Mammographic Density Classification using Multiresolution Histogram Information

Izzati Muhimmah, *Student Member, IEEE* , and Reyer Zwiggelaar

***Abstract*— Mammographic density is known to be an important indicator of breast cancer risk. Quantitative estimation approaches based on histogram information have been investigated previously. However, claims have been made that greylevel information might be insufficient to discriminate between complex density classes. A multi-resolution histogram technique, which was developed as a texture analysis approach, has been investigated as an alternative classification space. Using a DAG-SVM classifier on the MIAS database the result shows an agreement of 77.57% between automatic and expert radiologist manual classification.**

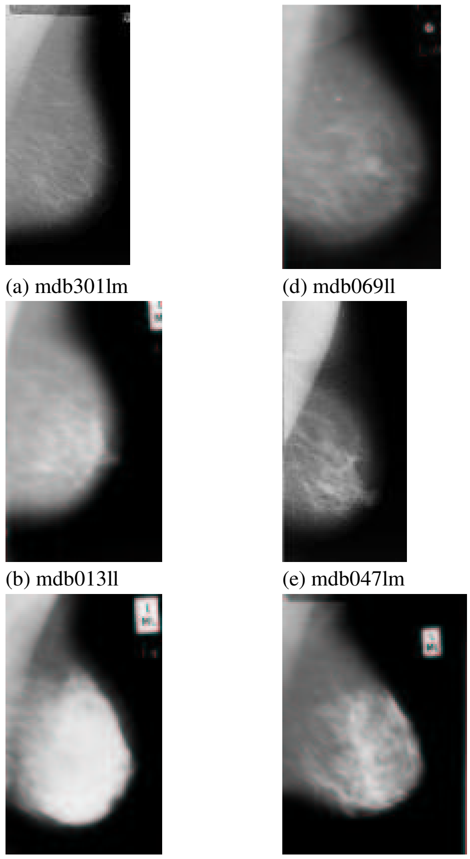
I. INTRODUCTION

MAMMOGRAPHIC density is known to be an important indicator of breast cancer risk. Some examples of different mammographic densities can be seen in Fig. 1. There are four metrics which are used in practice to relate the mammographyc parenchymal patterns and the risk of breast cancer, namely: Wolfe's four parenchymal patterns [1], Boyd's Six Class Categories [2], BI-RADS [3] and Tabár's Five Patterns [4]. The comparative study of these four mammographic based assessment approaches, especially in MIAS database [5], has been reported in [6].

Mammographic images are 2D projections of the x-ray attenuation properties of the 3D breast tissue along the path of the x-rays. The connective and ephiteal tissues attenuate more than fatty tissues. This phenomenon can be interpreted that brighter areas on mammographic films represent glandular tissues, whereas the dark areas represent fatty tissues. A brighter mammographic image appearance would suggest that the breast is composed mostly of glandular tissue rather than fatty tissue. This is referred to as a dense mammographic image. From an image analysis point of view, it can be seen that mammographic densities correspond to image intensities. This leads to the modeling based on histogram information.

Since Wolfe has introduced the mammographic risk assessment [1]; automatic breast parenchymal pattern classification methods have been investigated. Many studies have used (semi-) automatic histogram-based breast density segmentation and/or classification, for example: Byng *et al.* proposed an interactive thresholding technique [7], Karssemeijer proposed to use the skewness, rather than the standard deviation, of histograms of local breast areas [8], Sivaramakhrisna *et al.* introduced variation images to enhance the contrast between dense and fatty areas and subsequently Kittler's optimal threshold was used to segment the densities [9], Zhou *et al.* suggested an adaptive dynamic range compression technique [10], and Masek *et al.* presented results based on the comparison of four histogram distance measures [11]. In general, the described histogram based approaches for automatic density estimation produce robust and reliable results.

However, overall classification results, mostly in comparison with expert assessment, tend to be low. Masek *et*



(a) mdb301lm (d) mdb069ll

(b) mdb013ll (e) mdb047lm

(c) mdb171ll (f) mdb145lx

Fig. 1. Example mammograms: (a,d) Fatty, (b,e) Gandular, and (c,f) Dense.

