MAMMOGRAMS CLASSIFICATION USING FEATURES VECTOR AND NAÏVE BAYES CLASSIFIER

NG Angries[1], Diaz D. Santika[2]
Magister Teknik Informatika
Universitas Bina Nusantara
Jl. Kebon Jeruk Raya No. 27, Jakarta Barat, 11530, Indonesia
Telp.: 08999941208,e-mail: [1]griezst@yahoo.com, [2]ddsantika@gmail.com2

Abstract

Choosing the best features is very important in order to get better result in accuracy when we classify mammogram images into normal, abnormal, benign, or malignant categories. This study is aimed to evaluate various set of features such as GLCM, GLRLM, and Chip Histogram in order to find the best possible features combination. The classifying process is performed by Naïve Bayes classifier and facilitated by WEKA. The experiment results show that the combination of entropy, energy, contrast, sum average, variance, correlation, maximum probability, inverse difference moment, cluster shade features applied to enhanced mammogram images can give best result (78%) in accuracy in normal/abnormal classifications. Further study, using the same combination features, on abnormal mammogram images shows that the maximum accuracy that can be obtained is only 68% for benign/malignant classifications.

Keyword: *Features, Mammograms, Naïve Bayes classifier, WEKA*

1. INTRODUCTION

Breast cancer is the most common cancer for women. Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer death among females, accounting for 23% (1.38 million) of the total cancer cases and 14% (458,400) of the cancer deaths. The factors that contribute to the international variation in incidence rates largely stem from differences in reproductive and hormonal factors and the availability of early detection services [1]. Scientific studies have shown that the mortality in breast cancer is decreased by early detection and treatment. Retrospective studies have shown that in current breast cancer screening between 10% and 25% of the tumors are missed by radiologists [2]. Mammographic screening allows early detection of non-palpable, non-invasive and early invasive tumors. Hence, it can reduce the mortality from breast cancer by 20-30% [3]. Thfere is an increasing need for automatic and accurate detection of cancer cells. However, the low contrast between the breast cancer cells and normal cells increases the difficulty of early detection [4].

Because of the difficulty of early detection, many studies proposed methods to help early detection cancers by training system model using computer. For training supervised system models, we need extracted features vector that are trained into system. So the most important thing for classification is the best combination features vector which can discriminating normal, abnormal, benign, and malignant categories in mammograms. But finding combination of features vector is not easy because of the low contrast between the breast cancer cells and normal cells.

2. METHODOLOGY

We use the MIAS MiniMammographic Database. The size of all the images is 1024 pixels x 1024 pixels. The images have been centered in the matrix. When calcifications are present, centre locations and radii apply to clusters rather than individual calcifications. Coordinate system origin is the bottom-left corner. In some cases calcifications are widely distributed throughout the image rather than concentrated at a single site. In these cases centre locations and radii are inappropriate and have been omitted. Total sample images which are used for experiment are 292 images: 198 normal images, 52 abnormal-benign images, 42 abnormal-malignant images.

The extension of images in MIAS Database are *.pgm.

Figure 1. (a) Right Side Normal Breast Mammogram; (b) Left Side Normal Breast Mammogram

Severity of abnormality:
B – Benign
M – Malignant

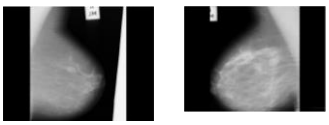

Figure 2. (a) Left Side Benign Breast Mammogram; (b) Right Side Benign Breast Mammogram

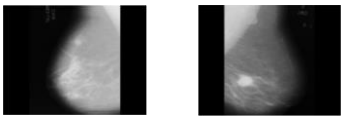

Figure 3. (a) Right Side Malignant Breast Mammogram; (b) Left Side Malignant Breast Mammogram

