

Segmentation Mammograms based on Level Set for Detection of Breast Cancer as a Second Opinion Radiologist

Endang Supriyati¹, Tutik Khotimah¹, Mohammad Iqbal², Tri Listyorini¹, E Evanita¹
{endang.supriyati@umk.ac.id¹, tutik.khotimah@umk.ac.id², mohammad.iqbal@umk.ac.id³,
trilistyorini@umk.ac.id⁴, evanita@umk.ac.id⁵ }

¹Informatics Engineering, Universitas Muria Kudus, Kudus, Indonesia

² Electrical Engineering, Universitas Muria Kudus, Kudus, Indonesia

Abstract. Detection of signs of cancer using mammograms is a difficult job, this is due to a pathological disorder and noise structures that appear in the image. A mammogram is an X-ray image of the breast that can reveal abnormalities at an early stage. Problems in mammography screening is error a high rate. Early detection of cancer can reduce the death rate. Segmentation is used to separate objects from the background. Segmentation Mass separates the mass of the background and captures the contours of mass. This study has several stages: pre-processing, segmentation, feature extraction and classification. In improving the quality of the image in which it uses the operating CLAHE (Contrast Limited Adaptive Histogram Equalization). This is done in a pre-processing stage. Segmentation is the method of Binary and Gaussian Filtering Selective Regularized Level Set (SBGFRLS). From the test data, the average obtained by using segmentation SBGFRLS RMSE of 8.67, whereas with traditional segmentation level set at 11.4725. For the feature extraction method in this study using the Discrete Wavelet Transformation (DWT), grey Level Co-occurrence Matrix (GLCM), Gabor-Wavelet (GW). The next stage is the classification performed using the method of artificial neural network (ANN) - Lavenberg Marquard (LM). This research resulted in a classification test GW-JST, where the results are given for the better in tests using the training data. In addition, testing is also performed using the data-ANN GLCM testing, where the results of these tests are less stable. This problem is caused because the amount of data that diversification is applied on the size and structure of the mammary pathological.

Keywords: segmentation, mammography, level set, classification.

1 Introduction

Cancer is the uncontrolled growth of cells in the body. Breast cancer is the uncontrolled cell growth that occurs in the breast area. This disease becomes deadly disease for women in Indonesia. This disease is not only suffered by women. Men are also likely to suffer from breast cancer. Early detection of cancer can reduce the death rate. Early detection of breast cancer can be successfully treated 100%. Therefore, every effort should be made to detect breast cancer at an early stage. Significant problem in mammography is variability among radiologists diagnosis and methods of *Computer Aided Diagnosis* (CAD) (which consistently highlights areas on the mammogram that may need further examination) is being developed to help radiologists (Xiangying, 2011). Image segmentation is one of the stages in the process of

analysing an image, which has a goal to get information on that image. This is done by splitting the image in a separate section, where each section is homogeneous and directed to a clear criteria of uniformity. Because in the process of image segmentation is the base and is very important in computer vision. Segmenting the image should be appropriate due to the information that is on in it can be translated either. The methods carried out at this stage is a segmentation of the image, such as the detection point, line detection, and detection side. Methods ACM(Active Contour Model)with the name of Selective Binary and Gaussian Filtering Regularized Level Set (SBGFRLS)(Zang, 2010).

2 Methodology

The research method in this study as in Figure 1, include: (1) Collecting data (input image), (2) Determining the ROI (Region of Interest), (3) Pre-processing, which is a process to improve the quality of image, (4) segmentation used method SBGRLS, (5) Feature extraction using a DWT (Discrete Wavelet Transformation), GLCM (Grey Level Co-occurrence Matrix, Gabor-Wavelet (6) Classification using artificial neural networks

