctly classified 97% of the cancerous cases and 42% of the non-cancerous cases with the two-cluster model. Neither the subtlety nor the density of the cancer had an effect on the results. On the training set, the cancerous cases were correctly classified 100% of the time while the non-cancerous cases were classified correctly just 33% of the time.

The comparison with a commercial system shows that the results are surprisingly good. Correct classification results of 96% of the cancerous cases and 33% of non-cancerous cases are possible using the R2 ImageChecker system [3] when doing pre-screening. Our method showed correct classification results on 97% of cancerous cases and 42% of the non-cancerous cases. This demonstrates the importance of asymmetry and image comparisons in pre-screening. The inclusion of additional factors other than asymmetry in the clustering method should improve the
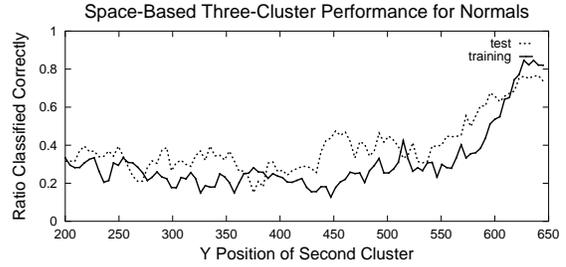

Figure 11: The comparison of the analysis on the training and test data for the three cluster space-defined clustered comparison for only cancerous cases. The Y-axis is the ratio of the number correctly classified cancerous mammograms to the total number of cancerous mammograms, while the X-axis is the position of the top of the middle cluster. For this plot, the bottom of the top cluster is set at 645, the high end of the range. The strong response of the data suggest that adding the third cluster is very important for the diagnosis of normality. Upon inspection of the data we found that this region is actually very empty in many normal mammograms, implying that our analysis has determined an important area for the diagnosis as non-cancerous in mammograms.

results. However, the data sets used are different, as the R2 ImageChecker data contains all cancer types and our method has only the difficult to detect spiculated lesions. The R2 ImageChecker data set also had a much higher proportion of non-cancerous mammograms to cancerous cases.

Since the neural net that we used to learn the parameters of the clusters relies on steepest descent to find the optimal parameters, there is concern about finding a local maxima instead of the global maxima. To evaluate the potential of this data set to become trapped in a local maxima, an exhaustive search over the y position of the top cluster was performed for the two-cluster model, and the results shown in Figure 6 demonstrate a clear maxima with relatively shallow local maxima. Additionally, the analysis is shown to generalize well from training to test data. One of the parameters that was learned was the optimal number of suspicious points to use in the analysis, and the results were always at or near the top of the range that we used, varying from 29 to 32 points depending on the model and weightings as shown in Figure 9. This was surprising because the cancer was usually in the top sixteen if not the top eight points. However, the suspicious points do tend to cluster around a cancer, so including more suspicious points may create a greater distortion of the underlying distribution than fewer points. The learning algorithm does not get the number of points directly, only the cluster differences, so the inclusion of more data should not result in overfitting. An interesting result from the three-cluster analysis showed that space-defined clustering could discover important regions in images, and this is demonstrated in Figures 8, 10, and 11. The analysis found a region of interest for diagnos-

Method	Test Cancerous	Test Normal	Training Cancerous	Training Normal
Space-Defined Two Cluster	87%	71%	79%	78%
Space-Defined Two Cluster Weighted Toward Cancer	97%	42%	100%	33%
Space-Defined Three Cluster	78%	90%	84%	78%
Data-Defined	51%	56%	67%	50%
Naive Bayes	51%	48%	48%	51%
R2 ImageChecker	96%	33%	96%	33%

Table 1: Results Table. The space-defined clustered comparisons performed significantly better than both the data-defined clustered comparisons and a Naive Bayes approach. However, the data-defined comparisons barely outperformed the Naive Bayes approach, demonstrating the inadequacy of this approach as we defined it. Adding a third cluster did improve the overall performance of the space-defined clustered comparisons.

ing a mammogram as non-cancerous.

