Available online at www.sciencerepository.org

Science Repository

SCIENCE REPOSITORY

Research Article

A Digital Mammogram Auto Classification Method Based on Fibroglandular Breast Tissue Density Evaluation by Image Similarity

Takuji Tsuchida[1,2], Toru Negishi[2] and Toshihiro Kai[3]*

[1]*Department of Radiological Technology, Saitama Saiseikai Kawaguchi General Hospital, Nishikawaguchi, Kawaguchi City, Saitama, Japan*
[2]*Department of Radiological Sciences, Graduate School of Human Health Sciences, Tokyo Metropolitan University, Higashi-Ogu, Arakawa-ku, Tokyo, Japan*
[3]*Shintoshin Ladies Mammo Clinic, 3F Capital Building, Yoshinomachi, Oomiya-ku, Saitama City, Saitama, Japan*

ARTICLE INFO

ABSTRACT

It is crucial to assess the fibroglandular breast tissue density to define the degree of the risk that the healthy breast tissue will obscure the lesions. Subjective assessment criteria, proceed by the reading physicians by using the mammary gland concentrations on mammograms, are defined as the breast classification method. However, due to the existence of between observer's variability, a computer-based quantitative classification method is required. The conventional method classifies according to the ratio of the Dmg region (mammary gland region) to the Dc region (fibroglandular breast tissue region). However, this does not include subjective evaluation elements. The purpose of this study is to improve the concordance rate with the subjective assessment by performing an automated classification based on image similarity. First, 130 cases of right MLO (Medio-Lateral Oblique) images, subjectively classified as fatty tissue, mammary gland diffuseness, non-uniform high density, and high density, were reclassified to two groups; fatty tissue and mammary gland diffuseness as Non-Dense breast, and non-uniform high density and high density as Dense breast. Next, as for evaluation images, 33 cases of both sides MLO images taken by different mammography devices were used. Finally, the image similarity analysis result using Normalized Cross-Correlation between the search image and the evaluation image was derived, and the degree of coincidence of subjective breast classification was calculated. As a result, the concordance rate between the conventional method and the subjective evaluation results of this method improved from 73 % to 91 %, and the kappa coefficient improved from 0.49 to 0.81. This result indicates that our approach is more useful for the automated classification of mammograms based on fibroglandular breast tissue density.

Introduction

The relationship between breast density and breast cancer risk was the first to reported in 1976 by Wolf *et al.* [1-3]. Since then, several studies show that the risk of developing breast cancer is two to six times higher for women with dense breasts than women in the lowest density category [4-7]. These studies suggested that, with the introduction of digital mammography, breast density classification has become more difficult, and this means that high-density mammary glands may hide breast cancer [8-10]. Evaluation of the fibroglandular breast tissue density has been based on the subjective evaluation of observers that Wolfe classification, Breast Imaging Reporting and Data System (BI-RADS) and Tabár classification [2, 11, 12].

The most common breast density reporting method is the BI-RADS which uses four categories and the American College of Radiology (ACR) proposes a modified version of the Wolfe patterns [11]. The previous edition BI-RADS used to categorize by the percentile of the mammary gland tissue in each breast (<25 %, 25 %-50 %, 51 %-75 %, or >75 %). With the new edition BI-RADS, the percentile has eliminated, and the breast density evaluation was preceded only by the subjective evaluation [11]. Previous studies showed a considerable inter- and intra-reader variability when using BI-RADS [13, 14]. To correct

these variabilities, a semi-automated and automated method to quantify breast density has been studied [15-17]. Also, the accuracy of the breast density assessment software on the market has been reported [18, 19].