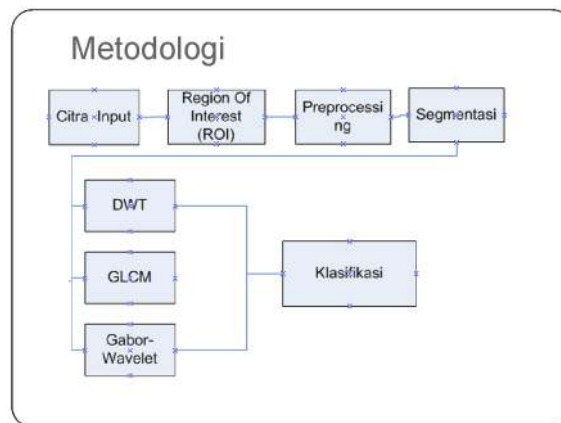


Fig. 1. Research Methodology

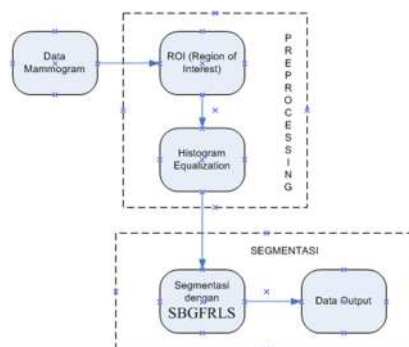


Fig. 2. Process Segmentation

2.1 Collecting Data

Data taken from MIAS (Mammographic image Analysis Society). The image is sized 1024x1024 pixels. Databases contains 322 images mammogram, comprised of three categories: normal, malignant and benign. Malignant and benign categorized as abnormal. In the abnormal data is divided into 6 categories: micro calcifications, masses (bLOB's), speculated mass (a lumps network with a point on the surface), the mass of which is not clear, architectural distortion and asymmetry. In this data has also been determined that the type of fatty breast tissue, fat granules (fatty-glandular) and density. It has also been determined the type of cancer/ tumor is benign and malignant(Mislav, n.d.).

2.2 Region of Interest

Process is to determine the area contained abnormality. Region of Interest (ROI) on the mammogram image (Figure 3) is governed by the size of 500x500 pixels.

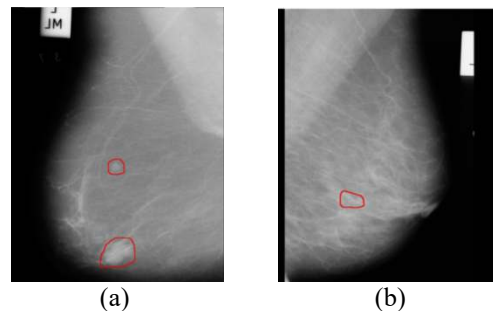


Fig. 3. (a) mammogram by mass (mdb005),
b) a mammogram with benign micro calcifications
(MDB 252)

2.3 Pre-processing

Pre-processing techniques is a step to improve the quality of the image to separate the detail that made become blurred. Or take part hallmark interesting picture. There are two categories at this stage of image enhancement techniques, namely:

1. spatial domain methods In this method operates directly on a pixel.
2. frequency domain methods

In the operate Fourier transformation method that shows on the image.

CLAHE (Contrast Limited Adaptive Histogram equalization) which is an image enhancement technique used in this study. CLAHE (Rajest, 2011) is a spatial domain method. Using CLAHE be one other way to change the histogram equalization. CLAHE operates on a small section in the image, which is often called a block. While different histogram equalization, because it works on all the existing image. The contrast value of each block are enhanced, as the histogram on the parts around which to a certain histogram. Value contrast has been improved, the next step is the equalization. How it works CLAHE is neighbouring blocks are combined using bilinear interpolation. This is done in order to artificially limit removed. In addition, CLAHE can also be used to avoid the noise in the image, by giving

restrictions on the homogeneous contrast. Figure 4 shows the process CLAHE used to improve image quality.

