

EKSTRAKSI FITUR DAN KLASIFIKASI MENGGUNAKAN METODE GLCM DAN SVM PADA CITRA MAMMOGRAM UNTUK IDENTIFIKASI KANKER PAYUDARA

Bebby Dwi Junita

*Jurusan Teknik Informatika, Fakultas Teknologi Industri, Universitas Gunadarma
bebbby_dj12@yahoo.com*

Abstrak

Kanker payudara merupakan salah satu penyakit kanker paling tua pada manusia dan menempati urutan kedua pada perempuan setelah kanker rahim. Berdasarkan Sistem Informasi RS (SIRS), jumlah penderita kanker payudara di Indonesia sekitar 28.7%. Mammogram adalah pemeriksaan radiologi khusus menggunakan sinar-X dosis rendah. Keabnormalan kanker payudara dapat dikenali dengan keberadaan massa atau mikroklasifikasi pada citra mammogram. Computer-Aided Diagnosis (CAD) dikembangkan untuk meningkatkan deteksi utama dari penyakit ini dimana proses tersebut antara lain preprocessing, deteksi dan segmentasi, ekstraksi fitur, serta klasifikasi. Dalam penelitian ini dilakukan ekstraksi fitur dengan menggunakan metode Gray Level Co-occurrence Matrix (GLCM), metode tersebut mengekstraksi empat statistik tekstur yaitu, Contrasts, Correlation, Energy dan Homogeneity yang dapat digunakan dalam proses ekstraksi nilai pada citra. Setelah itu dilakukan klasifikasi menggunakan metode Support Vector Machine (SVM). Metode analisis tekstur menggunakan GLCM dan proses klasifikasi menggunakan SVM pada citra mammogram menghasilkan nilai probabilitas sebesar 60% dengan data pengujian yang digunakan 10 citra gabungan (5 jinak dan 5 ganas) dan data pembelajaran yang digunakan 10 data jinak dan 10 data ganas.

Kata Kunci : *Image Processing, Ekstraksi Fitur, SVM, Kanker Payudara, Matlab.*

FEATURE EXTRACTION AND CLASSIFICATION BY USING SVM AND GLCM METHOD ON THE MAMMOGRAM IMAGE FOR BREAST CANCER IDENTIFICATION

Abstract

Breast cancer is one of the oldest cancers in humans and second rank most happened to women after cervical cancer. Based on Hospital Information System (SIRS), the number of breast cancer patients in Indonesia is around 28.7%. Mammogram is a special radiological examination by using low-dose X-rays. Abnormality of breast cancer can be identified by the presence of mass or micro-classification on the mammogram images. Computer-Aided Diagnosis (CAD) is developed to improve the primary detection of this disease where the processes are preprocessing, detection and segmentation, feature extraction and classification. In this research, the feature extraction is conducted by using Gray Level Co-occurrence Matrix (GLCM), the method extracts four texture statistics, namely, the Contrasts, Correlation, Energy and Homogeneity that can be used in the values extraction process in the image. Then, the classification using Support Vector Machine (SVM) is conducted. Texture analysis methods using GLCM and classification process using SVM on the mammogram image generate 60% probability value with using 10 combined images of test data (5 benign and 5 malignant) and using 10 benign data and 10 malignant data of the learning data.

Keywords: *Image Processing, Feature Extraction, SVM, Breast Cancer, Matlab.*

PENDAHULUAN

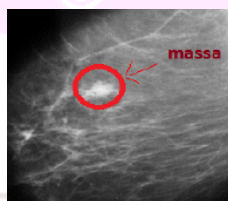
Kanker payudara menempati urutan kedua pada perempuan setelah kanker rahim. Kanker payudara merupakan penyebab kematian utama diantara wanita yang berusia antara 35-50 tahun. Jika skrining mamografi dilakukan secara teratur, dapat memberikan keuntungan pada wanita yang berusia dibawah 50 tahun. Berdasarkan Sistem Informasi RS (SIRS), jumlah penderita kanker payudara di Indonesia (28.7%). Pengobatan kanker payudara pada stadium awal secara umum memberi hasil yang lebih baik dibanding stadium lanjut, oleh karena itu perempuan sebaiknya memahami tentang kanker payudara sehingga bisa mendeteksi secara dini terhadap gejala-gejala terjadinya kanker payudara. Kanker terjadi ketika sel abnormal membagi dan tidak terkontrol, mereka dapat menjadi besar dengan membentuk jaringan ekstra, atau tumor, yang dapat menjadi jinak atau ganas [Sugita, 2012].

