PROGRESS IN BIOMEDICAL OPTICS AND IMAGING

Reprinted from

Medical Imaging 2001

Image Processing

19–22 February 2001
San Diego, USA

Proceedings of SPIE
Volume 4322

Detection of clustered microcalcifications in masses on mammograms by artificial neural networks

Xuejun Zhang[a], Takeshi Hara[a], Hiroshi Fujita[a], Takuji Iwase[b] and Tokiko Endo[c]

[a] Dept. of Inform. Sci., Gifu Univ., Gifu City, 501-1193, Japan

[b] Dept. of Breast Surgery, Aichi Cancer Center Hospital, Aichi, 464-8681, Japan

[c] Dept. of Radiology, Nat. Hosp. of Nagoya, Nagoya, 460-0001, Japan

ABSTRACT

The existence of a cluster of microcalcifications in mass area on mammogram is one of important features for distinguishing the breast cancer between benign and malignant. However, missed detections often occur because of its low subject contrast in denser background and small quantity of microcalcifications. To get a higher performance of detecting the cluster in mass area, we combined the shift-invariant artificial neural network (SIANN) with triple-ring filter (TRF) method in our computer-aided diagnosis (CAD) system. 150 region-of-interests around mass containing both of positive and negative microcalcifications were selected for training the network by a modified error-back-propagation algorithm. A variable-ring filter was used for eliminating the false-positive (FP) detections after the outputs of SIANN and TRF. The remained FPs were then reduced by a conventional three-layer artificial neural network. Finally, the program identified clustered microcalcifications form individual microcalcifications. In a practical detection of 30 cases with 40 clusters in masses, the sensitivity of detecting clusters was improved from 90% by our previous method to 95% by using both SIANN and TRF, while the number of FP clusters was decreased from 0.85 to 0.40 cluster per image.

Keywords: mammogram, microcalcification, mass, shift-invariant artificial neural network, triple-ring filter analysis, computer-aided diagnosis (CAD)

1. INTRODUCTION

Breast cancer is a dangerous disease in women, approximately 185,700 new breast cancer cases were diagnosed and 44,300 U.S. women died from this disease. Women take the risk of 1 in 8 being diagnosed with breast cancer, and 1 in 30 will die of this disease in her lifetime [1]. The occurrence of breast cancer in Japan is also increasing due to the westernization of life style. Early detection of the disease prolongs life expectancy and may carry out thymectomy before metastasis. Mammography using x-rays is currently considered to be a major significant way for detecting the abnormalities in breast as early as possible, which has been proven to reduce 30 to 35 percent in breast cancer mortality rates. However, radiologists do not detect all breast cancers that are visible on the mammograms, especially to the inexperienced residents or general radiologists. Double reading of mammograms is very helpful and may improve radiologists' performance. Unfortunately, the growing increase in the number of screening mammograms makes this option unpractical. Alternatively, a computer-aided diagnosis (CAD) system may act as a "second reader" to assist the radiologist in detecting and diagnosing lesions. A number of research groups have attempted to analyze mammographic abnormalities with digital computers [2,3].

Microcalcifications are often the first and sometimes the only radiographic findings in early, curable breast cancers. Furthermore, the existence of a cluster of microcalcifications in mass on mammograms is one of important features for distinguishing the breast cancer between benign and malignant. However, it often fails to be detected because of its low subject contrast in denser background and small quantity of microcalcifications. To assist radiologists in detecting clustered microcalcifications on mammograms, an automated computerized scheme based on triple-ring filter and feature extraction

Medical Imaging 2001: Image Processing, Milan Sonka, Kenneth M. Hanson, Editors,
Proceedings of SPIE Vol. 4322 (2001) © 2001 SPIE · 1605-7422/01/$15.00