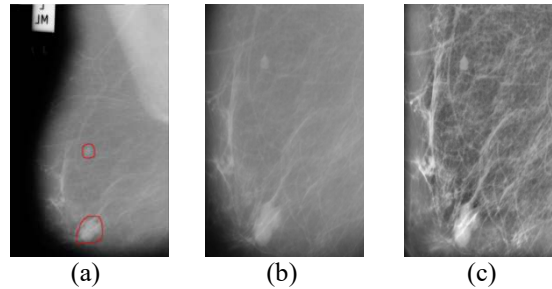


Fig. 4. (a) The ROI mass, (b) Crop mammogram, (c) CLAHE with limit Clip 0:02

2.4 Segmentation

Analysis process on the mammogram image is an important stage. Because in the process differentiating dubious part of the surroundings. In this case use the method of segmentation process used to release the Region of Interest on the usual background. The process of mass segmentation separates the mass of the background and captures the contours of mass. After the completing mass segment, shape features and the level of mass can be calculated to classify benign and malignant mass.

2.4.1 Level Set Traditional

Level traditional set is a method for detecting a movement of the curve. Figure 5 is the result of the traditional segmentation level set method.

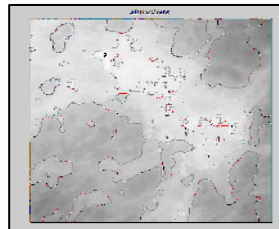


Fig. 5. Segmentation with Level Set Methods Traditional

2.4.2 Method of Selective Segmentation by Binary and Gaussian Filtering Regularized Level Set (SBGFRLS)

Selective method and Gaussian Filtering Binary Level Set Regularized change the function level set to binary, then use a Gaussian filter to set it. Figure 6 is a piece of script segmentation with SBGFRLS (Zang, 2010). The method used is the method of process selective and Gaussian Filtering Regularized Binary Level Set (SBGFRLS).

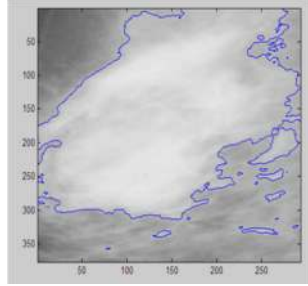


Fig. 6. Results of segmentation using SBGRS

The complete algorithm is:

1. Initialize the level set function ϕ as

$$\phi(x, t = 0) = \begin{cases} -\rho & x \in \Omega_0 - \partial \Omega_0 \\ 0 & x \in \partial \Omega_0 \\ \rho & x \in \Omega - \Omega_0 \end{cases} \quad (1)$$

where $\rho > 0$ is a constant, Ω_0 in which a subset of the domain image Ω and $\partial \Omega_0$ is the boundary of the Ω_0

2. Count $C_1(\phi)$ and $C_2(\phi)$ by equation

$$C_1(\phi) = \frac{\int_{\Omega} I(x) \cdot H(\phi) dx}{\int_{\Omega} H(\phi) dx} \quad (2)$$

where c_1 = the average intensity of the contour, c_2 = the average intensity of the outer contour, H = Heaviside function.

$$C_2(\phi) = \frac{\int_{\Omega} I(x) \cdot (1 - H(\phi)) dx}{\int_{\Omega} (1 - H(\phi)) dx} \quad (3)$$

3. Calculate Level function set with

$$\frac{\partial \phi}{\partial t} = \text{spf}(I(x)) \cdot \alpha |\nabla \phi|, \quad x \in \Omega \quad (4)$$

4. Calculate $\phi = 1$ if, $\phi > 0$ otherwise $\phi = -1$ it.

This step is necessary to obtain local segmentation. If not then this step is not necessary.