Deteksi dini kanker payudara dapat dilakukan dengan teknologi pencitraan medis yang saat ini berkembang dan teknologi tersebut memungkinkan pemanfaatan citra medis untuk membantu para-

medis dalam melakukan diagnosa penyakit, dalam kasus ini adalah untuk mendeteksi kanker payudara. Citra medis tersebut dapat dihasilkan dari berbagai macam peralatan yang digunakan di bidang kedokteran, seperti Ultra Sound Graphic (USG), MRI, CT-Scan/CAT-Scan, dan Diagnostic Mammogram [Nagaraj, Paga, Lamichhane, 2014].

Mammogram adalah pemeriksaan radiologi khusus menggunakan sinar-X dosis rendah untuk deteksi kelainan pada payudara atau ada tidaknya abnormalan pada payudara, bahkan sebelum adanya gejala yang terlihat pada payudara seperti benjolan yang dapat dirasakan, kanker payudara dapat dikenali dengan keberadaan massa atau keberadaan mikrokalsifikasi pada citra mammogram [Rizqi, 2012], yaitu :

1. Massa adalah area terdapatnya pola tekstur dengan bentuk serta batas area tertentu pada proyeksi foto mamografi (Gambar 1).
2. Mikrokalsifikasi yaitu deposit-deposit kecil kalsium dalam jaringan payudara yang terlihat sebagai titik-titik kecil putih di sekitar jaringan payudara (Gambar 2).



Gambar 1. Massa [Lisha V, Anitha, 2013]



Gambar 2. Mikrokalsifikasi [Lisha V, et al., 2013]

Dalam rangka meningkatkan kinerja diagnostik ahli radiologi, beberapa skema *Computer-Aided Diagnosis* (CAD) telah dikembangkan untuk meningkatkan deteksi utama dari penyakit ini. *Computer-Aided Diagnosis* (CAD) dikembangkan untuk meningkatkan deteksi utama dari penyakit ini dimana proses tersebut antara lain *preprocessing*, deteksi dan segmentasi, ekstraksi fitur, serta klasifikasi. Dalam penelitian ini dilakukan ekstraksi fitur dengan menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM)

METODE PENELITIAN

Dalam setiap metode penelitian pastinya memiliki langkah-langkah umum dalam pelaksanaannya, seperti proses pengambilan citra, *preprocessing*, ekstraksi fitur dan klasifikasi dimana nantinya proses-proses tersebut menghasilkan sebuah data yang dapat diolah.

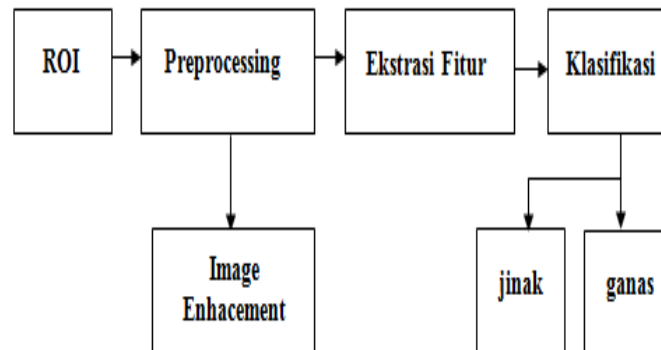
Gambar 3 adalah gambaran metode penelitian secara umum.

Region of Interest (ROI)

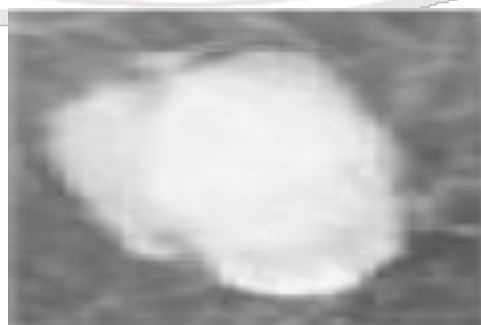
Suatu citra diperoleh dari hasil mammogram kanker payudara, pengambilan citra untuk analisa tekstur sendiri harus memiliki kriteria yaitu citra mammogram dengan format .bmp yang dipilih dengan *bitdepth* 24 dengan 256 warna dan pada proses pengambilan citra hanya dipilih bagian *region of interest* (ROI) (Gambar4).