The correlation between the software and subjective evaluation for these studies had positive feedbacks. However, these methods use pectoralis major muscle or dense breast on the mammogram as the threshold values and calculate the breast density, and it cannot evaluate between different mammography devices [15-17]. The breast density assessment software currently on the market is expensive and has not yet become popular. As for the conventional method, the classification method of Matsubara *et al*. and Yamazaki *et al*., which uses the ratio of Dmg region (mammary gland region) to Dc region (fibroglandular breast tissue region) reported a high concordance rate with the subjective evaluation result based on the BI-RADS [16, 17]. However, this does not include an evaluation image taken by a different model nor positioning influence. Therefore, we performed an analysis using image similarity measured by template matching as a new decision index for breast density classification [20, 21].

Materials and Methods

Evaluation objects were the mammograms taken at Shintoshin Ladies Mammo Clinic and reclassified to Non-Dense breast group (fatty, mammary gland diffuseness) and Dense breast group (non-uniform high density, high density) from subjective breast classification result (fatty, mammary gland diffuseness, non-uniform high density, and high density) based on BI-RADS [11].

Table 1: Subjective breast classification result of the evaluation images.

Mammographic breast composition	Non- Dense Breast		Dense Breast	
	Fatty	Mammary gland diffuseness	Non-uniform high density	High density
Evaluation image (n=33)	0	11	20	2

Thirty-three cases of MLO, age between 30 to 70, which were difficult to evaluate as Non-Dense breast or Dense breast with the subjective breast classification by the 22 reading physician (kappa coefficient between physicians were 0.34), were set as an evaluation image. The breast classification result was 11 Non-Dense breast cases and 22 Dense breast cases (Table 1). Also, we used the higher density side when the evaluation of the breast classification differs bilaterally. Mammography X-ray devices used for this study were AMULET FDR-3000AWS (Fuji Film Corporation, Tokyo, Japan) and a combination of MGU-1000D (Canon Medical, Tokyo, Japan) and FCR PROFECT CS (Fuji Film Corporation, Tokyo, Japan). AMULET FDR-3000AWS adopts direct conversion FPD on the X-ray detection part, and its matrix size is 3540×4740 matrix size (0.05mm pixel size) with a 14-bit grayscale output. FCR PROFECT CS is a CR system with 3540×4740 matrix size (0.05 mm pixel size) and 10-bit grayscale output.

For the conventional method, we followed the analysis method of Matsubara *et al*. and Yamazaki *et al*. and set the pixel value of the pectoralis major as a standard. Skin line and pectoralis major region extract, and dynamic range compression were done as a preprocessing. Next, the Dmg region was determine based on the surface of the skin

line (Dmg: the area where mammary gland thought to exist primarily). Then the Dc region will be extracted from the Dmg region. We set the definition for the Dc region as mammary gland density equivalent to the pectoralis major muscle density, and there is a high possibility that cancer may be hiding in that area. The case was defined as high-density breast when the Dc region ration within the Dmg region was 50 % and up [16, 17].

Our method proceeded template matching between the evaluation image and the search image group, then analyzed using the image similarity by Normalized cross-correlation values (NCC) as an index. The range of NCC is from -1.0 to 1.0 and shows the higher similarity between the two images when the NCC value is higher. PC spec for image analysis was Windows 10, a 64-bit operating system with Intel® Corei5 Processor with 4GB of mounted memory. Also, with application development, we used MATLAB R2017a 64-bit.
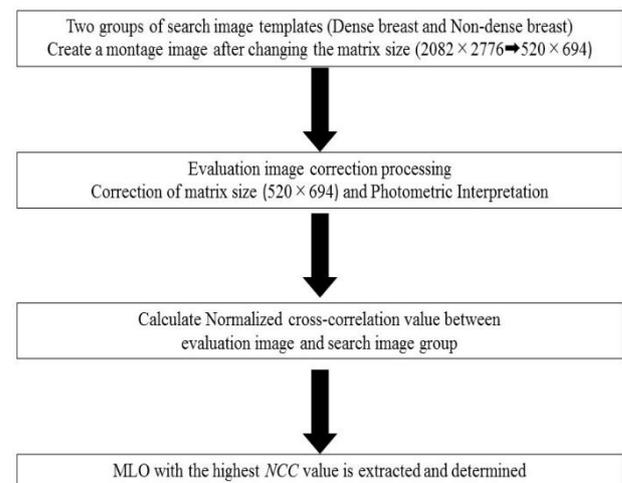

Figure 1: Abstract of the auto-classification method.

