G-008

Automatic Mass Detection for Breast Cancer from Mammography with Texture Analysis Technique and Statistical Features

ティワタンサクン　ソムチャノク †　　内村　圭一 †
Somchanok Tivatansakul　　Keiichi Uchimura

1. Introduction

To detect breast abnormalities, classify their types and indicate their regions from mammography, several techniques in image processing and computer vision have been applied for a computer-aided diagnosis (CAD). The breast abnormalities can be classified into two major cases: micro-calcifications and masses. In recent years, the accuracy of micro-calcification detection is very high. However, the accuracy of mass detection still need an improvement because it has the difficulty to detect masses in dense breast tissue, since the intensity and contrast between mass regions and dense breast areas do not clearly differ from each other on mammography. Furthermore, there are many types of mass need to be detected with unclear boundaries and the variation of size and shape such as circumscribed, spiculated, microlobulated and ill-defined masses [1-2]. The types of mass are categorized from different sizes, shapes, or boundaries. This paper presents an improved method to detect and segment the region of breast mass abnormalities.

2. The proposed method

This study focused on precise detection of mass boundary from mammography. We adapted and applied a gray-level co-occurrence matrix (GLCM) with statistical features and edge detection which were previously used for color edges extraction [3]. We also improved the method with pre-processing and GLCM iterations to detect masses.

2.1 Pre-processing

After the radiologists selected one mass region in mammography to locate a region of interest (ROI). Then, the method cropped the ROI and pre-processed the mass image by enhancing the contrast using histogram equalization, removing some breast tissue (background) from mass region using exponential operation, and adjusting the brightness with Otsu's thresholding [4] and a half of standard deviation to make the mass region differ from remaining breast tissue. The method normalized the intensity of all pixels from [0, 255] to [0, 7] for constructing GLCM in next step.

2.2 GLCM iteration with statistical features

The gray-level co-occurrence matrix or GLCM presents the appearance frequency between couple of pixels in particular distances and directions from a 3x3 sliding window [5-6].

In this study, we improved the method of previous study [3] to satisfy our breast mass detection results. We created, summarized and normalized four GLCMs with distance of 1 and directions of 0^o, 45^o, 90^o and 135^o to get a global GLCM. In general, various statistical features can be extracted from GLCM. However, only six features: mean, diagonal moment, contrast, energy, inverse difference moment, and variance were extracted from global GLCM. After that, all statistical features were summarized and normalized to create an attributed image. Finally, four-time iteration of GLCM process was applied to the attributed image in order to further remove breast tissue (background).

2.3 Edge detection

Eight-directional edge detection with Robinson mask was applied to the attributed image. The maximum value (max) of eight edge responses in eight directions for each pixel was calculated and compared with Otsu's threshold to determine edge (if max≥ Otsu) or non-edge pixels (if max < Otsu) [3-4]. After we got an edge response image, we selected a large region again on the edge image as a final mass region.

3. Preliminary evaluation

Since our method improved from the previous one [3], we conducted this preliminary evaluation to explain the enhancement points and reasons as following:

- Our method applied only six statistical features without the features of directivity and entropy because they made noise when summarizing all features together that affected to the smoothness of an attributed image. The edge detection result using six statistical features was more clearly.
- Our method applied iteration of GLCM with statistical features because the iterations technique gradually decreased the values of statistical features to near zero. In every step of the iterations, breast tissue (background) which had low GLCM and low statistical features values was gradually removed but the mass region which had high values was remained. Using four-time iteration of GLCM with edge detection was better to detect breast mass region.

Therefore, the enhancement of our improved method was currently suitable for breast mass detection.

4. Evaluation

We evaluated the mass detection results of our improved method with ground truth images in pixel level. We also calculated sensitivity, specificity and accuracy [7].

4.1 Dataset

The mini-MIAS database of mammograms (MIAS) [8] are applied to evaluate our improved method for breast mass detection. Thus, 55 mammography from MIAS which contain the following mass types: CIRC (well-defined/circumscribed masses), SPIC (spiculated masses), and MISC (other, ill-defined masses) were selected. Finally, 58 tumors from 55 mammography image were tested and analyzed with the ground

† Graduate School of Science and Technology, Kumamoto University

Table 1 Evaluation results of breast mass detection

Type	Sensitivity	Specificity	Accuracy
All	76.14%	83.00%	79.90%
CIRC	68.72%	85.85%	80.05%
MISC	82.62%	84.66%	82.70%
SPIC	80.40%	78.09%	77.50%

truth images generated by a medical doctor from Kumamoto University.

4.2 Results and Discussions

Table 1 shows the evaluation results of breast mass detection for three mass types. It indicated that our method achieved 79.90% of accuracy with 76.14% of sensitivity and 83.00% of specificity to detect three mass types. Moreover, MISC detection got higher accuracy than CIRC and SPIC with 82.70%. The accuracy of MISC and CIRC detections were higher than the accuracy of SPIC. This means that our method is better to detect MISC and CIRC than SPIC because the boundaries of MISC and CIRC are more clearly than SPIC. SPIC is the most difficult mass type to detect since its boundary is poorly defined in breast tissue [9]. Furthermore, we realized that if any masses had clear boundary and well distinguished from breast tissue, our method can accurately detect them (Figure 1: CIRC). If they infiltrated into high dense breast area with unclear boundary, we need to enhance contrast, remove breast tissue, and emphasize mass region. If these were good, our method can well detect them (Figure 1: MISC). However, if some breast tissue still remained in the image, our method might incorrectly detect them since the intensity of masses and breast tissue (background) were quite similar in mammography images (Figure 1: SPIC). Thus, we take this case into account as our future works.