Preprocessing

Pada tahap *preprocessing* terdapat dua buah proses, pertama adalah proses normalisasi citra, didalam proses normalisasi citra terdapat proses *resizing* untuk mengatur ukuran citra. Tahap kedua ialah proses *image enhancement*, yaitu proses peningkatan kualitas citra menggunakan teknik *sharpening*.



Gambar 3. Gambaran Umum Metodologi Penelitian Ekstraksi Fitur dan Klasifikasi



Gambar 4. ROI Mammogram Kanker Payudara

Ekstrasi Fitur

Dalam tahapan ekstrasi fitur ini menggunakan metode GLCM. GLCM merupakan langkah penting, dikarenakan fitur yang diekstrak tersebut digunakan sebagai informasi yang menentukan kriteria nilai dari citra kanker itu sendiri. Didalam ekstrasi fitur ini dilakukan 10 kali pengambilan nilai untuk citra pembelajaran dan citra pengujian. Contoh hasil ekstraksi fitur dapat dilihat pada Gambar 5.

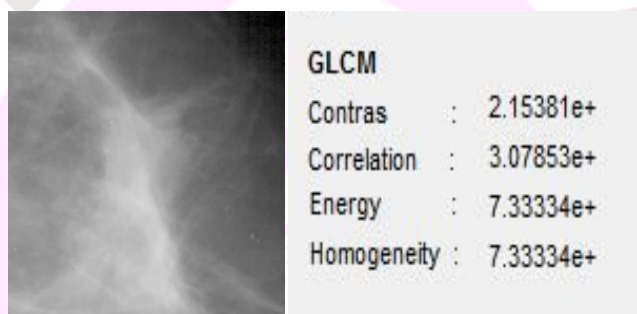
Klasifikasi

Pada proses klasifikasi terdapat beberapa rangkaian proses seperti pro-

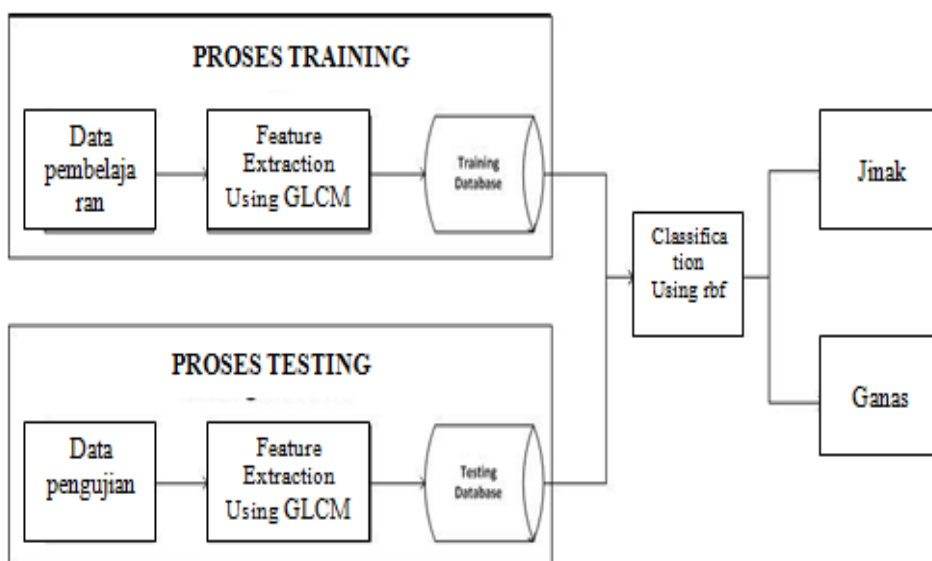
ses pengklasifikasi nilai pada citra pembelajaran dan proses pengklasifikasian nilai pada citra pengujian dimana nantinya akan menghasilkan sebuah hasil apakah citra pengujian tersebut termasuk kategori citra jinak atau ganas. Rangkaian langkah-langkahnya dapat dilihat pada Gambar 6.

Perancangan Aplikasi

Aplikasi yang dibuat juga didukung dan berasal dari rancangan hardware dan software menggunakan pendekatan metode SDLC (*Software Development Live Cycle*).