Figure 1 shows the outline of our method. First, we create a search image group. The result of subjective breast classification, 130 cases of left MLO, were recategorized to the Non-Dense breast group and the Dense breast group (Table 2). Mammomat Inspiration Software Versions(s) VB41A (SIEMENS Healthineers Japan.) was used to take the search images. Mammomat Inspiration is a direct conversion FPD, with 2082×2776 matrix size (0.085 mm pixel size) and 12-bit grayscale output.

Table 2: Subjective breast classification result of the search images.

Mammographic breast composition	Non- Dense Breast		Dense Breast	
	Fatty	Mammary gland diffuseness	Non-uniform high density	High density
Search image (n=130)	0	66	61	3

The analysis accuracy of the NCC has considerable influence on the image noise, and also, needs to shorten the analysis time and lessen the memory consumption depending on the number of search images. As for these correction methods, we set each image scale size to 0.25 times size and changed the matrix size to 520×694 from 2082×2776. We employed Bilinear interpolation for pixel interpolation processing. Classified Non-Dense breast and Dense breasts were converted to montage images

within the group. Creating the montage can shorten the processing time by reducing the number of loops while NCC analysis. Secondary, the matrix size of the evaluation images was changed to 520×694 to match the search images as a correction process. At the same time, the correction process is necessary for the left MLO case and reversed cases with Photometric Interpretation, which shows the relationship between pixel value and display intensity.

Next, to obtain the image similarity between the evaluation image (A) and the search image (B), the Normalized cross-correlation values (NCC) were calculated. The formula is defined below: (1) & (2). I and J express the height and width of the region within the selected evaluation image, and $\bar{a}$ and $\bar{b}$ represent the average pixel value. σ_A and σ_B shows as the standard deviation of each evaluation image (A) and search image (B). NCC was calculated only with the overlapped domain between two images. Higher NCC value indicates a higher similarity compare to the lower values. Therefore, by comparing the highest NCC value from each Non-dense breast and Dense breast group, the higher NCC value was considered as a density classification result (Figure 2).
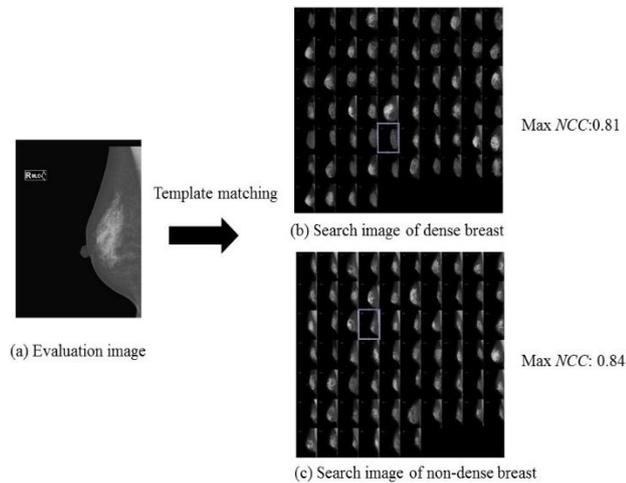

Figure 2: An example of a correctly identified case with the highest correlation value obtained by Normalized cross-correlation values for illustration of template matching in the evaluation image and the search images. The reading physician classification was Non-dense breast **(a)**. The dense breast **(b)** and Non-dense breast **(c)** used as search montage images. The above example was classified as a non-dense breast with a higher NCC value.