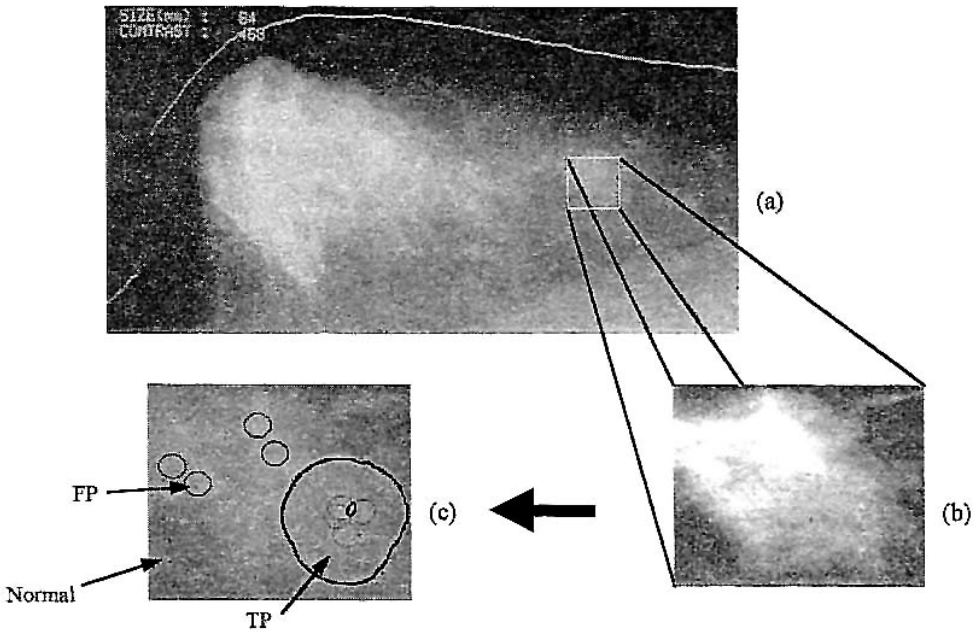


Figure 1: The detection flowchart of the experiment in our mammography CAD system. The mass area (b) was firstly extracted from the breast region on the digitized mammogram (a) as a region of interest (ROI), then the cluster area was picked up according to the number of detected microcalcifications. TP: true-positive cluster of microcalcifications, FP: false-positive cluster of microcalcifications. Normal: normal tissues.

techniques had been developed by our group [3,4]. We also applied a conventional artificial neural network (ANN) to improve the detection performance [5]. However, the issue to be solved was to efficiently detect the clustered microcalcifications in mass area with least number of false positives (FPs). The shift-invariant artificial neural network (SIANN) had been shown to be a powerful tool to eliminate over half of FP clusters without any loss of the true-positive (TP) clusters [6]. In this work, we combined the SIANN with TRF method in our CAD system to get a higher performance of detecting clusters in mass area, and a conventional ANN was used for the purpose of reducing FPs.

2. MATERIALS AND METHOD

2.1 Case selection

40 cases with clusters in mass were selected from the database at Aichi Cancer Center Hospital in Japan. Digitized mammograms with a pixel size of 0.1 mm in 10 bits were used for detecting microcalcifications, and the compressed 0.4-mm-pixel-size images were employed for extracting mass candidates in our CAD scheme [7].

2.2 The steps of detection

In Figure 1, a mass area was extracted in a mammogram with 0.4 mm effective pixel size, while the detection of clustered microcalcifications was performed on an image with "original" 0.1 mm pixel size. Our overall scheme proposed for detecting microcalcifications in masses includs the steps written below:

1. Extraction of Breast Region
2. Detection and Extraction of Mass Area
3. Contrast Correction
4. Density Gradient Calculation

5. Triple-ring Filter (TRF) and SIANN Analysis
6. Variable-ring Filter Analysis
7. ANN Analysis for Eliminating FP Microcalcifications
8. Feature Analysis
9. Classification of Cluster and Display

2.3 Extraction of breast region

The border of the skinline in each mammogram was extracted by investigating the change of density profile for eliminating the textual background and non-mammary area. The smallest rectangle only containing the breast region was automatically extracted for further processing.

