

## ABSTRACT

Breast cancer is the number one killer of women in the world. By the early detection of breast cancer the life expectancy for the patient can be improved about 95%. In the United States the early detection of breast cancer patients were able to save the lives approximately 12 to 37 people per day. While in Indonesia, according to the health profile of the Ministry of Health of the Republic of Indonesia Year 2012 the highest cancer which is suffered by Indonesian women was breast cancer with the incidence rate of 2,2% per 1000 women. If this condition is not controlled, in 2030 there will be 26 million people suffering of breast cancer and 17 million people will die.

Detection of breast cancer by radiologists all this time was based on direct observation by naked eye and experience of the radiologist. By using that method, there were weaknesses in analyzing the mammogram image because there were some condition that were difficult to distinguish using the direct observation, because of the noise and low contrast images and other human factors, such as fatigue, mood, etc. This condition causes the need for a method that is able to assist radiologist in diagnosing breast screening results more precisely. The higher the value of accuracy is found, it will be the smaller abnormality detection errors that occur thereby it can minimize the circumstances which could endanger the lives of patients. Based on the data obtained, there were approximately 5-10% misdiagnoses in analyzing the mammogram image.

This study used the intact image of the breast without crop on the ROI and without segmentation, then did the pre-processing and texture-based feature extraction to obtain accurate classification results. Classifications of breast abnormality in this study were divided into three classes: normal, benign and malignant. Phase of the study consisted of a pre-processing stage, the removal of a

label or artifacts that was found on the mammogram image, contrast enhancement by comparing several methods (CLAHE, NSCT, CS) to obtain the best method, image cropping, filtering and smoothing with a median filter to the image. Furthermore, the output of the pre-processing stage will be input to the feature extraction stage. The Feature extraction that was used in this study were texture-based namely GLCM, GLRLM and histogram.

The results from feature extraction will be input to classification process. The classification process used the MLP with BEP learning. For classification, by testing techniques using k-fold cross validation, the optimal values was obtained at 10th fold. MLP was compared to other classifiers namely SVM and Naïve Bayes. The maximum value of classification results was obtained using CLAHE contrast enhancement methods, GLCM-based feature extraction method and MLP classifier with accuracy of 98,33%, sensitivity of 100% and a specificity of 97,5% for 2 classes normal and abnormal. If we used hierarchical classification, then performed more abnormal classification (benign and malignant), we obtained the accuracy of 82,5%, sensitivity of 80% and specificity of 85%. While for the classification of three classes (normal, malignant and benign) directly without hierarchy obtained accuracy of 90%, sensitivity of 85% and specificity of 87,5%. In this study, to examine the influence of enhancement process (contrast improvement and filtering), we conducted a two-class classification of normal and abnormal images without enhancement. By that examination, we obtained that the results were worse if the image didn't begin with the enhancement process namely accuracy of 83,333%, sensitivity of 92,5% and specificity of 80%.

**Keywords** - Early detection, screening, breast cancer, CLAHE, GLCM, MLP.

## INTISARI

Kanker payudara merupakan pembunuh nomor satu pada wanita di dunia. Dengan adanya deteksi dini kanker payudara kesempatan untuk bertahan hidup bagi penderita dapat ditingkatkan sekitar 95%. Di Amerika deteksi dini kanker payudara mampu menyelamatkan nyawa pasien sekitar 12 sampai 37 jiwa perhari. Sedangkan di Indonesia menurut profil kesehatan Departemen Kesehatan Republik Indonesia Tahun 2012 kanker tertinggi yang diderita wanita Indonesia adalah kanker payudara dengan angka kejadian 2,2 % dari 1000 perempuan. Jika hal ini tidak terkendali, maka diperkirakan pada tahun 2030 akan ada 26 juta orang menderita kanker payudara dan 17 juta meninggal dunia.

Deteksi kanker payudara oleh ahli radiologi selama ini hanya secara kasat mata dan berdasarkan jam terbang dari para radiolog. Dengan cara tersebut, terdapat kelemahan dalam menganalisis citra *mammogram* karena ada beberapa kasus yang sulit dibedakan dengan mata telanjang, karena tertutup noise dan kontras citra yang rendah serta faktor manusiawi lainnya, seperti kelelahan, mood, dll. Kondisi tersebut menyebabkan dibutuhkan sebuah metode yang mampu membantu dokter dalam mendiagnosis hasil skirining payudara secara lebih akurat. Karena semakin tinggi akurasi yang didapatkan, maka akan semakin kecil kesalahan deteksi keabnormalan yang terjadi yang bisa membahayakan jiwa pasien itu sendiri berdasarkan data yang didapat masih terdapat kesalahan diagnosis sekitar 5-10 % dalam menganalisis citra *mammogram*.

Pada penelitian ini, dipakai citra utuh bagian payudara yang dilakukan proses pra pengolahan dan ekstraksi fitur berbasis tekstur untuk mendapatkan hasil klasifikasi yang akurat. Klasifikasi jenis keabnormalan payudara dalam penelitian ini dibagi dalam 3 kelas yaitu normal, jinak (*benign*) dan ganas (*malignant*). Tahap dari penelitian ini terdiri dari proses pra pengolahan, yaitu penghilangan label atau artifak yang terdapat pada citra *mammogram*, peningkatan kontras dengan membandingkan beberapa metode (CLAHE, NSCT, CS) untuk mendapatkan metode terbaik, *cropping*

citra, *filtering* serta *smoothing* dengan tapis *median* pada citra. selanjutnya keluaran dari tahap pra pengolahan akan menjadi masukan pada tahap ekstraksi fitur. Ekstraksi fitur yang dipakai di penelitian ini berbasis tekstur yaitu GLCM, GLRLM dan *histogram*.

Hasil dari ekstraksi fitur akan dijadikan masukan untuk proses klasifikasi. Proses klasifikasi disini dengan menggunakan metode MLP dengan pembelajaran BEP. Untuk klasifikasi, dengan teknik *testing* memakai *k-fold cross validation*, nilai optimal didapatkan pada *fold* yang ke 10. MLP dibandingkan dengan *classifier* lain yaitu SVM dan *Naïve Bayes*. Hasil dari klasifikasi didapatkan nilai maksimal dengan metode peningkatan kontras CLAHE, metode ekstraksi fitur berbasis GLCM dan *classifier* MLP, dengan nilai akurasi 98,33%, sensitifitas 100% dan spesifisitas 97,5% untuk 2 kelas normal dan abnormal. Jika dilakukan klasifikasi bertingkat (hirarki), dilakukan lagi klasifikasi untuk kelas abnormal (*benign* dan *malignant*), didapatkan nilai akurasi 82,5%, sensitifitas 80% dan spesifisitas 85%. Sedangkan untuk klasifikasi 3 kelas (normal, *malignant* dan *benign*) secara langsung tanpa hirarki didapatkan nilai akurasi 90%, sensitifitas 85% dan spesifisitas 87,5%. Untuk melihat sejauh mana pengaruh proses *enhancement* (peningkatan kontras dan *filtering*), pada penelitian ini juga dilakukan klasifikasi dua kelas citra normal dan abnormal tanpa *enhancement* dan citra dengan tahapan *enhancement*. Didapatkan hasil yang lebih buruk jika citra tidak diawali dengan proses *enhancement* yaitu dengan nilai akurasi 83,333%, sensitifitas 92,5% dan spesifisitas 80%.

**Kata Kunci** - Deteksi dini, Skrining, Kanker payudara, CLAHE, GLCM, MLP.