Finally, the breast density classification result by the conventional method, and our proposed method was compared by concordance rate and the kappa coefficient with subjectively evaluated breast classification [16, 17].

$$NCC = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}\frac{\{A(i,j)-\bar{a}\}\{B(i,j)-\bar{b}\}}{\sigma_A \cdot \sigma_B} \qquad (1)$$

where

$$\bar{a} = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}A(i,j)$$

$$\bar{b} = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}B(i,j)$$

$$\sigma_A = \sqrt{\frac{\sum_{j=1}^{J}\sum_{i=1}^{I}\{A(i,j)-\bar{a}\}^2}{IJ}} \qquad (2)$$

$$\sigma_B = \sqrt{\frac{\sum_{j=1}^{J}\sum_{i=1}^{I}\{B(i,j)-\bar{b}\}^2}{IJ}}$$

Results

Thirty-three cases of evaluation images were classified using the database with 133 cases of search images to evaluate our proposed method. The subjective breast classification result of the employed images was proceeded by 22 reading physicians who are certified by The Japan Central Organization on Quality Assurance of Breast Cancer Screening. Table 3 shows the classification performance of the conventional method and our method. The concordance rate of the conventional method and the subjective breast classification result was 73 % with a 0.49 kappa coefficient, while the concordance rate improved to 91 % with a 0.81 kappa coefficient with our proposed method.

As for the Non-dense breast, the concordance rate with subjective breast classification result was 100 % (11 out of 11 cases matched) with both the conventional method and our method. On the other hand, the concordance rate of the Dense breast was 59 % (13 out of 22 cases matched) with the conventional method and 86 % (19 out of 22 cases matched) with our method.

Table 3: The classification performance of the conventional method and our proposed method.

		Computerized classification					
		Previous method			New method		
		Non- Dense Breast	Dense Breast	Match rate [%]	Non- Dense Breast	Dense Breast	Match rate [%]
Reading Physicians' classification	Non- Dense Breast	11	0	100 (11/11)	11	0	100 (11/11)
	Dense Breast	9	13	59.1 (13/22)	3	19	86.4 (19/22)
Total				72.7 (24/33)			90.9 (30/33)

Discussion

Our method corresponds well with the breast density classification result provided by the reading physician. This result shows that it is useful as an objective and quantitative alternative method for breast cancer evaluation. It is also automated, easily reproducible, and able to eliminate the inter- and intra-reader variability. Nine cases of mismatch occurred with the conventional method, and this thought to happen since the pixel value of the pectoralis major muscle is set as a threshold value to extract the Dc region and could not eliminate the effect from the positioning and photographic conditions.

Our method had three mismatch cases. Figure 3a shows the Dense breast case which matched with the subjective classification result. NCC=0.862 with the matching case within the Dense breast search images, and NCC=0.821 with the matching case within the Non-dense breast search images. Also, (Figures 3b & 3c) shows the mismatched case of the same patient with the subjective evaluation result. Figure 3b was taken with FPD, and (Figure 3c) was with CR. The classification results, Dc/Dmg, with the conventional method were 7.7 % and 14.7 % and classified as Non-dense breasts. (b) showed NCC=0.859 with the matching case within the Dense breast search images, and NCC=0.883 with the matching case within the Non-dense breast search images. (c) showed NCC=0.779 with the matching case within the Dense breast search images, and NCC=0.810 with the matching case within the Non-dense breast search images.
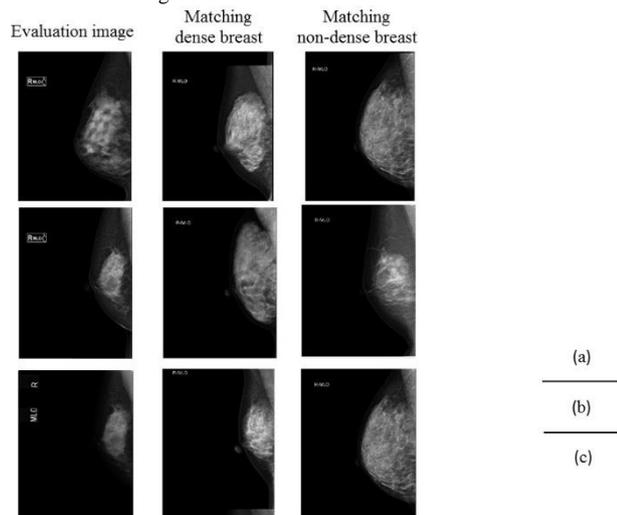