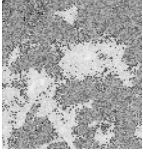

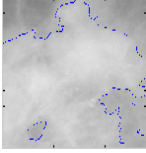
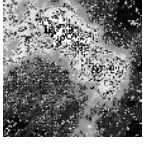

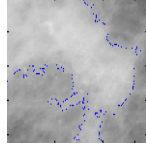
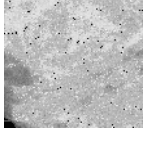

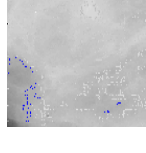
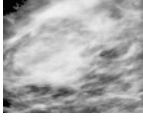

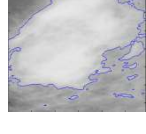
5. Use the level set function with Gaussian filter, which $\phi = \phi * G_{\sigma}$

6. Checks whether the results have been convergent level set function, if not repeat step 2.

Step 4 serves as a selective segmentation procedure, because it makes deviations $|\nabla\phi|$ remote level set function ϕ and closer to zero. Similarly, evolution has proportion local segmentation. Where it can initiate contour approaching an object of interest in order to obtain segmentation is needed. But if we want to detect all objects, stage 4 should be omitted/ deleted.

At stage five which is the standard deviation σ in the Gaussian filter is an important parameter and the election must be done properly. If the value of σ is too low, then the method performed on sensitivity to noise, and evolution to become less stable. However, if σ is too large, then what happens is the leakage edge. In addition to the detection limits enable inaccurate values.

Table 1: Results Segmentation SBGRS

Name File	Histogram	Segmentation Level Set Traditional	Level Set SBGRS
mdb265 G MISC M 593 498 60			
mdb209 G CALC M 647 503 87			
mdb241 D CALC M 453 678 38			
mdb315 D CIRC B 516 447 93			

2.5 Feature extraction

Feature extraction is used to show appropriate information to handle the computing tasks that have linkages with an application. In particular, these characteristics can refer to results of operations in the general environment (characteristic extractor/ detector characteristics) are applied to the image. In addition, a special chart in the picture, starting the simple part, for example points / edge to structure more sort of object. Trial stage in this study on feature extraction method consisting of three methods: DWT, GLCM and Gabor Wavelet.

2.5.1 DWT (Discrete Wavelet Transformation)

In DWT is divided into two parts. The second part is the low and high frequencies. And using a filter that is low pass filter and a high pass filter. Furthermore, the results of the filter taken some then carried down sampling process. In Figure 5 is the result of the calculation DWT. In a process known as decomposition (Gonzalez and Woods, 2002). To calculate the low frequency/ high 1 level set forth in the following formula:

$$y_{low}[k] = \sum_n x[n]h[2k - 1] \quad (5)$$

$$y_{high}[k] = \sum_n x[n]g[2k - n] \quad (6)$$

Where:

$y_{low}[k]$ = result of the low pass filter

$y_{high}[k]$ = result of the high pass filter

$x[n]$ = the starting signal

$g[n]$ = high pass filter

$h[n]$ = low pass filter

2.5.2 GLCM (Grey Level Co-occurrence Matrix) Order Two

Methods GLCM (grey-level co-occurrence matrix) is one way to extract the texture characteristics of the second-order statistics (Michael, 2012). The characteristics used in this study among other things: Contrast, Correlation, Variance, Different Inverse Moment and Entropy. Co-occurrence matrix is a square matrix with the number of elements as the square of the number of pixels in the image intensity level. Each point (p, q) in matrix-oriented co-occurrence provides opportunities incidence θ p-value pixels of neighbouring pixels at a distance d q-value as well as the orientation θ and $(180-\theta)$. Here is an overview GLCM formation on the image with 4 levels of grey (grey level) at a distance of $d = 1$ and the direction of 0° . Figures 5 and 6 show the description of the feature extraction calculation GLCM.