Gambar 5: Hasil Ekstrasi Fitur Menggunakan Metode GLCM



Gambar 6. Blok Diagram Ekstrasi Fitur

Pendefinisian Kebutuhan

Aplikasi pendeteksian citra kanker payudara ini memiliki beberapa syarat agar dapat digunakan. Aplikasi ini membutuhkan:

1. Sebuah citra yang diambil kamera harus murni dari pencahayaan langsung dimana dapat merusak pengambilan nilai input yang diperoleh melalui metode GLCM.
2. Sebuah program terpisah yang berfungsi merekonstruksi ketajaman citra *gray* sebelumnya dengan menampilkan grafik perubahannya.
3. Sebuah program terpisah yang berfungsi untuk membangun ekastrasi fitur dari hasil dengan inputan yang diperoleh, melakukan simulasi dari hasil pembelajaran menggunakan SVM dan metode pengambilan keputusan.

Proses Aplikasi

Pada tahap ini dilakukan perancangan antarmuka grafis untuk pengguna. Perancangan antarmuka ini melibatkan sedikit penggunaan alat bantu

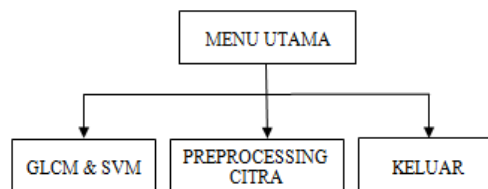
pembuatan citra seperti kamera dan sebagainya sehingga diusahakan semaksimal mungkin untuk memperoleh tampilan antarmuka yang memudahkan para pengguna, termasuk pengecekan *bug* dan perbaikan sehingga menghasilkan program sesuai kebutuhan. Berikut merupakan langkah-langkah yang digunakan dalam proses perancangan aplikasi.

PEMBAHASAN

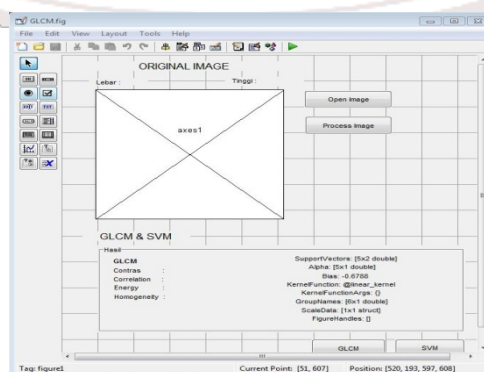
Pembahasan berisi tentang pembuatan GUI dan hasil pendeteksian kanker menggunakan metode GLCM.

Pembuatan GUI

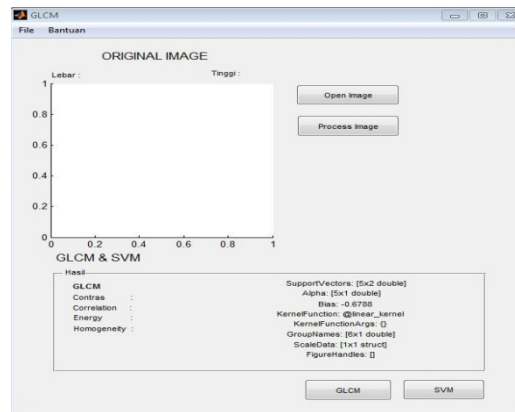
Jendela ini merupakan yang muncul pertama kali pada saat aplikasi ini dijalankan, jendela ini merupakan pusat dari kegiatan yang dilakukan pada aplikasi ini. Untuk membuat tampilan jendela utama ini sesuai dengan (Gambar 8 dan Gambar 9) maka rancangan GUI nya seperti pada gambar berikut.