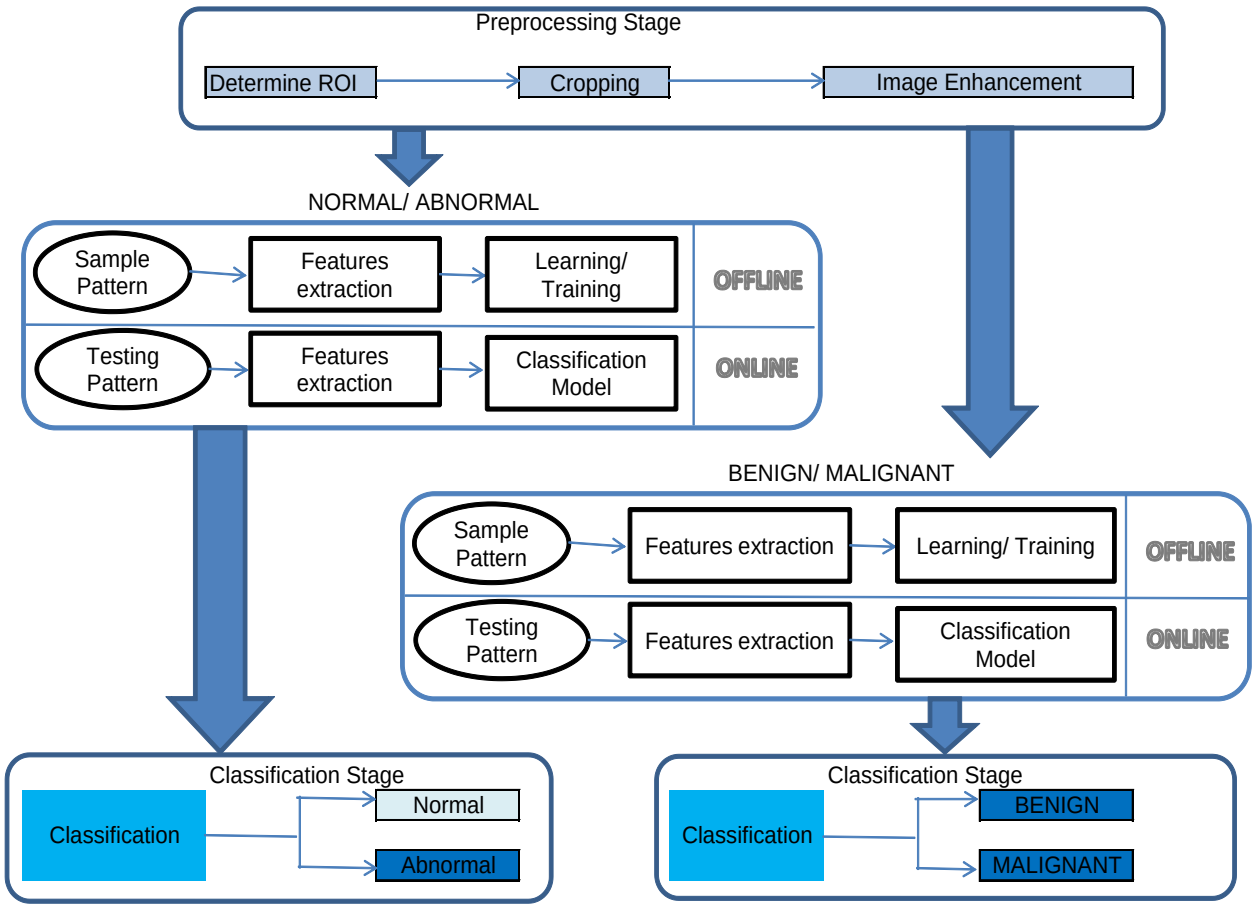

Figure 4. Methodology

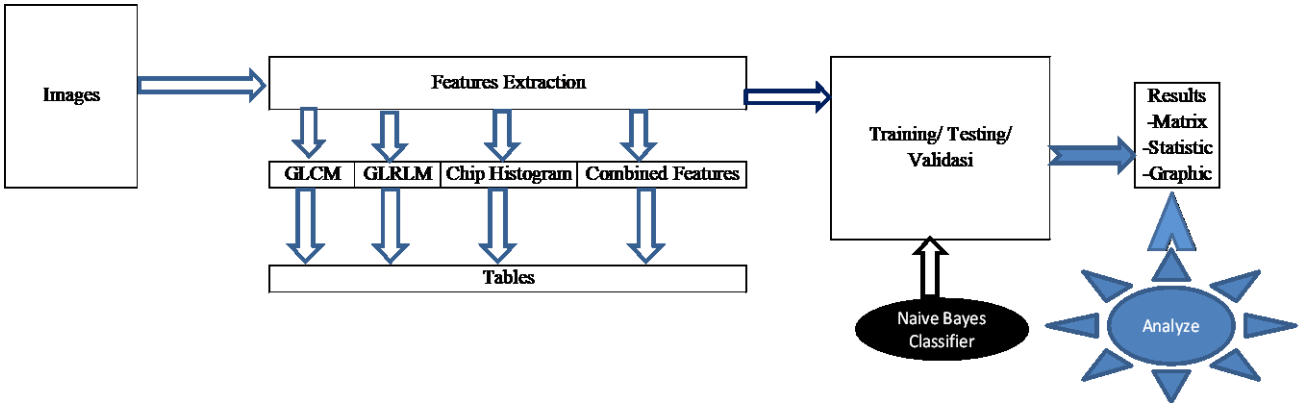

Figure 5. Model and Analysis Method

This study is done by following steps:

1. Preprocessing stage

All images data from MIAS database that will be used in this study before being processed is large, which is 1024 X 1024. Many parts that do not show valid information for research. Specific area of the image wa s not showing the information that we want investigated. The picture quality is not necessarily good. All of these things can make data being unvalid. Therefore, we must prepare the images before the extraction process to obtain the research results. In this stage, we determine Region of Interest (ROI), do cropping, and do enhancing images (option).

2. Features extraction from images

In features extraction stage, we use Grey Level Co-occurance Matrices (GLCM): autocorrelation, contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, sum of sqaures, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation1, information measure of correlation2, inverse difference normalized, inverse difference moment normalized; Grey Level Run Length Matrices (GLRLM): SRE, LRE, GLN, RP, RLN, LGRE, HGRE; Chip Histogram: mean, variance, skewness, kurtosis, energy. Features vector extracted from images with multiple methods to analyze the features of the strongest and most dominant as descriptors for the classification of normal/ abnormal and benign/ malignant categories. For extraction the features from images, we use Matlab tools and Microsoft Excel.

3. Training/ testing/ validation system