2.4 Detection and extraction of mass area

Our program for detecting mass candidates was based on single-image segmentation by employing a standard adaptive thresholding technique. Then a method by comparing both right and left images was employed to eliminate FPs. The suspicious mass area was extracted for further detecting clustered microcalcifications.

2.5 Contrast correction

The contrast correction was used to isolate the signals from heterogeneous background. In this study, the contrast was defined as the difference between the minimum pixel value of a microcalcification and its background pixel value. In our previous papers [4], we found that the contrast of the microcalcifications was less than it should be if microcalcifications exist in a denser mammary gland area, higher absorption and dispersion of the x-ray beam in the surrounding tissues. So, the detection of microcalcifications in mass areas became more difficult than in a fatty area. Here we applied a contrast correction technique to the enhancement of contrast by using an approximated contrast correction curve. 423 microcalcifications confirmed by an experienced radiologist from 20 mammograms were plotted on coordinate with axis of background pixel value and contrast. As shown in Figure 2, we could draw an approximated one-dimensional curve expressed by:

$$y = \alpha x + \beta \qquad (1)$$

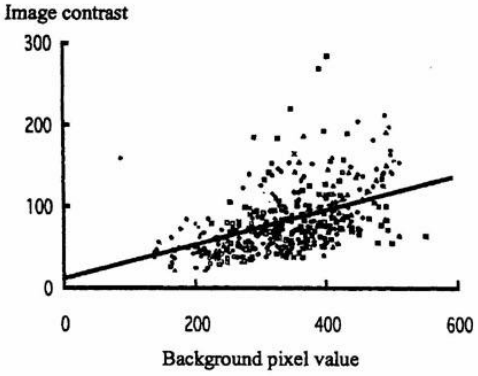


Figure 2: The approximate line was drawn according to the dispersion of contrast vs. background pixel value data.

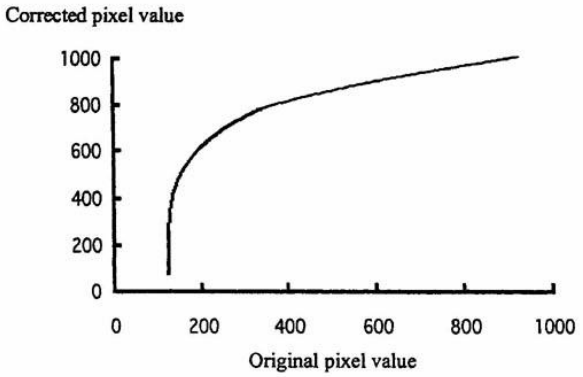


Figure 3: The contrast correction curve obtained by integral calculation of ratio *r* for the line shown in Figure 2.

The ratio r of the curve to the basic contrast s was then calculated by:

$$r = s/(\alpha x+\beta) \quad (2)$$

The approximated contrast correction curve was given by:

$$y= s(1-\alpha/2) \log(\alpha x+ \beta) +\gamma \quad (3)$$

where s = 45.0, α = 0.18719, β = 10.217 and γ was a constant obtained from the integral calculation of equation (2). Applying every pixel values in the original image to the contrast curve shown in Figure 3, we can get a background corrected image with a homogeneous contrast s in every microcalcifications.

2.6 Density gradient calculation and TRF analysis

2.6.1 Density gradient calculation

In digitized Images, generally microcalcifications absorbed more x-rays than the glandular tissues or blood vessels, so the density gradient and the magnitude value decrease greatly in microcalcifications. The magnitude of a gradient vector was less affected by the size of microcalcifications but concentrated in the central area and constructs a particular cone-shaped vector pattern as shown in Figure 4. According to this property, the microcalcification candidates can be defined as if the region with a structure close to the circular cone. Since direction and magnitude of the gradient vector were good features representing this structure [3], they were calculated by using Sobel filter in every pixels.