Gambar 7. Struktur Navigasi Aplikasi GLCM & SVM



Gambar 8. Rancangan GUI Jendela Utama



Gambar 9. Hasil dari Program GLCM.m

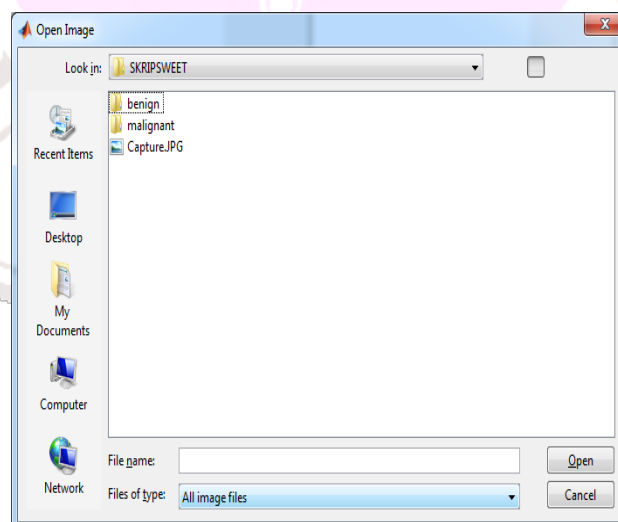
Uji Coba Aplikasi

Aplikasi pendeteksian kanker payudara ini mengalami beberapa ujian *runtime* seperti mengambil nilai GLCM dari data pembelajaran dan data testing yang akan digunakan untuk parameter data SVM pada saat aplikasi selesai dikompilasi. Pengujian dilakukan untuk melihat apakah *exception* yang ditetapkan oleh kode sumber program dapat menangani kesalahan perintah terjadi saat aplikasi berjalan untuk pertama kalinya (sebelum proses *debug*). Setelah mengalami serangkaian perbaikan

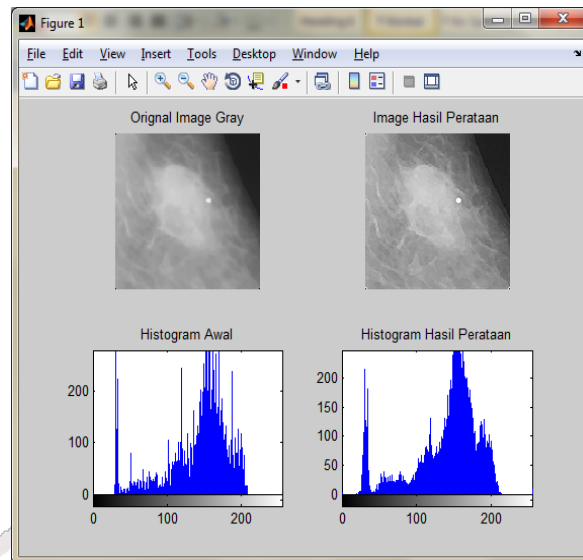
kode sumber program sesuai dengan kebutuhan dan layak untuk dijalankan secara *runtime*. Pertama adalah proses pengambilan citra sebagai contoh pemrosesan secara GUI dengan cara menekan tombol *Open Image* maka akan nampak seperti gambar 10.

Kemudian tekan tombol *Proses Image* untuk mendapatkan hasil *pre-processing* citra seperti gambar 11.

Setelah itu dilakukan ekstraksi fitur dan klasifikasi menggunakan GLCM. Hasil klasifikasi dapat dilihat pada Tabel 1.



Gambar 10. Mengambil Data Citra



Gambar 11. Preprocessing Citra

Tabel 1. Data Range GLCM Klaisifikasi

Citra	Nilai Sebenarnya	Nilai Setelah Klasifikasi	Keterangan
1	1	1	Berhasil
2	0	1	Gagal
3	1	1	Berhasil
4	0	0	Berhasil
5	0	1	Gagal
6	0	1	Gagal
7	0	0	Berhasil
8	1	1	Berhasil
9	1	0	Gagal
10	0	0	Berhasil
Probabilitas			60%

Dari data (Tabel1) dapat dihitung probabilitas yaitu dengan cara menjumlahkan jumlah data yang berhasil diujicoba dibagi dengan total data pengujian (10 buah) dan dikalikan 100. Jumlah data yang digunakan adalah 10 citra, yang berhasil dalam uji coba aplikasi adalah 6 dari 10 citra tersebut. Sehingga mendapatkan nilai probabilitas sebesar 60%. Berikut adalah rumus yang digunakan untuk menghitung probabilitas.

$$\frac{\text{Edata berhasil}}{\text{Edata selurus}} \times 100\% \dots [1]$$

Sehingga dapat disimpulkan bahwa metode *Gray Level Co-occurrence Matrix* GLCM) dan *Support Vector Machine* (SVM) memiliki tingkat keberhasilan sebesar 60% dari hasil uji coba tersebut dalam identifikasi kanker payudara menggunakan citra mammogram.