We choose and combine various of features from features extraction. Then we train/test/validate system model using those combination features. The classifying process is performed by Naïve Bayes classifier and facilitated by WEKA. Weka (Waikato Environment for Knowledge Analysis) is a machine learning software written in Java, developed at the University of Waikato, New Zealand. Weka is a free software under the GNU general public license. Weka contains a collection of visualization tools and algorithms for data analysis and predictive models, along with a display face for easy user access to the functional (Bouckaert, dkk, 2010). There are normal/abnormal system models and benign/malignant system models in this study. We use 10 folds cross-validation for validate data.

4. Classification stage

This study is aimed to classify mammograms into normal/ abnormal categories. From abnormal mammograms, we classify them into benign/ malignant categories. We analyze the accuracy every set of combination features.

5. Measurement Results

We use confusion matrix for describing the accuracy of classification, in which the matrix can be seen from the results of the percentage of overall accuracy of the model. The higher percentage of accuracy of training and validation, the better combination of features vector that we use as discriminator of classification.

Model and analysis methods that we use are descriptive: model such as table, graphic, diagram, and the explanation and analytic: estimation, statistic testing, and so on. The variables in this study are more ordinal or qualitative. Global variables in this study is a variable features vector mammogram. Attributes used in this study were GLCM feature vectors, GLRLM, Chip Histogram,

and a combination of some vectors of these features.