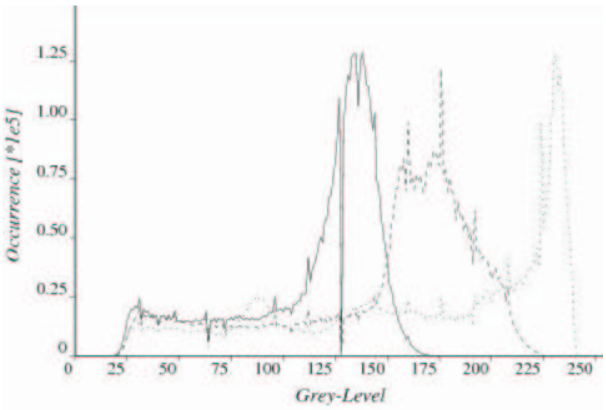
Manuscript received June 30, 2006. This work was supported in part by the Islamic Development Bank under the IDB Merit Scholarship and University of Wales Aberystwyth Studentship.

Izzati Muhimmah is with Department of Computer Science, Univeristy of Wales Aberystwyth, SY23 3DB UK, on leave from Department of Informatics, Universitas Islam Indonesia, Yogyakarta, 55584 Indonesia (e-mail: emma@ fti.uii.ac.id).

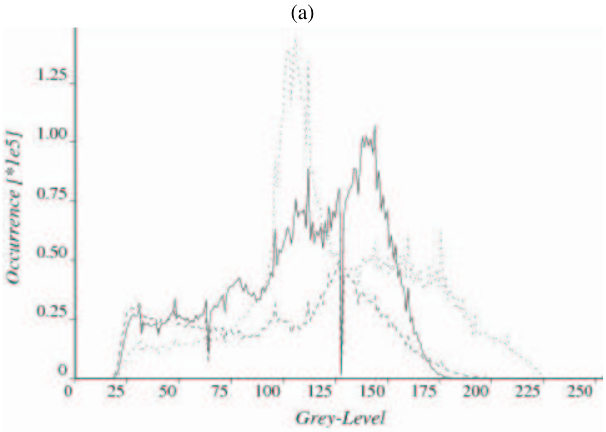
Reyer Zwiggelaar is with the Department of Computer Science, University of Wales, Aberystwyth, SY23 3DB UK (phone: +44-1970-628691; fax: +44-1970-628536; e-mail: rrz@aber.ac.uk).

al. used average histograms of original images of each density class as a feature for triple MIAS classification and an agreement of 62.42% was achieved using a Euclidean distance measure [11]. Zwiggelaar *et al.* reported statistical grey-level histogram modeling (PCA) for triple MIAS classification and agreement of 71.5% was reported [12].

The study by Zhou *et al.* [10] showed that there were some typical histogram patterns for each density class (see Figure 2 (a) for three distinct histogram patterns from the MIAS database [5]). Yet, due to the statistical nature of histograms, they also pointed out that there are relatively similar histograms that represent different risks (see Figure 2 (b) for examples).



(a)



(b)

Fig. 2. Example histograms for the MIAS density classes. The three histograms in (a) represent the fatty (continous line), glandular (dashed line), and dense (dotted line) mammographic examples from Fig. 1. a-c and show good separation with respect to greylevel. The three histograms in (b) represent three other mammograms (using the same key for fatty, glandular and dense mammograms) from the MIAS database which are displayed in Fig. 1 d-f for which it might be difficult to separate them just on greylevel information.

In practice, radiologists assess the breast density not solely based on brightness. They also take into account the breast tissues patterns and distribution (textures). So, it is expected that the inclusion of texture information in the classification process will lead to improvements. In addition, a recently published paper by Hadjidemetriou *et al.* [13] showed that different generic texture images with similar histograms can be discriminated by a multi-resolution histogram approach.

The main aim of this study is to investigate whether the multi-resolution histogram technique proposed by Hadjidemetriou *et al.* [13] does capture a combination of intensities and textures features sufficient to automatically classify mammographic densities. We also investigated if this multi-histogram approach gains significantly in mammographic density classification compared to single-scale histogram information.

The remainder of this paper is outlined as follows: data and the proposed methodology are described in Section II. Section III gives results of the proposed method and a discussion on our findings. Finally, conclusions appear in Section IV.