2.6.2 TRF analysis

In order to find out the suspected cone-shaped microcalcifications, we used a feature extracting filter called triple-ring filter (TRF) for analyzing the calculated density gradient vectors on direction and magnitude images, considering that micro-calcifications were always smaller than 10 pixels, the diameters of the three ring subfilters were selected as 3, 5 and 7 pixels, respectively (Figure 5), which correspond to various sizes of microcalcifications. Calculated vector patterns, direction and magnitude values in each subfilter were then compared with the basic vector pattern. In this procedure, two features called the "directional feature D" and the "magnitude feature I" were given by:

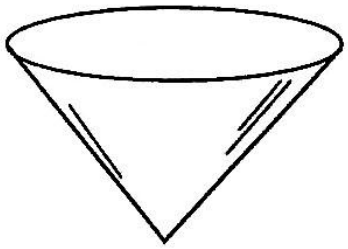


Figure 4: The basic structure of density shape of microcalcification is less affected by the size of microcalcifications, which can be expressed by the magnitude of a gradient vector in this typical shape.

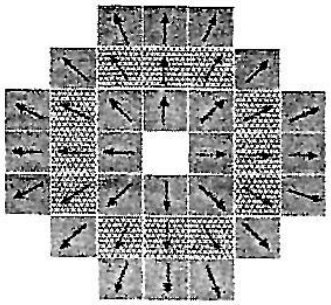


Figure 5: Vector patterns of triple-ring filter (TRF) contained three sub-ring filters: A,B and C filter with diameters of 3,5 and 7 pixels, respectively. The arrows show the vectors of a standard microcalcification.

$$D = \frac{1}{n}\sum_{k=1}^{n}(1+\delta\sin\theta_k)\cos\theta_k \qquad (4)$$

and

$$I = \frac{1}{n}\sum_{k=1}^{n}\left|\vec{V}_k\right|(1+\delta\sin\theta_k)\cos\theta_k \qquad (5)$$

where $\vec{V}_k$ represents a vector at a pixel with k th address on the subfilter, n was the number of pixels in subfilter and the difference of angle between $\vec{V}_k$ and basic vector pattern is written by $\theta_k (0 \le \theta_k \le \pi)$. δ was the factor of direction feature.

Then these features were recalculated by the combination of the three subfilters, and the threshold values for detecting candidates microcalcifications were employed for these values. Three types of the subfilter combinations were used as given below:

directional feature of (filter A + filter B)/2 = a
directional feature of (filter B + filter C)/2 = b
directional feature of (filter A + filter B + filter C)/3 = c
magnitude feature of (filter A + filter B)/2 = d
magnitude feature of (filter B + filter C)/2 = e
magnitude feature of (filter A + filter B + filter C)/3 = f (6)

In this experiment, the threshold values for a, b and c were set as 0.81, 0.81 and 0.79, respectively. Threshold values for d, e and f were equal at 30. These threshold values were determined corresponding to a vector feature which became larger when the deepest part of the microcalcification was being analysed by the triple-ring filter. Microcalcifications with the shapes varied from the basic circular cone by adjusting these features were also able to be detected.

2.7 The shift-invariant neural network

The TRF perform well in the shape-based detection of microcalcifications, however, it failed to detect some irregular