In summary, our improved method based on color edge detection technique using four-time iteration of GLCM with six statistical features and eight-directional edge detection is more suitable for detection of MISC and CIRC mass types. Moreover, it might be useful for radiologists to confirm the diagnosis results.

5. Conclusion

This study improved a previous color edge extraction method to accurately detect breast mass region from mammography using pre-processing, six statistical features, and four-time GLCM iteration technique. Our improved method can detect CIRC, SPIC, and MISC with 79.90% of accuracy, 76.14% of sensitivity and 83.00% of specificity. We realized that our method is more suitable for detecting CIRC and MISC. However, we need to improve it to detect masses that infiltrated into high dense breast area with unclear boundary such as SPIC because the intensity of masses and breast tissue (background) were quite similar. Thus, we take this case into account as our future works.

Acknowledgement

We would like to thank Assoc. Prof. Uchiyama Yoshikaza, Faculty of Life Science, Kumamoto University for making ground truth images. We also would like to thank Assoc. Prof. Koutaki Gou, Priority Organization for Innovation and Excellence, Kumamoto University.

Reference

[1] A. Oliver, J. Freixenet, J. Marti, E. Pérez, J. Pont, E.R. Denton, and R. Zwiggelaar, "A review of automatic mass detection and segmentation in mammographic images," Medical image analysis, vol. 14, no. 2, 2010, pp.87-110.
[2] J. Chu, H. Min, L. Liu, and W.Lu, "A novel computer aided breast mass detection scheme based on morphological enhancement and SLIC superpixel segmentation," Medical physics, vol. 42, no.7, 2015, pp.3859-3869.
[3] S.B. Chaabane, and F. Fnaiech, "Color edges extraction using statistical features and automatic threshold technique: application to the breast cancer cells," Biomed. Eng, vol. 13, no. 4, 2014, p.1-18.
[4] N. Otsu, "A threshold selection method from gray-level histograms,"Automatica, vol. 11, no.285-296, 1975, pp. 23-27.
[5] F. Albregtsen, "Statistical texture measures computed from gray level coocurrence matrices," Image processing laboratory, department of informatics, university of oslo, 2008, pp. 1-14.
[6] C.G. Eichkitz, J. Davies, J. Amtmann, M.G. Schreilechner, and P. de Groot, "Grey level co-occurrence matrix and its application to seismic data," FIRST BREAK, vol. 33, no.3, 2015, pp.71-77.
[7] Wikipedia, "Sensitivity and Specificity," Available at: https://en.wikipedia.org/wiki/Sensitivity_and_specificity [Accessed: May, 2016].
[8] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, and P. Taylor, "The mammographic image analysis society digital mammogram database," InExerpta Medica, International Congress Series, vol. 1069, 1994, pp. 375-378.
[9] S. Halls, "Mass shape, margin, and density as found with screening mammography," Available at: http://breast-cancer.ca/mass-chars/ [Accessed: May, 2016].

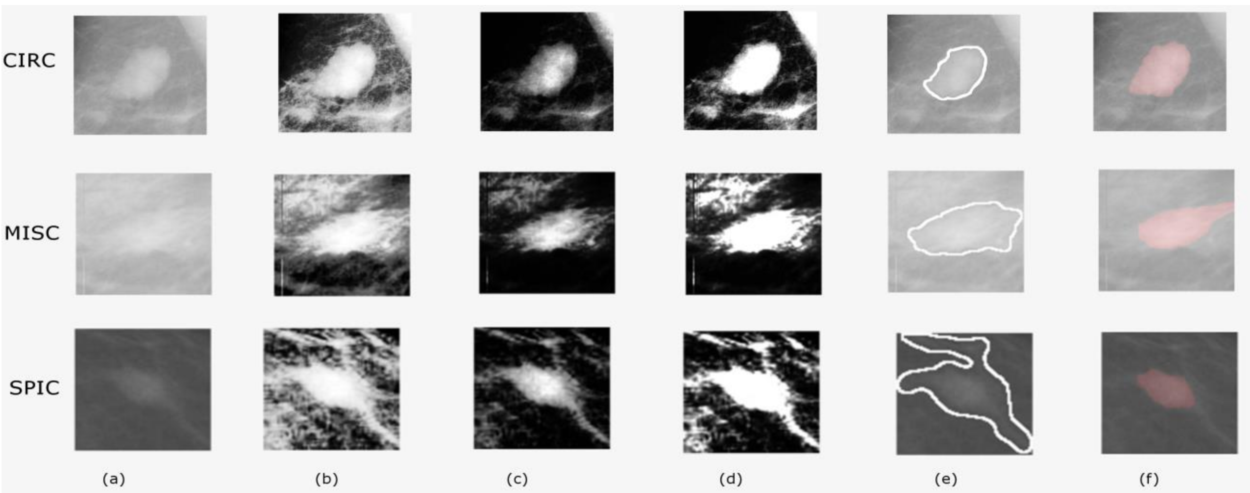

Figure 1 Detection results of masses: (a) original images, (b) images after histogram equalization, (c) images after exponential operation, (d) images after adjust brightness, (e) mass detection results from our method, (f) ground truth images