II. Methods and Data

A. *Multi-resolution Histogram Technique*

The main aim is to obtain a feature space which can be used to discriminate between the various mammographic density classes. We followed Hadjidemetriou's algorithm [13] to form these feature vectors (see Fig. 4):

- Construct image pyramid. We have built a five-level Gaussian pyramid [14] of the breast area (ignoring the pectoral muscle and background areas) of digitised mammograms. In our pyramid structure, level 0 is the original image (base). See Fig. 3.

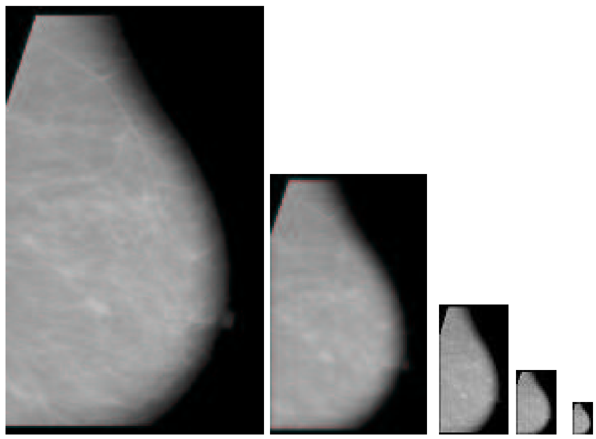
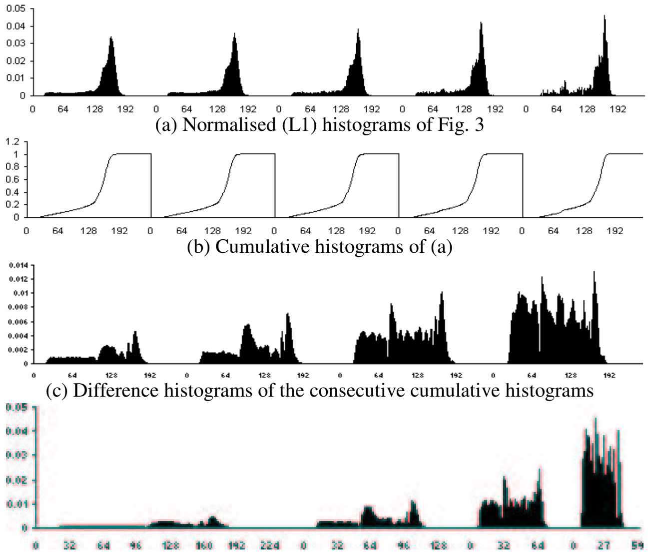


Fig. 3. Five-level Gaussian pyramid of mammogram mdb097ll.

- Compute the histograms $\{h_0, h_1, h_2, h_3, h_4\}$, where h_i is a row vector related to the histogram of resolution i and i = 0 is the highest resolution. In total, the combined histogram length was 1280.
- Normalise histograms with L_1 norm. This is similar to normalise the histograms with respect to the breast area. See Fig. 4 (a).
- Compute cumulative histograms. Prior to this, the histograms are smoothed with a Gaussian filter of window size of 5. See Fig. 4 (b).
- Compute difference histogram between consecutive levels. The difference histogram has length of 1024. See Fig. 4 (c).

- Sub-sample difference histograms and renormalise. We used a sub-sampling factor of $2^{2/3}$. In other words, the consecutive histogram lengths were shorter by the factor of $2^{2/3}$. See Fig. 4 (d).
- Concatenate the difference histograms to form the feature vector. The dimensionality of the resulting feature space is equal to 580. This will be referred as the multihistogram feature.



(a) Normalised (L1) histograms of Fig. 3

(b) Cumulative histograms of (a)

(c) Difference histograms of the consecutive cumulative histograms

(d) Sub-sampled difference histograms to form the features vector

Fig. 4. Formation of the feature vector.

B. Data

We evaluated the proposed technique on 321 mammogram images of the Mammographic Image Analysis Society (MIAS) database and classified them based on the provided three class categories (Fatty, Glandular, Dense) [5]. It should be noted that the database has equal proportions for these three classes.

C. Evaluation

For classification, an automatic method is developed based on the feature vectors in combination with a multi-class Directed Acyclic Graph - Support Vector Machine (DAG-SVM) classifier [15, 16].