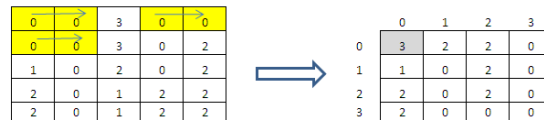


Fig. 7. (a) The image with 4 levels of grey, (b) GLCM at a distance of $d = 1$, at 0°

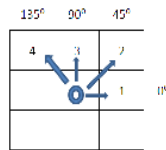


Fig. 8. pixel

Figure 8. From the pixel in the centre, pixel 1 indicates the direction $= 0^\circ$ with a distance $d = 1$, pixel 2 directions $= 45^\circ$ with $d = 1$, pixel 3 directions $= 90^\circ$ with a distance $d = 1$, from pixel 4 directions $= 135^\circ$ with a distance $d = 1$.

a. Contrast

Contrast shows deployment size (moment of inertia) elements of the image matrix. When the location is far from the main diagonal, then the greater the value of the contrast. Judging from the visual, the contrast value is a measure of the variation between the degrees of grey in a section of the image.

$$CON = \sum_k k^2 [\sum_i \sum_j p(i, j)] \quad (7)$$

Where $|i - j| = k$

b. Correlation

Correlation shows the degree of grey linear dependency size of an image. Therefore, it is used as the direction of the linear structure of an image.

$$COR = \frac{\sum_i \sum_j (i - \mu_x)(j - \mu_y) p(i, j)}{\sigma_x \sigma_y} \quad (8)$$

c. Variance

At this stage provides some elements of variance matrix co-occurrence. Small variances become a reference for the image which has a small degree of grey transition anyway.

$$VAR = \sum_i \sum_j (i - \mu_x)(j - \mu_y) p(i, j) \quad (9)$$

d. Different Inverse Moment

image with grey similar degree shown in homogeneity. In this homogeneous image has a large IDM prices.

$$IDM = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (10)$$

e. Entropy

Irregularities form shown in entropy measure. The magnitude of the price of ENT at image transitions and degrees of grey uniformly irregular structure of the image of little value / variable.

$$ENT_2 = - \sum_i \sum_j p(i, j) \cdot \log_2 p(i, j) \quad (11)$$

2.5.3 GW (Gabor Wavelet)

Gabor wavelet / GW is part of the existing wavelet sorts. Wavelet consists of two words, namely wave and let, where the meaning of these two words is a short wave. The waves here have a sense of limited duration. Or other words that contain the object is localized. Gabor Wavelet is the one used in the technique of feature extraction. Which have to extract characteristic can benefit from the normalized image. Benefits of 2-D Gabor wavelet (Rangayyan M, 2007) is a spatial filter undercooling maximum field in far less important trait in the spatial scope and frequency.

Gabor wavelet functions used are defined as:

$$h(x, y) = \frac{1}{2\pi} \exp \left[-\alpha^2 j \frac{x^2 + y^2}{2} \right] \exp [j\pi\alpha(x \cos \theta + y \sin \theta)] \quad (9)$$

with $\alpha = \frac{1}{\sqrt{2}}$, $j = 0, 1, 2 \dots$ dan $\theta \in [0, 2\pi]$

where σ is the variance of the Gaussian distribution both in the x and y directions, ω_0 is the frequency sinusoidal and θ is the direction sinusoidal. Actually the basic functions of a 2D Gaussian envelope Gabor is modulated with frequency ω_0 and orientation θ .

2.6 Artificial Neural Network - Lavenberg Marquard

Backpropagation algorithm (Cemil et al., 2007) was used as a training model of ANN (Artificial Neural Network). Training the neural network calculates the weight to the output in

response to input up to the limit specified error. Input value and the target referred to as a training partner. Backpropagation algorithm is as follows:

- a. Select a training partner (input and target), the input value of the input on the network.
 - b. Calculate the output network
 - c. compare the output value to the target and count error
 - d. Recalculate weights to reduce the error
- above steps (1-4) is repeated until the specified error limit.