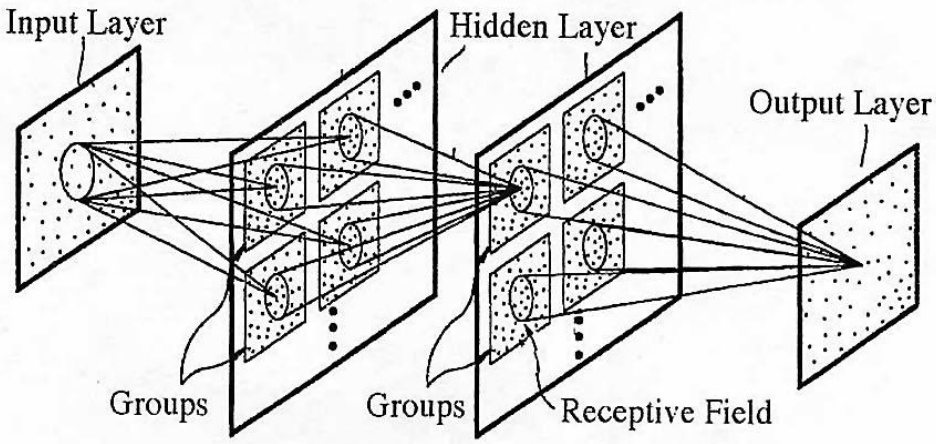


Figure 6: The structure of a shift-invariant artificial neural network (SIANN)

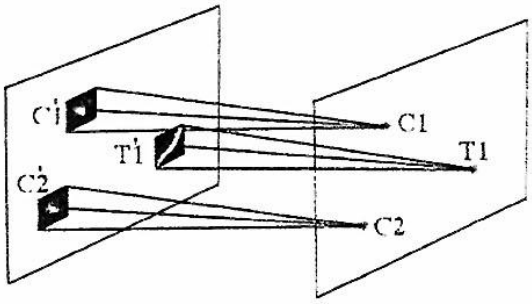


Figure 7: The interconnection between two groups in two layers. C1, C2 and T1 are disconnected each other.

candidates, which were hard to model into a predetermined mathematical algorithm. Artificial neural networks had been shown to be a powerful tool for pattern recognition and data classification. The major difference between neural networks and conventional algorithmic approaches to information processing was that the problems were not solved by use of a predetermined algorithm, but rather by "training" using examples repeatedly. A shift-invariant artificial neural network had been described by W. Zhang et al. [6], which was a layered feed-forward neural network with local interconnections as illustrated in Figure 6. The basic structure of the SIANN was similar to that of the Neocognitron model developed by Fukushima et al.[8], which was quite successfully applied to the handwriting recognition. However, the SIANN was a feed-forward neural network trained with the error-back-propagation (EBP). As shown in Figure 6, units in the input layer and output layer corresponding to pixels of the input and output were divided into groups. Every unit in a subsequent layer was connected with the units of a small region called the receptive field in every group in the preceding. In Figure 7, unit C1 in the subsequent layer receives inputs from the receptive field C1' in the preceding layer with black square illustrated. We should notice that the connections of C1 were localized. Therefore, a local variation such as C2 and T1 in the input image did not affect the outputs of the units whose receptive fields did not include the local variation, so the detection of clustered was highly independent of the quantity and position of the microcalcifications. In learning procedure, we can train the SIANN with individual microcalcification C1 and tissue T1 or with cluster C1, C2 and tissue T1 in any place of the input image, which was impossible to a conventional full-connected neural network.

A back-propagation algorithm with a generalized delta rule was modified to train the neural network with the shift-invariant-connection constraint, and the sigmoid function was used as the activation function for each processing unit in the neural network. In the training process, the internal parameters of the connections between layers were adjusted iteratively so that the difference between the output values and the desired results was minimized. The training was stopped at 6,000 iterations to avoid overtrained problem.

In this experiment we selected 150 region-of-interests (ROIs) around mass containing both of positive and negative microcalcifications for training the network. The SIANN was trained by an improved error back-propagation algorithm [6] with a 21x21-17x17-13x13 structure. Another 30 cases were selected for testing the trained ANN.

2.8 Elimination of FPs

2.8.1 Variable-ring filter

To eliminate the cross-linear pattern candidates (blood vessels or glandular tissue), a variable-ring filter based on a region growing technique was superimposed on the findings which were detected by the TRF and SIANN.