Figure 3: The cases classified as Dense breasts with the subjective evaluation. Case **(a)** matched with the subjective evaluation. With our method, NCC=0.862 with the matching case within the Dense breast search images, and NCC=0.821 with the matching case within the Non-dense breast search images. On the other hand, cases **(b)** and **(c)**, taken from the same patient, were the mismatching case with the subjective evaluation. **(b)** was taken with FPD, and **(c)** was with CR. **(b)** showed NCC=0.859 with the matching case within the Dense breast search images, and NCC=0.883 with the matching case within the Non-dense breast search images. **(c)** showed NCC=0.779 with the matching case within the Dense breast search images, and NCC=0.810 with the matching case within the Non-dense breast search images.

The Normalized Cross-Correlation, which we employed, performs the matching by characterizing the high contrast of the mammogram.

Therefore, direct X-ray, the shape of the papillary area and the breast margin, mammary gland parenchyma and fat distribution become the main component of the matching and lead to the result of (Figures 3b & 3c). Also, the previous study reported a higher matching rate with local regions compared to the whole pulmonary field with the template matching of the chest images [20]. Therefore, as an additional study, we attempted a template matching with the local regions, such as below the nipple and Dc region, where thought to have a high residual rate of the mammary gland, but the matching rate was low with the breast density classification result, given by the reading physician. As for the cause, mammogram has a strong influence from the breast shape and positioning, and unable to fix the trimming of the region of interest. Also, the case determined as a dense breast with the subjective evaluation by the reading physician, even with the local high mammary gland density region, this case will be classified as a Dense breast. As for future improvement, it is necessary to consider the detail of the trimming settings in the region of interest to be able to process the breast density classification. Additionally, it is clear that to construct a large search image database is possible due to the improvement of the processing capability of the analyzing PC.

Conclusion

To define the degree of the risk that the mammary gland may obscure the lesions, we have constructed an algorithm that automatically classifies the mammogram to each breast density category. When applying our method to 33 cases, taken with the different model from the search image database, almost 90 % matched with the classification result driven by the skilled reading physicians. The usefulness of our method for the automated digital mammogram classification became clear in this study.

Acknowledgements

Conflicts of Interest

None.

Ethical Approval

Application approval was obtained by following the code of ethics of Tokyo Metropolitan University Graduate School and Saitama Kawaguchi General Hospital.

Consent

Informed consent for all images, including the database used for this study, was renunciation by the Institution Review Board.

REFERENCES

1. Wolfe JN (1976) Breast patterns as an index of risk for developing breast cancer. *AJR Am J Roentgenol* 126: 1130-1137. [Crossref]

2. Wolfe JN, Saftlas AF, Salane M (1987) Mammographic parenchymal patterns and quantitative evaluation of mammographic densities：a case-control study. *AJR Am J Roentgenol* 148: 1087-1092. [Crossref]

3. Huo CW, Chew GL, Britt KL, Ingman WV, Henderson MA et al. (2014) Mammographic density-a review on the current understanding of its association with breast cancer. *Breast Cancer Res Treat* 144: 479-502. [Crossref]

4. Boyd NF, Martin LJ, Yaffe MJ, Minkin S (2011) Mammographic density and breast cancer risk: current understanding and prospects. *Breast Cancer Res* 13: 223. [Crossref]

5. Eng A, Gallant Z, Shepherd J, McCormack VA, Li J et al. (2014) Digital mammographic density and breast cancer risk: a case-control study of six alternative density assessment methods. *Breast Cancer Res* 16: 439. [Crossref]