We apply the composition of training abount 50% of the population data, validation approximately 30% of the population data, and testing about 20% of the population data. So, suppose that the data population numbered 100, then 50 data will be used as training data, 30 data will be used as data validation, and 20 data will be used as trial data. However, we should pay attention to elements of the balance of the composition. When classifying mammograms into normal and abnormal categories, the number of data samples of a normal mammograms should be equal to the number of data samples of abnormal mammograms. The case also when classifying mammograms into the category of abnormal benign and malignant, amount of benign data samples must be equal to the number of malignant data samples.

3. DISCUSSION

In recent years, several studies for classification mammograms have been published. They proposed their own methods to improve quality of classification mammograms. As mammograms are difficult to interpret, preprocessing is necessary to improve the quality of image and make the feature extraction phase as an easier and reliable one [5].

Sivaramakrishna, et al [1] compared the performance of mammographic enhancement algorithms. For microcalcifications, the adaptive neighborhood contrast enhancement algorithm was the most preferred in 49% of the interpretations, the wavelet-based enhancement in 28%, and the unenhanced image in 13%. For masses, the unenhanced image was the most preferred in 58% of cases, followed by the unsharp masking algorithm (28%).

Pereira, et al. [18] developed a CBIR system for mammograms to aid the diagnosis of breast lesions in 5518 images of regions of interest, which were obtained from the Digital Database for Screening Mammography that included microcalcifications, masses, and normal cases. Sixteen texture features were used, 13 were based on the spatial gray-level dependence matrix and 3 on the wavelet transform. The results obtained from receiver operating characteristic analysis indicated that the texture features can be used for separating normal regions and lesions with masses and microcalcifications. However, the texture features were not very effective for distinguishing between malignant and benign lesions. The study showed that the texture features can be used for the detection of suspicious regions in mammograms.

GLCM features of the difference entropy, local homogeneity, and the differences and GLRLM features of SRE, LRE are used by Karahaliou (Karahaliou, et al, 2007). Meanwhile, GLCM features used by Chia Hung are the energy, correlation, sum variance, and difference entropy (Wei, et al, 2006). In addition, Felipe using GLCM features of entropy, energy, contrast, sum average, variance, correlation, maximum probability, and inverse difference moment (Felipe, 2003). Then, Nithya using GLCM features contrast, cluster shade, dissimilarity, difference entropy, and information correlation1.

Chia-Hung Wei, et al [6] used methodology which is divided into two parts-image analysis and image retrieval. In the image analysis part, 19 abnormal regions of interest (ROI) and 20 normal ROIs are selected as samples for the whole ROI dataset. These two groups of ROIs are used to analyze 11 textural features based on gray level co-occurrence matrices. A maximum precision of 51% and recall of 19% were obtained using the gray level co-occurrence matrices. The averages of precision and recall are 49% and 18% in this experiment.

4. RESULTS

Table 1. Validation and Testing Normal-Abnormal Mammograms Results

Features Vector	Type	Enhancement Method	Description	Confusion Matrix	Correctly Classified Instances	Incorrectly Classified Instances
Entropy Energy	Training and cross	No	100 data (50 normal and 50	a b <-- classified as	75%	25%

Contrast Sum Average Variance Correlation Maximum Probability Inverse Difference Moment Cluster Shade SRE1	validation 10 folds		abnormal data)	39 11 | a = NORMAL 14 36 | b = ABNORMAL		
Entropy Energy Contrast Sum Average Variance Correlation Maximum Probability Inverse Difference Moment Cluster Shade SRE1	Validation	No	60 data	a b <-- classified as 39 11 | a = NORMAL 14 36 | b = ABNORMAL	71.67%	28.33%
Entropy Energy Contrast Sum Average Variance Correlation Maximum Probability Inverse Difference Moment Cluster Shade SRE1	Testing	No	40 data	a b <-- classified as 39 11 | a = NORMAL 14 36 | b = ABNORMAL	60%	40%
Entropy Energy Contrast Sum Average Variance Correlation Maximum Probability Inverse Difference Moment Cluster Shade	Training and cross validation 10 folds	Adaphisteq	100 data (50 normal and 50 abnormal data)	a b <-- classified as 43 7 | a = NORMAL 15 35 | b = ABNORMAL	78%	22%
Entropy Energy	Validation	Adaphisteq	60 data	a b <-- classified as	86.67%	13.33%