2.8.2 ANN for eliminating FP microcalcifications

Since the TRF gave the coordinate of microcalcification candidates, we can train the ANN without learning any information on location, as they were all centralized in every ROIs. Size of the ANN was limited to 11 times 11, which was small enough to learn the bright dots from background. Instead of SIANN, a conventional full connected three-layer (121 - 60 - 1) neural network was used for the purpose of eliminating FPs. The training data were 120 ROIs extracted form the

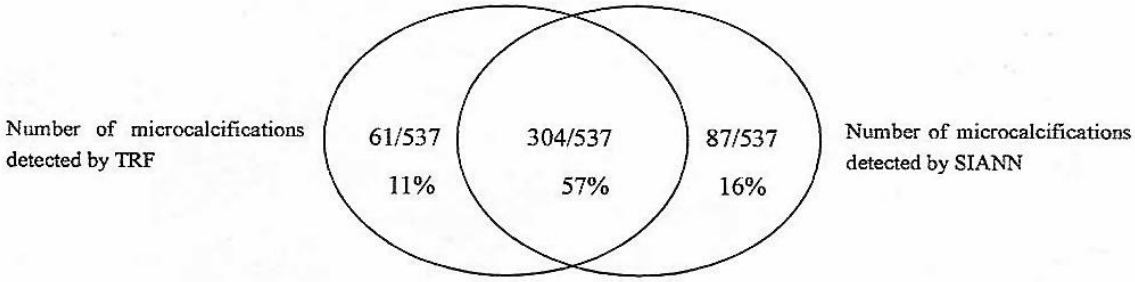


Figure 8: The sensitivity of detecting microcalcifications by triple-ring filter (TRF) and shift-invariant artificial neural network (SIANN) was 68% and 73% , respectively

output of variable-ring filter on detecting 30 mammograms. The ANN was trained by BP algorithm to produce output value of 1 corresponding to a TP input, otherwise 0 for FP input.

2.8.3 Feature analysis

The FPs still existed even after variable-ring filter were re-analyzed by contrast feature and circularity analysis. The size of a microcalcification was limited to less than 50 pixels so that big benign calcifications will not be detected.

2.9 Classification of cluster area and display

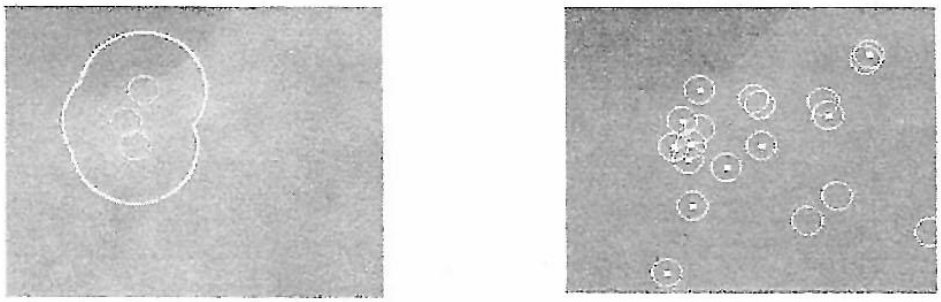
The remaining microcalcifications were then divided into individual or clustered microcalcifications. A cluster area was defined when more than three microcalcifications existed in an area of 50 mm squared. A cluster area was displayed by a cloud shape. The malignant cluster by our scheme was mainly based on the microcalcifications which were (i) very small in size, (ii) concentrated at a point, (iii) numerous, (iv) different in size, and (v) irregular in shape. The cancerous level of the clusters was shown by three different colors of the cloud shape, where red was for malignant cases, yellow for intermediate cases and blue for benign cases.