6. McCormack VA, dos Santos Silva I (2006) Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 15: 1159-1169. [Crossref]

7. Vachon CM, van Gils CH, Sellers TA, Ghosh K, Pruthi S et al. (2007) Mammographic density, breast cancer risk and risk prediction. *Breast Cancer Res* 9: 217. [Crossref]

8. Kerlikowske K, Hubbard RA, Miglioretti DI, Geller BM, Yankaskas BC et al. (2011) Comparative effectiveness of digital versus film-screen mammography in community practice in the United States: a cohort study. *Ann Intern Med* 155: 493-502. [Crossref]

9. Kerlikowske K, Zhu W, Tosteson AN, Sprague BL, Tice JA et al. (2015) Identifying women with dense breasts at high risk for interval cancer: a cohort study. *Ann Intern Med* 162: 673-681. [Crossref]

10. Prummel MV, Muradali D, Shumak R, Majpruz V, Brown P et al. (2016) Digital Compared with Screen-Film Mammography: Measures of Diagnostic Accuracy Among Women Screened in the Ontario Breast Screening Program. *Radiology* 278: 365-373. [Crossref]

11. Reston VA (2013) Breast imaging reporting and data system (BI-RADS) atlas. 5th ed. *American College of Radiology*.

12. He W, Denton ERE, Stafford K, Zwiggelaar R (2011) Mammographic image segmentation and risk classification based on mammographic parenchymal patterns and geometric moments. *Biomed Signal Process Control* 6: 321-329.

13. Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B et al. (2005) Categorizing breast mammographic density; intra- and interobserver reproducibility of BI-RADS density categories. *Breast* 14: 269-275. [Crossref]

14. Gard CC, Aiello Bowles EJ, Miglioretti DI, Taplin SH, Rutter CM (2015) Misclassification of Breast Imaging Reporting and Data System (BI-RADS) Mammographic Density and Implications for Breast Density Reporting Legislation. *Breast J* 2: 481-489. [Crossref]

15. Jamal N, Ng KH, Looi LM, McLean D, Zulfiqar A et al. (2006) Quantitative assessment of breast density from digitized mammograms into Tabar's patterns. *Phys Med Biol* 51: 5843-5857. [Crossref]

16. Matsubara T, Tsuchimoto T, Hara T, Fujimoto H (2000) A Classification Scheme for Mammograms Based on the Evaluation of Fibroglandular Breast Tissue Density. *Japan J Med Electron Biol Engine* 38: 93-101.

17. Yamazaki D, Matsubara T, Hara T, Fujita H (2000) Improvement of a Computerized Automatic Classification Method of Digital Mammograms Based on Assessment of Fibroglandular Breast Tissue Density. *Med Imag Informat Sci* 17: 161-166.

18. Christine N Damases, Patrick C Brennan, Claudia Mello Thoms, Mark F McEntee (2016) Mammographic Breast Density Assessment Using Automated Volumetric Software and Breast Imaging Reporting and Data System (BIRADS) Categorization by Expert Radiologists. *Acad Radiol* 23: 70-77. [Crossref]

19. Holland K, van Zelst J, den Heeten GJ, Imhof Tas M, Mann RM et al. (2016) Consistency of breast density categories in serial screening mammograms: A comparison between automated and human assessment. *Breast* 29: 49-54. [Crossref]

20. Toge R, Morishita J, Sasaki Y, Doi K (2013) Computerized image-searching method for finding correct patients for misfiled chest radiographs in a PACS server by use of biological fingerprints. *Radiol Phys Technol* 6: 437-443. [Crossref]

21. Murakawa S, Tanigawa A, Uchiyama Y, Muramatsu C (2015) Kernel eigensapace template matching for detection of lacunar infarcts on MR images. *Nihon Hoshasen Gijutsu Gakkai Zasshi* 71: 85-91. [Crossref]