In this study, a multi-layer Perceptron use and trained with the Lavenberg-Marquardt algorithm (LM) (Rachmad S and Mauridhi, 2006). Algorithm (LM) is a development Backpropagation algorithm (BP) standard. In the standard BP algorithm weights and biases of the update process using a negative gradient descent directly whereas the algorithm Lavenberg-Marquardt using matrix approach Hessian. Lavenberg-Marquardt algorithm is an enhancement develop standard back propagation algorithm. In the back propagation algorithm, the update process using the weights and biases negative gradient descent directly while algorithm Lavenberg-Marquardt approach Hessian matrix (H) which can be calculated:

$$H = J^T e \quad (12)$$

while the gradient can be calculating with,

$$g = J^T J \quad (13)$$

In this case J is a Jacobian matrix containing the first derivative of network error to the weights and biases of the network. Changes weighting can be calculated,

$$\Delta X = [J^T J + \mu I]^{-1} J^T e \quad (14)$$

so that corrective weighting can be determined by

$$X = X + \Delta X \quad (15)$$

$$X = \Delta X + [J^T J + \mu I]^{-1} J^T e \quad (16)$$

X = function-network weight weights and biases

e is a vector or an error on the network output

$e = [t_1 - y_1, t_2 - y_2, \dots, t_p - y_p]^T$

μ = constant learning

I = identity matrix

3 Result Discussion

Amounts to 18 pieces of training data used in this test. Of 18 was divided into several parts, namely 6 masses of data mammogram, 6 pieces of data is the data mammogram benign and 6 pieces of data are malignant mammogram. The usefulness of these data are applied to this test as many as 12 pieces of data. From these data 12 divided into 4 data for mass mammogram, 4 pieces of data for benign mammogram and 4 pieces of data for malignant mammogram.

- a. RMSE (Root Mean Squared Error)

Root Mean Squared Error is used to determine how the error rate is generated. The smaller the RMSE value, the better the results

Table 2: RMSE

Name File	Level	Set	SBGRSL
	Traditional		
mdb265 G MISC	12.16		8.90
M 593 498 60			
mdb209 G CALC	12.26		10.72
M 647 503 87			
mdb241 D CALC	12.00		5.01
M 453 678 38			
mdb315 D CIRC	9.47		10.05
B 516 447 93			

b. Classification

Test this would result in three classes, namely mass, benign and malignant. Training method used is the Lavenberg-Marquard.

1. Trial Extraction Feature GLCM with JST

Setting artificial neural networks this experiment can be seen in Table 3:

Table 3: Results from Pilot to Feature GLCM Order2

No	Accuracy Training	Accuracy Testing
1	61%	41%
2	61%	41%
3	50%	25%
4	39%	42%

Value *Contrast*, *Correlation*, *Variance*, *Different Inverse Moment* and *Entropy*, amounting to 5 is at the input node. Where all these values are values of feature extraction. Value 18 is the number of training data. Node which lies in the hidden layer are 20, but there is no determination of the number of nodes specifically. Logsig function is used for the activation function of hidden layer and output layer. Where the output layer has three nodes that mass class, benign and malignant.

2. Trial Extraction Feature DWT with JST

Setting artificial neural networks this experiment can be seen in Table 4:

Table 4: Results from Pilot to Feature DWT

No	Accuracy Training	Accuracy Testing
1	100%	33%
2	94%	33%
3	89%	33%
4	67%	33%

Node inputs are numbered 32 represents the value of feature extraction. Value 18 is the number of training data. Node which lies in the hidden layer are 20, but there is no determination of the number of nodes specifically. Logsig function is used for the activation function of hidden layer and output layer. Where the output layer has three nodes that mass class, benign and malignant.

3. Feature Extraction TestGW by ANN

Settings artificial neural networks this experiment can be seen in table 5.

Table 5. Results of Trial to FeatureGW

No	Accuracy Training	Accuracy Testing
1	100%	33%
2	100%	33%
3	100%	33%
4	72%	33%

Node inputs are numbered 48 represents the value of feature extraction. Value 18 is the number of training data. Node which lies in the hidden layer are 20, but there is no determination of the number of nodes specifically. Logsig function is used for the activation function of hidden layer and output layer. Where the output layer has three nodes that mass class, benign and malignant.