It is known that mammographic intensities vary with exposure levels and film characteristics [17, 18]. And, in an imaging session, a woman likely had the mammogram taken using similar films and/or exposure levels. The MIAS database consists of pairs of mammograms; hence, to minimise bias, we treated the left and right mammograms independently. Thus, we trained the classifier on all the left (right) mammograms according to a leave-one-image-out methodology.

For evaluation, we combine the results of the left and right mammograms in the dataset. The results are presented in the form of confusion matrices. Percentage agreements and linear-weighted Kappa values (κ) were calculated [19].

III. RESULTS

The results of the multi-class DAG-SVM classifier for single histogram information h_0 and h_4 can be found in Table I. The multihistogram feature results are shown in Table II.

TABLE I

DAG-SVM, SINGLE-HISTOGRAM BASED CLASSIFICATION.

The proportion of breast tissue is represented as F: Fatty, G: Fatty-glandular, and D: Dense-glandular

		MIAS Classification		
		F	G	D
Auto Class	F	88	22	7
	G	14	65	40
	D	4	17	64

(a) h_0

		MIAS Classification		
		F	G	D
Auto Class	F	91	19	9
	G	11	73	37
	D	4	12	65

(b) h_4

Agreement based on normalised single histogram information h_0 and h_4 were 67.60% (κ = 0.594) and 71.34% (κ = 0.628), respectively (see Table I). It can be seen that feature h_4 gave improvement in classifying the fatty and the glandular classes compare to h_0. These two features gave similar classification rate for the dense group.

TABLE II

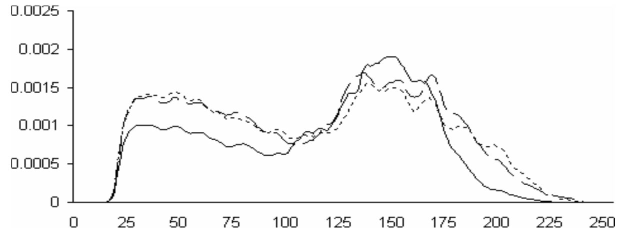
DAG-SVM, MULTI-RESOLUTION HISTOGRAM BASED CLASSIFICATION.

The proportion of breast tissue is represented as F: Fatty, G: Fatty-glandular, and D: Dense-glandular

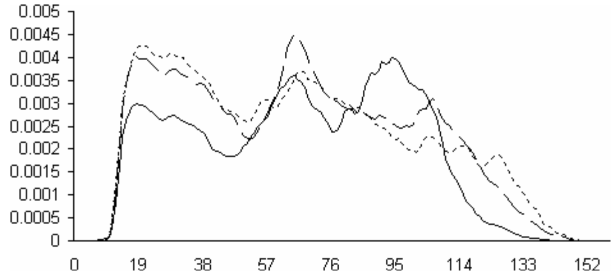
		MIAS Classification		
		F	G	D
Auto Class	F	97	19	5
	G	3	65	19
	D	6	20	87

On the other hand, the agreement based on multihistogram information was 77.57% (κ = 0.717), see Table II. There are significant improvement in classification rates for the fatty and the dense groups in comparisons to either h_4 or h_0. The multihistogram feature seemed to perform slightly less in recognizing the glandular group than h_4, but similar to the h_0 results.

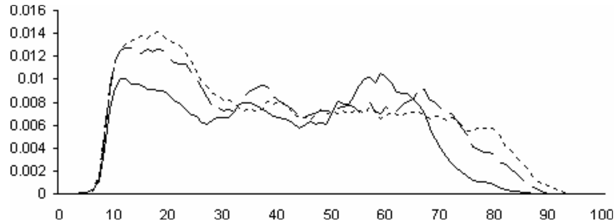
The average of the multi-histogram features of each classes (fatty, glandular, and dense) can bee seen in Fig. 4. The fatty histogram patterns shows differences with those of glandular and dense multi-histogram patterns (χ^2 statistic between fatty-glandular was 0.174 and as for fatty-dense was 0.291). The multi-histogram pattern of the glandular class was quite similar to those of dense mammograms (χ^2 statistic between glandular and dense groups was 0.046).



