ABSTRACT

Breast cancer in women is among the top five deadliest types of cancer in thirteen of the fifteen countries of Asia. Breast cancer is the most frequently diagnosed cancer and the leading cause of death by cancer in women in the ASEAN region, as well as throughout the world. Early detection of cancer in breast examination is important to assure that someone still has normal breast. If there are abnormalities, they can be found early. The possibility of recovery of breast cancer which is treated at an early stage is approaching 95%. One technique for early detection of breast cancer is mammography. Mammography is x-ray examination with a low dose x-ray. In examining the results of mammography image, which is called mamogram, the overall parenchyma pattern of the left and right breast placed side by side for symmetry assessed breast tissue left and right by a radiologist. Thus, in building Computer Aided Diagnosis (CAD) system for screening mammography, it is necessary to adapt the working procedure of the radiologist.

In this study, 60 images from Oncology Clinic Kotabaru Yogyakarta were used, which was consisted of 30 training images and 30 testing images. The first step was enhancing the image quality with the *Median Filter* operation and contrast limited adaptive histogram equalization (CLAHE). Then, feature extraction was processed by histogram-based and by Gray Level Co-Occurrence Matrix (GLCM) based. Furthermore, the similarity measurement process was used to measure the difference value between features on left and right mamogram image. This process was intended to appraise symmetry of left and right mamogram image as radiologists did in screening mammography. From the 11 features extracted, feature selection was done using Weka software to look for features that produce the best classifier performance when used as input. Three methods were tested, namely Best First, Random Search, and Ranker. Using the Ranker method, there were five features used for classifier inputs and produced best outputs if compared with Best First method and Random search method namely, IDM, ASM, Energy, contrast, and entropy-based GLCM. The results of the classification between normal and abnormal images with Error Back Propapagation (EBP) algorithm were: accuracy 93.3%, sensitivity 100%, and specificity 83.3%.

Keywords : Screening Mammography, Mammogram, CAD, Feature Extracion, GLCM, Histogram, Similarity Measurement, EBP

INTISARI

Kanker payudara pada wanita berada di antara kelima jenis kanker penyebab kematian tertinggi di tiga belas dari lima belas negara Asia. Kanker payudara adalah kanker yang paling sering didiagnosis dan menjadi penyebab utama kematian akibat kanker pada wanita di kawasan ASEAN, serta di seluruh dunia. Pemeriksaan deteksi dini pada payudara menjadi penting karena hal ini berguna untuk memastikan bahwa payudara seseorang masih normal. Bila ada kelainan dapat ditemukan lebih awal. Kanker payudara yang diobati pada stadium dini kemungkinan sembuh mendekati 95%. Salah satu teknik deteksi dini kanker payudara adalah dengan mamografi. Mamografi adalah pemeriksaan rontgen dengan dosis x-ray rendah. Dalam pemeriksaan citra hasil mamografi, yaitu mamogram, keseluruhan pola parenkim payudara kiri dan kanan diletakkan berdampingan untuk dinilai simetri jaringan payudara kiri dan kanan oleh ahli radiologi. Sehingga, dalam membangun sistem *Computer Aided Diagnosis* (CAD) untuk skrining mamografi, perlu diperhatikan prosedur yang mengadaptasi cara kerja radiolog.

Pada penelitian ini digunakan 60 citra mamogram dari Klinik Onkologi Kotabaru Yogyakarta, yang terdiri dari 30 citra latih dan 30 citra uji. Pertama dilakukan peningkatan kualitas citradengan operasi *Median Filter*dan *contrast limited adaptive histogram equalization* (CLAHE). Kemudian dilakukan proses ekstraksi fitur berbasis histogram dan *Gray Level Co-Occurence Matrix* (GLCM). Selanjutnya, dilakukan *similarity measurement* yaitu tahapan mengukur perbedaan nilai fitur-fitur antara citra mamogram kiri dan kanan. Tahapan ini dilakukan untuk menilai simetri mamogram kiri dan mamogram kanan seperti yang dilakukan radiolog dalam skrining mamografi. Dari 11 fitur yang terekstraksi, dilakukan seleksi fitur dengan menggunakan software Weka untuk mencari fitur-fitur yang menghasilkan kinerja *classifier* terbaik ketika dijadikan input. Tiga metode pencarian yang dicobakan yaitu *Best First, Random Search,* dan *Ranker.* Hasilnya dengan menggunakan metode *Ranker* terdapat lima fitur yang digunakan untuk input *classifier* dan menghasilkan output paling baik dibandingkan metode *Best First* dan *Random Search*, yaitu IDM, ASM, Energi, Kontras, dan Entropi berbasis GLCM. Hasil klasifikasi antara citra normal dan abnormal dengan algoritma *Error Back Propapagation* (EBP) adalah akurasi 93,3%, sensitifitas 100%, dan spesifisitas 83,3%.

Kata Kunci – Skrining Mamografi, Mamogram, CAD, Ekstraksi Fitur, GLCM, Histogram, *Similarity Measurement,* EBP