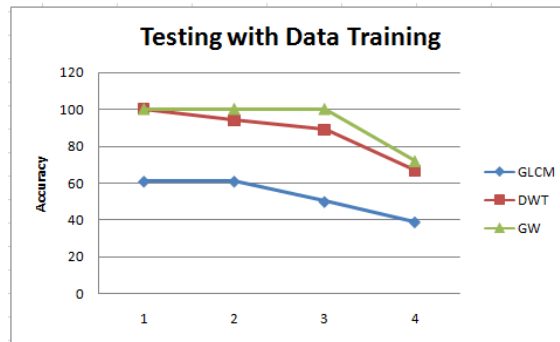


Fig. 9. The results of testing with data training

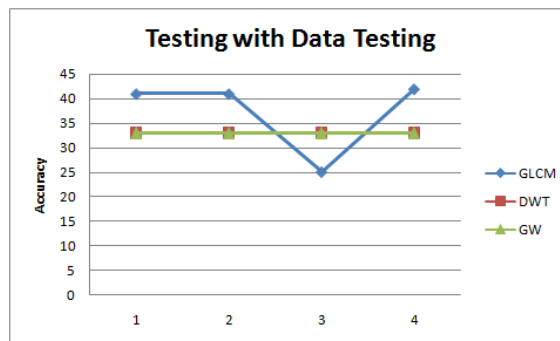


Fig. 10. The results of testing with the data testing

Figure 9 and 10 shows the results of testing the accuracy of training data and testing data testing. From the graph 6. (a) can be seen best results bail is Gabor Wavelet (GW). In the graph 6. (b) shows the results unstable. This is because for two things: first, the diversification of test data that is great, this is due to the data mammograms taken from patients of different, where each patient has the size and structure of different pathological, the classification used is a neural network Knockoffs which has several drawbacks which are trapped in a local minimum.

4 Conclusions

For the with segmentation process uses 4 data for testing RMSE values obtained from traditional segmentation using level set by an average of 11.4725 while SBGRSL segmentation with an average of 8.67. Segmentation by SBGRSL better than the traditional level set segmentation of the three extraction features that have been described, the feature extraction Gabor wavelet have accuracy values that are less good training. Results of testing accuracy no results were striking; this situation is quite high due to the data verified mammogram. In which the data is taken from several different patients. Therefore, different sizes and pathological structures mammogram. The result of better classification required optimization feature extraction and measurement of tumour area.

Acknowledgements

Thanks to those who helped in this research. To the Muria Kudus University Research Institute which has supported the research on mammogram segmentation. And also we are grateful to Ministry of Research, Technology and Higher Education for supporting the research on self-extraction in this article.

References

- [1] Cemil, O., Ming, C., Leu, 2007. Linguistic properties based on American Sign Language isolated word recognition with the sense official neural networks using a sensory glove and motion tracker. *Neuro computing* 70.
- [2] Gonzale, R., Woods, R., 2002. *Digital Image Processing*, 2nd ed. Prentice Hall, New Jersey.
- [3] Michael, B., 2012. *Algorithms*.
- [4] Mislav, G., n.d. *Mammographics Image Analysis Homepage*.
- [5] Rachmad S, G., Mauridhi, H., 2006. Levenberg-Marquardt algorithm Comparison With Method Backpropagation Learning process Neural Networks for Pattern Recognition Signal Electrocardiograph, in: SNATI. Presented at the National Seminar on Information Technology Application, Universitas Islam Indonesia, Yogyakarta.
- [6] Rajest, G., 2011. Histogram Equalization Techniques for Image Enhancement. *IJECT* 2.
- [7] Rangayyan M, R., 2007. A review of computer-aided diagnosis of breast cancer. *Journal of the Franklin Institute* 312–348.
- [8] Xiangying, W., 2011. An Interactive Segmentation Method Using Graph Cuts for mammographic Masses. *IEEE*.
- [9] Zang, K., 2010. Active Contour with Selective Local or Global Segmentation: A New Formulation and Level Set Method. *Image and Vision Computing* 28 668–676.