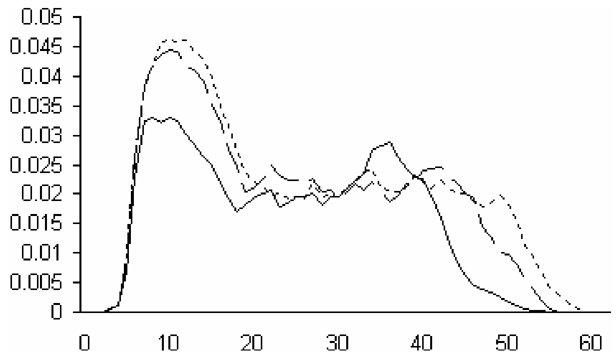
(a) level 0, full resolution



(b) level 1



(c) level 2



(d) level 3

Fig. 4. Detail views of the average multihistogram feature for the MIAS density classes. The three histograms in (a)-(d) represent the fatty (continous line), glandular (dashed line), and dense (dotted line).

IV. Discussion

A. Correct Classification

An example of image which was misclassified when using single histogram features, but correctly classified when multihistogram features were used, can be seen in Fig. 5. The overall multi-histogram feature outperforms the single histogram features, i.e. h_0 and h_4. Hence, we could infer that the addition of information yields improved classification results.

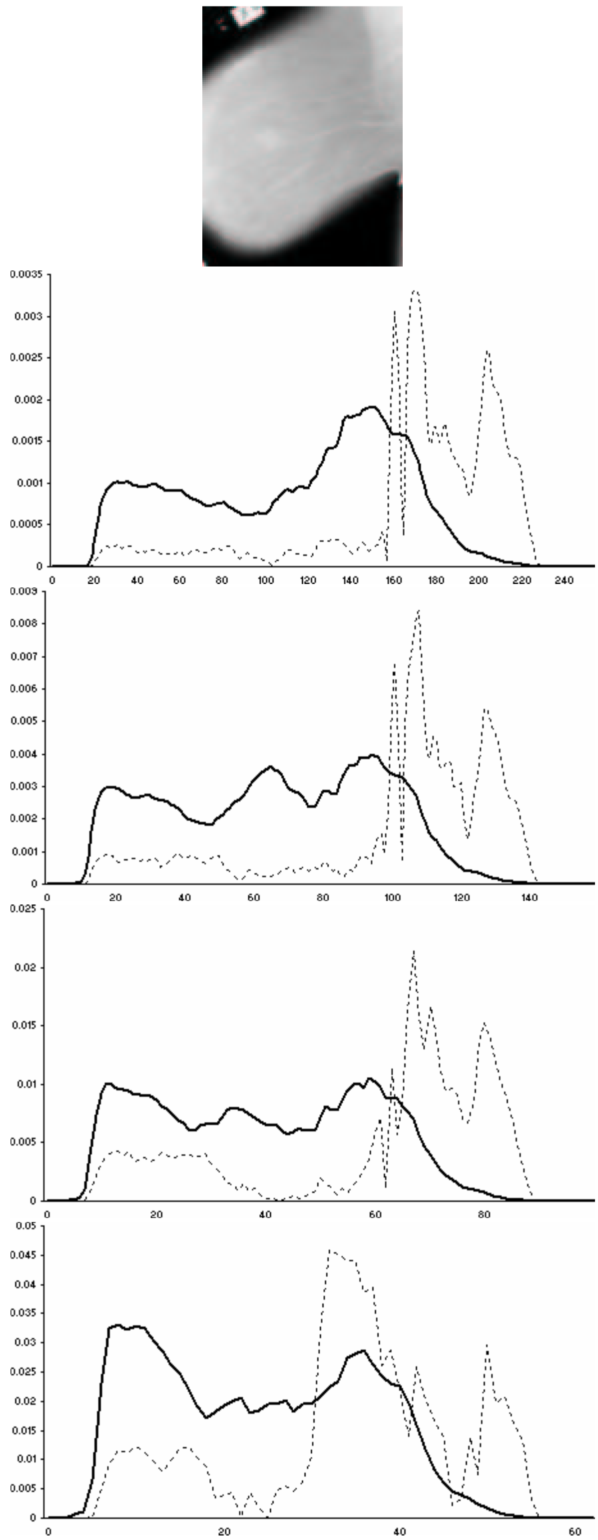


Fig. 5. Example of the image which was correctly classified using multihistogram feature while their single histogram features gave misclassification results. The thick continuous line represents the average of fatty features and the doted one represents multi-histogram feature of the above shown fatty image.