SIMPULAN

Metode analisis tekstur menggunakan GLCM dan proses klasifikasi menggunakan SVM pada citra mammogram berhasil dikembangkan menggunakan citra ROI mammogram

identifikasi kanker payudara. Aplikasi GLCM dan SVM ini berhasil menunjukkan nilai-nilai pengukuran statistik dalam bentuk histogram dari hasil ekstraksi fitur menggunakan GLCM dan klasifikasi metode SVM pada citra kanker yang jinak dan ganas yang menghasilkan nilai probabilitas sebesar 60% dengan data pengujian yang digunakan sebanyak 10 citra gabungan (5 jinak dan 5 ganas) dan data pembelajaran yang digunakan sebanyak 10 data jinak dan 10 data ganas. Nilai tersebut cukup baik dan nilai tersebut dapat berbeda-beda tergantung pada jumlah database dan metode yang digunakan. Hasil pembuatan aplikasi tersebut menampilkan nilai-nilai yang digunakan sebagai data dalam objek penelitian seperti nilai GLCM, histogram dan nilai klasifikasi SVM.

DAFTAR PUSTAKA

- [1] Adi Purnomo, Sulistyo Puspitodjati. 2013. Analisis Tekstur Kayu Parquet dengan Menggunakan Metode Grey Level Co-occurrence Matrix (GLCM).
- [2] Abdul Kadir, Adhi Susanto. 2013. Teori Dan Aplikasi Pengolahan Citra, Penerbit: Andi.
- [3] Beny Firman. 2010. Mesin Vektor Pendukung (svm) Kasus Non Linier.
- [4] Chang, Chih-Chung, Lin, Chih-Jen, 2013. *A Library For Support Vector Machines*. Libsvm.
- [5] Christianini. 2000. Teori Pembelajaran pada Support Vector Machine (SVM).
- [6] Dipankar, Hazra. 2011. *Texture Recognition With Combined glcm, Wavelet and Rotated Wavelet Features*.
- [7] Eko Prasetyo. 2011. Pengolahan Citra Digital dan Aplikasinya menggunakan Matlab, Penerbit: Andi.
- [8] Feriza, Irawan. 2012, Buku Pintar Pemrograman Matlab, Penerbit: MediaKom.
- [9] Guru Kumar, Mohanaiah, Sathyanarayana. 2013. *Image Texture Feature Extraction Using glcm For Human*.
- [10] Karatzoglou. 2004. *Kernel method*.
- [11] Kiswanto. 2012. Identifikasi Citra Untuk Mengidentifikasi Jenis Daging Sapi Dengan Menggunakan Transformasi Wavelet Haar.
- [12] Kusumanto. 2011. Pengolahan Citra Digital Untuk Mendeteksi Objek Menggunakan Pengolahan Warna Model Normalisasi RGB. Semantik.
- [13] Mehul. Sampat, Markey, Bovik. 2005. *Computer-Aided Detection and Diagnosis in Mammography*.
- [14] Meritxell, Arnau, Robert, Mariona, Sergi, Lidia, Melcior, Freixenet. 2012. *Adapting Breast Density Classification Form digitized to Full Field Digital Mammograms In Breast Imaging, pages 561-568*. Springer.
- [15] Nagaraj, Paga, Prasanna, Lamichhane. 2014. *Early Breast Cancer Detection Using Statistical Parameters*. International Journal of Research in Engineering & Technology (IMPACT: IJRET).
- [16] Nugroho. 2003. Metode SVM Dan Pengaplikasiannya.
- [17] Poongodi, Samuel, Manoharan, Narain, Ponraj, Evangelin, Jenifer. 2011. *A Survey On The Preprocessing Techniques Of Mammogram For The Detection of Breast Cancer*.
- [18] Santosa. 2007. *Metode One-Against-One*.
- [19] Sianipar. 2013. Pemrograman Matlab Dalam Contoh dan

- Penerapan, Penerbit: Informatika Bandung.
- [20] Sugita, Lina. 2012, *Tingkat Kecukupan Energi dan Protein Tingkat Pengetahuan Gizi Jenis Terapi Kanker dan Status Gizi Pasien Kanker Rawat Inap di Rumah Sakit Kanker Dharmas*. Departemen Gizi Masyarakat Fakultas Ekologi Manusia Institut Pertanian Bogor (IPB).
- [21] Sembiring. 2007. Metode Analisa dalam *Support Vector Machine (SVM)*.
- [22] Tanuja, Sarode, Suryawanshi, Kekre, Sudeep, Thepade. 2010. *Image Retrieval Using Texture Features Extracted*.
- [23] Rizqi Adnamazida. 2012. Apa sih mammogram itu
URL: <http://www.artikelpayudara.com/2009/04/15/mendeteksi-kanker-payudara-dengan-mamografi>, tanggal