3. RESULTS AND DISCUSSION

The consistency test indicated that the neural networks were able to learn all of the patterns that were used for the training. In a practical detection of 30 cases with 40 clusters in masses, 304 out of the whole 537 true microcalcifications were detected by both methods, and the other 61 and 87 of true microcalcifications were detected by TRF and SIANN methods, respectively, with about 120 FP microcalcifications per image (Figure 8). Using both methods, 84% (452/537) of true positives could be picked out. The number of FPs was decreased to 10.2 per image while preserving 432 ture positives after using the variable-ring filter. 36 clusters were obtained by the detected microcalcifications form TRF and variable-ring filter, and two other clusters can be found by adding SIANN. Figure 9 shows an example of a missed cluster by TRF in which only a few microcalcifications in a cluster exist. Therefore, the sensitivity of detecting clusters was improved from 90% by our previous method to 95% by using both SIANN and the TRF with the same number of FPs of 0.85 per image. Figure 10 (a) shows a non-clustered ROI with VRF output only, three false-positive microcalcifications were detected as a cluster, and in (b), output of TRF were marked in circles, microcalcifications were picked up as FPs with a point in the center of circles.

Original image TRF SIANN

Figure 9: An example of a misdetected cluster in a mass area by TRF, but correctly detected by SIANN



(a) (b)

Figure 10: Comparison of the performance between VRF and ANN. A mass contained no microcalcifications was misdetected as a cluster with three FP microcalcifications (a). The missed FP detections can be eliminated by ANN (b).

The other 6 FPs can be eliminated by VRF as reference to the corresponding position in (a). By using both ANN and VRF method, FPs in the above experiment were reduced to 0.4 with keeping the same TP rate of 95%.

4. CONCLUSIONS

This study demonstrated that the SIANN method combined with our previous TRF method was able to detect 5% additional true clusters compared with the TRF method only, because it detected microcalcifications by learning samples form datasets instead of calculating by explicit mathematical model. Further efficient training has to be done for gaining the potential high generation ability of SIANN. Moreover, it was found that the conventional ANN was also an effective way in eliminating FPs in mass area.

ACKNOWLEDGMENTS

This research was performed with the assistance of members in Fujita lab. We would like to express our appreciation to the Foundation for C&C Promotion (NEC Group) for their finance support on the project.

REFERENCES

[1] J. Harris, M. Lippman, U. Veronesi, et al., "Breast cancer," *N. Engl. J. Med.* **327**, pp. 319-328, 1992.

[2] K. Doi, M. Giger, R. Nishikawa, et al:, "Computer-aided diagnosis of breast cancer on mammograms, " *Breast Cancer* **4**, pp. 228-233, 1997.

[3] K. Hirako, H. Fujita, T. Hara, et al., "Development of detection filter for microcalcifications on mammograms: A method based on density gradient and triple-ring filter analysis," *Trans. Inst. Electron. Inform. Commun. Eng.* **J78-D-II**, pp. 1334-1335, 1995.

[4] K. Hirako, H. Fujita, T. Hara, et al., "Detection scheme for clustered microcalcifications with newly introduced contrast correction technique and variable-ring filter analysis, " *Med. Imag. Tech.* **14**, pp. 665-679, 1996.

[5] K. Seki, H. Fujita, K. Hirako, et al., "Detection of microcalcifications on mammograms using neural networks," *Med. Imag. Tech.* **15**, pp. 639-651, 1997.

[6] W. Zhang, K. Doi, ML. Giger, et al., "An improved shift-invariant artificial neural network for computerized detection of clustered microcalcifications in digital mammograms," *Med. Phys.* **23**, pp. 595-601, 1996.

[7] M. Matsubara, H. Fujita, T. Hara, et al., "New algorithm for mass detection in digital mammograms," *Proc. of the 12th International Symposium and Exhibition on Computer Assisted Radiology and Surgery (CAR'98)*, pp. 219-223, 1998.

[8] K. Fukushima, S. Miyake, and T. Ito, "Neocognition: A neural network model for a mechanism of visual patter recognition," *IEEE Trans. Sys. Man and Cybernetics* **SMC-13**, pp. 826-43, 1983.