In comparison with published results on histogram-based density classification using the MIAS database; Masek *et al.* reported an agreement of 62.42% [6] and Zwiggelaar *et al.* reported an agreement of 71.5% [7]. Our results based on the multi-class DAG-SVM classifier and multi-resolution histogram information of 77.57% show improvement over existing techniques. This also indicates that texture information, as described by multi-scale histogram information, does provide additional discrimination potential. It should be noted that the results are based on a large dataset of over 300 mammographic images and that bias in the results has been avoided by effectively using a leave-one-woman-out methodology for the classifier stage.

B. *Failed cases*

There were three images which were consistently over classified as dense (see Fig.6. a) and two images which always be classified as fatty (see Fig.6.b) regardless the features (h_0, h_4 and multihistogram) used. In Fig.6. the multihistogram features are shown along with their average features (thick continuous line).

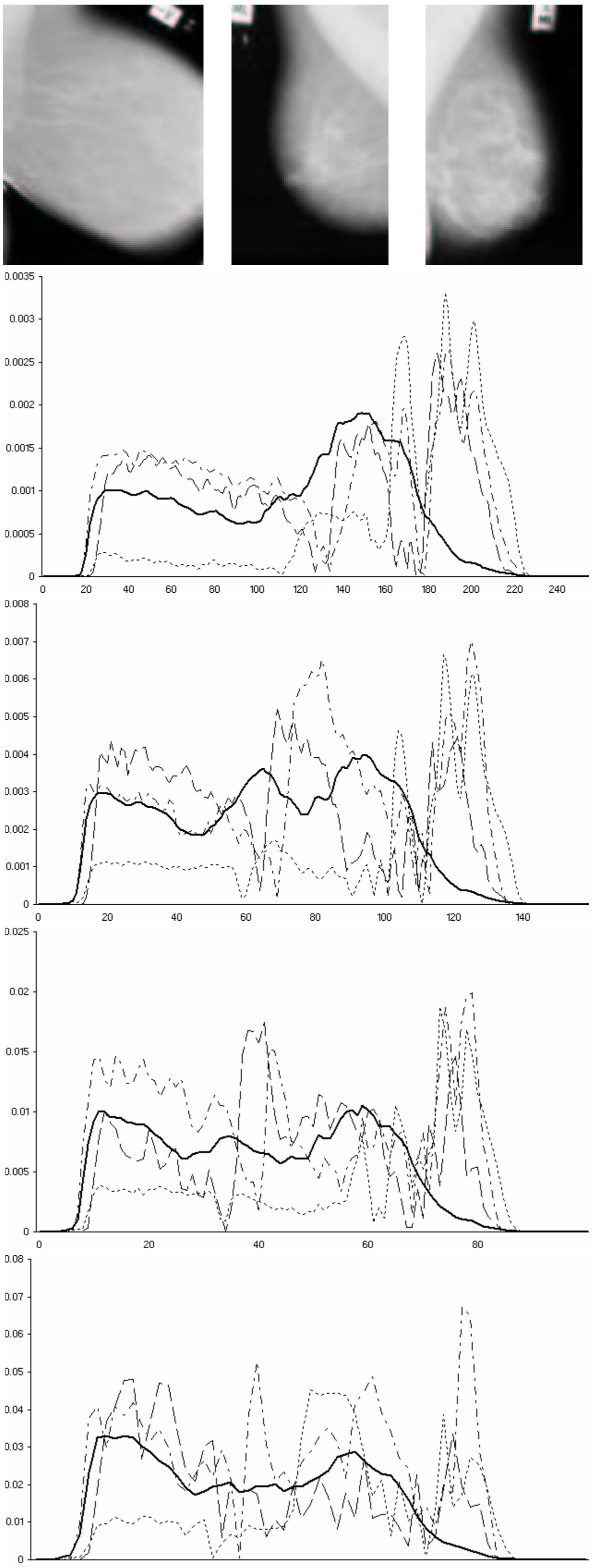


Fig. 6.a. Example of the failed cases: fatty as dense. Original images are shown at the top.