Our clustering method makes use of a spatial analysis of the suspicious points, counting the number of suspicious points in certain areas or clusters. Its success is an encouraging sign for the investigation and utilization of more complicated non-local analysis techniques in medical imaging and analysis.

5 Conclusion

The overall results of using our techniques are good. Our experiments on malignant cases yielded 97% accuracy suggesting that asymmetry is an important measure to incorporate into prescreening or CAD software. We have also shown that using the cluster comparisons to determine asymmetry is insensitive to the parameters of the clusters. We created and compared three models, demonstrating that three area clusters worked slightly better than two, and showed that comparing the number of small clusters was not an effective technique. The space-defined clusters also discovered an area of interest in mammogram comparisons which improved the diagnosis of mammograms that did not have cancer. More clusters might improve the technique, or, more importantly, they might lead to the discovery of more areas of interest. We suggest several ways to improve on the methods that we used to compare mammograms. One method is to convert a mammogram into a connected graph structure of suspicious points and utilize known graph comparison methods for the measure. Another is to use registration of suspicious points from one breast to the other and reduce the suspiciousness of points that have a similar counterpart.

Our work has demonstrated the potential of utilizing techniques like clustered comparisons and other non-local methods to avoid exact registration and discover interesting results in medical imaging. We have shown that we can effectively measure doctor-defined quantities like asymmetry. We believe that in the future, the combination of using com-

puter vision to capture doctor-defined quantities like asymmetry and machine learning of parameters could be a powerful method for improving the quality of research in medical imaging, and this is one of the avenues of research that we intend to pursue.

References

[1] American Cancer Society. Breast Cancer Facts and Figures 1999- 2000. Atlanta, GA: American Cancer Society, Inc., 1999.

[2] J.A. Anderson, 1995, An Introduction to Neural Networks, The MIT Press.

[3] S. Astley, F.J. Gilbert, Computer-aided detection in mammography, Clinical Radiology, vol 59, pp. 390-9, 2004.

[4] S. Astley, T. Mistry, C.R.M. Boggis, V.F. Hillier, Should we use humans or a machine to pre-screen mammograms? Proceedings of the Sixth International Workshop on Digital Mammography, pp. 476- 480, 2002.

[5] R. Campanini, A. Bazzani, A. Bevilacqua, D. Bollini, D. Dongiovanni, E. Iampieri, N. Lanconelli, A. Riccardi, M. Roffilli, and R. Tazzoli, A novel approach to mass detection in digital mammography based on Support Vector Machines, Proceedings of the 6th International Workshop on Digital Mammography, June 2002.

[6] M. Dettling and P. Buhlmann, Boosting for tumor classification with gene expression data, Bioinformatics, vol 19, no 9, pp. 1061-1069, 2003.

[7] I. El-Naqa, Y. Yang, N.P. Galatsanos, R.M. Nishikawa, M.N. Wernick, A similarity learning approach to content based image retrieval: application

to digital mammography, IEEE Transactions on Medical Imaging, vol 23 , issue 10 , pp. 1233-1244, Oct. 2004.

[8] R.J. Ferrari, R.M. Rangayyan, J.E.L. Desautels, A.F. Frere, Analysis of asymmetry in mammograms via directional filtering with Gabor wavelets, IEEE Transactions on Medical Imaging, vol 20, no. 9, Sept 2001.

[9] T.W. Freer, M.J. Ulissey, Screening mammography with computer- aided detection, Radiology, vol 220, pp. 781-786, 2001.

[10] J. Goldberger, S. Gordon, H. Greenspan, An efficient image similarity measure based on approximations of kl-divergence between two Gaussian mixtures, Proceedings of the Ninth IEEE International Conference on Computer Vision, pp. 487-493, Oct 2003.

[11] I. Gondra, D.R. Heisterkamp, Learning in region-based image retrieval with generalized support vector machines, Proceedings of the Computer Vision and Pattern Recognition, p. 149, 2004.