Contrast Sum Average Variance Correlation Maximum Probability Inverse Difference Moment Cluster Shade				25 2 | a = NORMAL 3 27 | b = ABNORMAL		
Entropy Energy Contrast Sum Average Variance Correlation Maximum Probability Inverse Difference Moment Cluster Shade	Testing	Adaphisteq	40 data	a b <-- classified as 14 6| a = NORMAL 0 20 | b = ABNORMAL	85%	15%
Difference Entropy Local Homogeneity Difference Variance	Training and cross validation 10 folds	Adaphisteq-Adjust	100 data (50 normal and 50 abnormal data)	a b <-- classified as 40 10 | a = NORMAL 12 38 | b = ABNORMAL	78%	22%
Difference Entropy Local Homogeneity Difference Variance	Validation	Adaphisteq-Adjust	60 data	a b <-- classified as 21 9 | a = NORMAL 15 15 | b = ABNORMAL	60%	40%
Difference Entropy Local Homogeneity Difference Variance	Testing	Adaphisteq-Adjust	40 data	a b <-- classified as 17 3 | a = NORMAL 4 16 | b = ABNORMAL	82.5%	17.5%

Table 2. Validation and Testing Benign-Malignant Mammograms Results

Features Vector	Type	Enhancement Method	Description	Confusion Matrix	Correctly Classified Instances	Incorrectly Classified Instances

Entropy Energy Contrast Sum Average Variance Correlation Maximum Probability Inverse Difference Moment Cluster Shade	Training and cross validation 10 folds	No	50 data (25 Benign data dan 25 Malignant data)	a b <-- classified as 11 14 | a=BENIGN 5 20 | b=MALIGNANT	62%	38%
Entropy Energy Contrast Sum Average Variance Correlation Maximum Probability Inverse Difference Moment Cluster Shade	Validation	No	30 data	a b <-- classified as 8 7 | a=BENIGN 1 14 | b=MALIGNANT	73.33%	26.67%
Entropy Energy Contrast Sum Average Variance Correlation Maximum Probability Inverse Difference Moment Cluster Shade	Testing	No	20 data	a b <-- classified as 3 7 | a=BENIGN 3 7 | b=MALIGNANT	50%	50%
Difference Entropy Local Homogeneity Difference Variance	Training and cross validation 10 folds	Adaphisteq	50 data (25 Benign data dan 25 Malignant data)	a b <-- classified as 15 10 | a=BENIGN 12 13 | b=MALIGNANT	56%	44%
Difference Entropy Local Homogeneity Difference Variance	Validation	Adaphisteq	30 data	a b <-- classified as 10 5 | a=BENIGN 5 10 | b=MALIGNANT	66.67%	33.33%
Difference Entropy Local Homogeneity Difference Variance	Testing	Adaphisteq	20 data	a b <-- classified as 11 14 | a=BENIGN 5 20 | b=MALIGNANT	75%	25%
Difference Entropy Local Homogeneity Difference Variance	Training and cross validation 10 folds	Adaphisteq-Adjust	50 data (25 Benign data dan 25 Malignant data)	a b <-- classified as 15 10 | a=BENIGN 12 13 | b=MALIGNANT	56%	44%
Difference Entropy Local Homogeneity Difference Variance	Validation	Adaphisteq-Adjust	30 data	a b <-- classified as 4 11 | a=BENIGN 2 3 | b=MALIGNANT	56.67%	43.33%
Difference Entropy	Testing	Adaphisteq-	20 data	a b <-- classified as	55%	45%