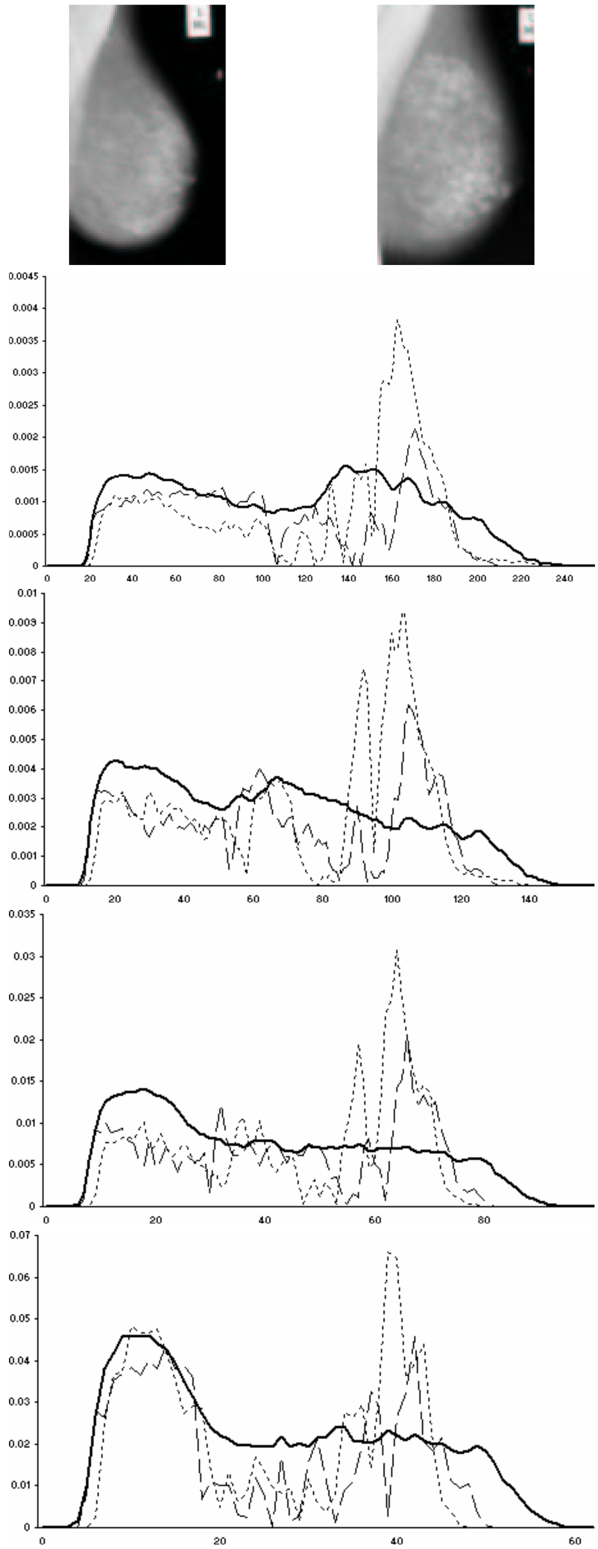


Fig. 6. b. Example of the failed cases: dense as fatty. Original images are shown at the top.

C. Future works

It should be noted that h_4 performed better in recognizing glandular group rather than either the multi-resolution feature or h_0. In order to achieve higher classification rates, the possible group-specific-features and/or weighting factor of the features will be investigated.

The intensities-texture combination is used by expert radiologist in practice to assess the mammographic risk. Several other techniques to obtain this combination features need to be investigated to get the most effective feature for breast density classification.

Future development will include two aspects. Evaluation using other mammographic density metrics, e.g. Boyd's Six Class Category and/or BIRADS, and on an alternative mammographic data will be investigated. Alternative classifiers will be investigated to provide a full comparison between the effects of various classification approaches and the selection of the most appropriate techniques.

V. Conclusion

In conclusion, we have investigated the combination of texture and intensities information for breast density classification using multi-resolution histogram technique. Based on the presented results, this approach gave an agreement of 77.57% when a DAG-SVM classifier was used. This shows improvements over published work.