[12] V. Gudivada and V. Raghavan. Design and evaluation of algorithms for image retrieval by spatial similarity. ACM Transactions on Information Systems, vol 13, issue 2, pp. 115-144, April 1995.

[13] A. Guimond and G. Subsol, Automatic MRI database exploration and applications, Pattern Recognition and Artificial Intelligence, vol 11, no 8, pp. 1345-1365, 1997.

[14] M.D. Heath and K.W. Bowyer, Mass detection by relative image intensity, The Proceedings of the 5th International Conference on Digital Mammography, Medical Physics Publishing (Madison, WI), ISBN 1-930524-00-5, June 2000.

[15] M.D. Heath, K.W. Bowyer, D. Kopans et al, Current status of the digital database for screening mammography, Digital Mammography, Kluwer Academic Publishers, pp 457-60, 1998.

[16] S.L. Horowitz and T. Pavlidis, Picture segmentation by a tree traversal algorithm. Journal of the ACM, vol 23, no 2, pp. 368-388, April 1976.

[17] B.L. Kalman, S.C. Kwasny, and W.R. Reinus, Diagnostic screening of digital mammograms using wavelets and neural networks to extract structure, Technical Report 98-20, Washington University, 1998.

[18] S. Lui, C.F. Babbs, and E.J. Delp, Multiresolution detection of spiculated lesions in digital mammograms, IEEE Transactions on Image Processing, no 6, pp. 874-884, June 2001.

[19] J. Linda, W. Burhenne, S.A. Wood, C.J. D'Orsi, S.A. Feig, D.B. Kopans, K.F. O'Shaughnessy, E.A. Sickles, L. Tabar, C.J. Vyborny, and R.A. Castellino, Potential contribution of computer- aided detection to the sensitivity of screening mammography, Radiology, vol 215, pp. 554-562, 2000.

[20] P. Miller, S. Astley, Detection of breast asymmetry using anatomical features, Proceedings of the International Society for Optical Engineering Conference on Biomedical Image Processing and Biomedical Visualization, vol 1905, p. 433-442, 1993.

[21] P. Sajda, C. Spense, and L. Parra, Capturing contextual dependencies in medical imagery using hierarchical multi-scale models, Proceedings of the IEEE International Symposium on Biomedical Imaging, pp. 165-168, 2002.

[22] M. Sallam and K.W. Bowyer, Registering time-sequences of mammograms using a two-dimensional unwarping technique, Digital Mammography, K.Doi, M.L. Giger, R.M. Nishikawa, and R.A. Schmidt, Eds. Amsterdam, the Netherlands: Elsevier, 1996, pp. 291-96.

[23] H.A. Swett and P.L.Miller, ICON: a computer-based approach to differential diagnosis in radiology, Radiology, vol 163, pp. 555-558, 1987.

[24] N. Vujovic and D. Brzakovic, Establishing the correspondence between control points in pairs of mammographic images, IEEE Transactions on Image Processing, vol 6, no 10, pp. 1388-1399, Oct. 1997.

[25] M.A. Wirth, A. Jennings, A nonrigid-body approach to matching mammograms, Proceedings of the IEEE Image Processing and its Applications, pp 484-7, 1999.

[26] M.A. Wirth, J. Narhan, and D. Gray, Nonrigid mammogram registration using mutual information, SPIE: Medical Imaging: Image Processing, San Diego, CA, vol 4684, pp. 562-73, Feb. 2002.

[27] Jianhua Yao, R. Taylor, Assessing accuracy factors in deformable 2D/3D medical image registration using a statistical pelvis model, Proceedings of the Ninth IEEE International Conference on Computer Vision, vol 2, pp. 1329-1334, Oct 2003.

[28] F.F. Yin, M.L. Giger, K. Doi, C.E. Metz, C.J. Vyborny, R.A. Schmidt, Computerized detection of masses in digital mammograms: analysis of bilateral subtraction images, Medical Physics, vol 18, pp 955-63.