Local Homogeneity Difference Variance		Adjust		1 9 | a=BENIGN 0 10 | b=MALIGNANT		
Contrast Cluster Shade Dissimilarity Difference Entropy Information measure of correlation1	Training and cross validation 10 folds	Adaphisteq-Adjust	50 data (25 Benign data dan 25 Malignant data)	a b <-- classified as 10 15 | a=BENIGN 7 18 | b=MALIGNANT	56%	44%
Contrast Cluster Shade Dissimilarity Difference Entropy Information measure of correlation1	Validation	Adaphisteq-Adjust	30 data	a b <-- classified as 6 9 | a=BENIGN 3 12 | b=MALIGNANT	60%	40%
Contrast Cluster Shade Dissimilarity Difference Entropy Information measure of correlation1	Testing	Adaphisteq-Adjust	20 data	a b <-- classified as 6 4 | a=BENIGN 2 8 | b=MALIGNANT	70%	30%

On mammogram classification into the category of normal/abnormal, testing and validation results showed an increase in the accuracy of the training and validation 10 folds from a combination of features of entropy, energy, contrast, sum average, variance, correlation, maximum probability, inverse difference moment, cluster shade and using images enhancement with adaphisteq method. The results of the training and validation 10 folds with 78% accuracy rate in increased to 86.67% with 60 samples of data and 85% with 40 different data samples.

On mammogram classification into categories developed benign/malignant, testing and validation results showed an increase in the accuracy of the training and validation 10 folds from a combination of features of entropy, energy, contrast, sum average, variance, correlation, maximum probability, inverse difference moment, cluster shade and using unenhanced images. The results of the training and validation 10 folds with 62% accuracy rate increased to 73.33% with 60 samples of data, but decreased to 50% accuracy rate with 20 sample data.

5. CONCLUSION

With a selection of different feature vectors, or by combining multiple feature vectors of several methods, this study have shown varied results. With a different image enhancement methods, the study also showed varied results. The experiment results show that the combination of entropy, energy, contrast, sum average, variance, correlation, maximum probability, inverse difference moment, cluster shade features and using adaphisteq-adjust enhanced images can give best result (78%) in accuracy in normal/ abnormal classifying. Further study, on abnormal mammogram images, show that better accuracy (68%) can be obtained through the combination of same features as normal/ abnormal classification, but with unenhanced images in benign/ malignant classifying.

REFERENCES

[1] Sivaramakrishna, Radhika, et al. (2000). Comparing the Performance of Mammographic Enhancement Algorithms. AJR July 2000 vol. 175 no. 1 45-51. Retrieved Dec 9, 2011 from http://www.ajronline.org/content/175/1/45.full.pdf

[2] N. Riyahi Alam, et al. (2009). Computer-Aided Mass Detection on Digitized Mammograms using a Novel Hybrid Segmentation System. International Journal of Biology and Biomedical Engineering, Issue 4, Volume 3, 2009. Retrieved in Nov 5, 2011 from http://www.naun.org/journals/bio/19-443.pdf

[3] Karahaliou, A., et al. (2007). Texture analysis of tissue surrounding microcalcifications on mammograms for breast cancer diagnosis. The British Journal of Radiology, 80 (2007), 648–656. Retrieved Nov 30, 2011 from

http://bjr.birjournals.org/cgi/reprint/80/956/648.pdf
[4] AbuBaker, Ayman A., et al. (2006). Efficient Pre-processing of USF and MIAS Mammogram Images. Journal of Computer Science 3 (2), 67-75, ISSN 1549-3636.
[5] Mohanty, Kumar Aswini, et al. (2010). Efficient Image Mining Technique for Classification of Mammograms to Detect Breast Cancer. Special Issue of IJCCT Vol. 2 Issue 2, 3, 4; 2010 for International Conference [ICCT-2010]. Retrieved Nov 17, 2011 from http://interscience.in/SpIss_ijcct_icct2010vol2_no234/19_ML87.pdf
[6] Wei, Chia-Hung, Chang-Tsun Li, and Roland Wilson. A General Framework for Content-Based Medical Image Retrieval with its Application to Mammograms. Department of Computer Science University of Warwick, Coventry, CV4 7AL, UK. Retrieved Oct 29, 2011 from www.dcs.warwick.ac.uk/research/improc/pubs/Wei2005b.PDF