References

[1] J.N. Wolfe, “Risk for breast cancer development determined by mammographic parenchymal pattern,” *Cancer*, vol. 37, pp. 2486-2492, 1976.

[2] N.F. Boyd, J.W. Byng, R.A. Jong, E. Fishell, L. Little, A. Miller, G. Lockwood, D. Tritcher and M.J. Yaffe, “Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening study,” *Journal of the National Cancer Institute*, vol. 87, pp. 670-675, 1995.

[3] ACR: Breast Imaging Reporting and Data Systems (BI-RADS) 2nd edition. Reston: American College of Radiology, 1995.

[4] L. Tabár, T. Tot, and P.B. Dean. *Breast Cancer: The Art and Science of Early Detection with Mammography*. Georg Thieme Verlag, 2005.

[5] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal, and J. Savage, “The Mammographic Image Analysis Society Digital Mammogram Database.,” *Excerpta Medica. International Congress Series*, vol. 1069, pp. 375-378, 1994.

[6] I. Muhimmah, E.R.E. Denton, and R. Zwiggelaar. “Comparison between Wolfe, Boyd, BI_RADS and Tabár based mammographic risk assessment,” in *LNCS 4046 Digital Mammography. P roc. 8th IWDM 2006*, Manchester, June 2006, pp. 407-415.

[7] J.W. Byng, N.F. Boyd, E. Fishell, R.A. Jong, and M.J. Yaffe, “The quantitative analysis of mammographic densities,” *Physics in Medicine and Biology*, vol. 39, pp. 1629-1638, 1994.

[8] N. Karssemeijer,, “Automated classification of parenchymal patterns in mammograms,” *Physics in Medicine and Biology*, vol. 43, pp. 365-378, 1998.

[9] R. Sivaramakhrishna, N.A. Obuchowski, W. A. Chilcote, and K. A. Powell, “Automatic Segmentation of Mammographic Density,” *Academic Radiology*, vol. 8, no. 3, pp. 250-256, March, 2001.

[10] C. Zhou, H.P. Chan, N. Petrick, M.A. Helvie, M. Goodsitt, B. Sahiner, and L.M. Hadjiiski, “Computerized image analysis: Estimation of breast density on mammograms,” *Medical Physics*, vol. 28, no. 6, pp. 1056-1069, June, 2001.

[11] M. Masek, S.M. Kwok, C.J.S. deSilva, and Y. Attikiouzel, “Classification of Mammographic Density Using Histogram Distance Measures,” in *CD-ROM Proceedings of the World Congress on Medical Physics and Biomedical Engineering*, August, 2003, pp. 1.

[12] R. Zwiggelaar, I. Muhimmah, and E.R.E. Denton, “Mammographic Density Classification Based on Statistical Grey-level Histogram Modelling,” in *Proc. Medical Image Understanding and Analysis 2005*, Bristol, July 2005, pp. 183-186.

[13] E. Hadjidemetriou, M.D. Grossberg, and S.K. Nayar, “Multiresolution Histograms and Their Use for Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 7, pp. 831-847, July, 2004.

[14] P.J. Burt and E.H. Adelson, “The Laplacian Pyramid as a Compact Image Code,” *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532-540, April, 1983.

[15] J.C. Platt, N. Cristianini, and J. Shawe-Taylor, “Large Margin DAGs for Multiclass Classification,” *Advances in Neural Information Processing Systems*, vol. 12, pp. 547-553, MIT Press, 2000. Available: http://research.microsoft.com/~jplatt/DAGSVM.pdf

[16] G.C. Cawley. (2000). MATLAB Support Vector Machine Toolbox (v0.55β). Available: http://theoval.sys.uea.ac.uk/~gcc/svm/toolbox

[17] *The fundamentals of medical radiography.* Fuji Photo Film Co., Ltd. *FUJIFILM Technical Handbook: Medical X-ray System, 2003.*

[18] *Fundamentals of sensitized materials for radiography.* Fuji Photo Film Co., Ltd. *FUJIFILM Technical Handbook: Medical X-ray System, 2003.*

[19] R. Lowry. KAPPA Calculator. (Accessed 7/3/06). Available: http://faculty.vassar.edu/lowry/kappa.html