

## **ANALISIS TEKSTUR CITRA MIKROSKOPIS KANKER PARU MENGGUNAKAN METODE *GRAY LEVEL CO-OCCURANCE MATRIX* (GLCM) DAN TRANSFORMASI WAVELET DENGAN KLASIFIKASI NAIVE BAYES**

*Rizky Ayomi Syifa<sup>1)</sup>, Kusworo Adi<sup>1)</sup> dan Catur Edi Widodo<sup>1)</sup>*

<sup>1)</sup> Departemen Fisika, Fakultas Sains dan Matematika, Universitas Diponegoro, Semarang

E-mail: rizky\_ayomi\_syifa@st.fisika.undip.ac.id

### **ABSTRACT**

*This research, conduct the lung cancer detection system on a microscopic image. The microscopic image used is the result from lung biopsy. If there is a cancerous tissue in the image of lung biopsy, the texture will be irregular, while the image of the normal lung biopsy will have a regular texture. The purpose of this research is to develop the lung cancer detection system and also to count the performance of the lung cancer detection system. The classification process uses two methods, Gray Level Co-Occurrence Matrix (GLCM) and Daubechies Wavelet Transform (db1). The Daubechies wavelet transformation is used in decomposition in level 4, while the offset of GLCM is 6. The feature extraction process is done in the transformation wavelet using the 4 subbands, approximation, horizontal Detils coefficients, vertical Detil coefficients and diagonal Detil coefficients, and the feature extraction of GLCM uses the contrast, correlation, homogeneity and energy as the parameter. The naïve bayes classification requires 2 parameter input, do a classification is 4 combination from each method of feature extraction. The result of this research is to extent the level of accuracy for the extraction of the feature extraction in 71,42% wavelet transformation method for the combination coefficients approximation-diagonal Detil coefficients and 80% accuracy of GLCM method for the combination of homogeneity-correlation.*

**Key words :** Lung Cancer, Microscopic Image, Wavelet Transform, Gray Level Co Occurrence Matrix. Naive Bayes Classification

### **ABSTRAK**

Pada penelitian ini dilakukan sistem deteksi kanker paru pada citra mikroskopis. Citra mikroskopis yang digunakan adalah hasil biopsi paru. Citra biopsi paru jika terdapat jaringan kanker, memiliki tekstur yang tidak teratur, sedangkan pada citra biopsi paru normal akan memiliki tekstur yang teratur. Penelitian ini bertujuan mengembangkan sistem deteksi kanker paru dan menghitung kinerja sistem deteksi kanker paru. Proses klasifikasi dilakukan dengan menggunakan dua metode ekstraksi ciri yaitu *Gray Level Co-Occurrence Matrix* (GLCM) dan Transformasi wavelet. Transformasi wavelet yang digunakan wavelet daubechies dekomposisi level 4 dan GLCM jarak antar piksel (*offset*) yang digunakan sebesar 6. Proses ekstraksi ciri yang dilakukan pada transformasi wavelet menggunakan 4 subband yaitu aproksimasi, koefisien detail horizontal, koefisien detail vertikal dan koefisien detail diagonal dan ekstraksi ciri GLCM menggunakan parameter kontras, korelasi, homogenitas dan energi. Proses klasifikasi naive bayes memerlukan dua parameter masukan, dilakukan kombinasi 4 parameter dari masing-masing metode ekstraksi ciri tersebut. Penelitian ini menghasilkan tingkat akurasi sebesar 71,42% untuk ekstraksi ciri metode transformasi wavelet pasangan kombinasi koefisien aproksimasi-koefisien detail vertikal dan akurasi pada metode GLCM sebesar 80% untuk kombinasi homogenitas-korelasi.

**Kata Kunci :** Kanker Paru, Citra Mikroskopis, Transformasi Wavelet, Gray Level Co-Occurrence Matrix, Klasifikasi Naive bayes,

## **PENDAHULUAN**

Kanker paru-paru merupakan salah satu penyebab mortalitas terbesar di dunia. Berdasarkan data GLOBOCAN, *International Agency for Research Center* (IARC) tercatat tahun 2012 sebesar 19,4% masyarakat dunia meninggal akibat kanker paru-paru [1]. Menurut *World Health Organization* (WHO) pada tahun 2030 di prediksi sekitar 12 juta jiwa meninggal akibat kanker paru-paru [2]. Deteksi dini dengan menggunakan pemeriksaan mikroskopis mengharuskan memiliki tiga kriteria penting, yaitu, sudah ada fase pra klinis, teknologi untuk pemeriksaan secara pra klinis sudah tersedia, dan lebih efektif jika diagnosa ditemukan pada pemeriksaan [3]. Pemeriksaan kanker paru dilakukan dengan tiga proses, yaitu CT Scan, pemeriksaan secara mikroskopis (*Sputum Cytology*) dan biopsi.

Biopsi paru merupakan pengambilan jaringan paru yang kemudian diamati dibawah mikroskop. Proses diagnosa menggunakan biopsi, praktisi medis masih menggunakan pengamatan secara visual dan dibutuhkan waktu yang lama, sehingga diperlukan teknologi yang dapat meningkatkan tingkat akurasi dengan waktu yang cepat dan akurat pada proses pendekripsi kanker paru-paru. Proses deteksi dapat dilakukan melalui pengolahan citra dengan ekstraksi ciri menggunakan transformasi wavelet dan *Gray Level Co-Occurrence Matrix* (GLCM), serta proses klasifikasi menggunakan Naive Bayes. Ekstraksi ciri menggunakan transformasi wavelet, GLCM dan klasifikasi Naive Bayes sudah banyak diaplikasikan dibidang medis seperti yang telah dilakukan oleh Rahayuda (2016).

## **DASAR TEORI**

### **Kanker Paru**

Kanker merupakan pertumbuhan dan penyebaran sel-sel abnormal yang memiliki karakter yang khas. Kanker yang sudah menyebar dan tidak terkontrol dapat menyebabkan kematian [4]. Gejala kanker paru terdiri dari sesak nafas, nafsu makan berkurang,

nyeri pada dada yang tidak membaik, batuk disertai darah, kesulitan menelan dan kehilangan berat badan [2]. Keluhan penurunan dini kanker paru jarang terjadi, biasanya keluhan yang ringan terjadi pada pasien yang telah memasuki stadium II dan III [5]. Biopsi merupakan salah satu prosedur yang dilakukan untuk mengetahui area kanker. Pada biopsi, jaringan atau cairan diambil dari bagian tubuh (misal;paru-paru) dan diamati dibawah mikroskop yang disebut dengan patologi. Jika biopsi terindikasi terdapat kanker maka identifikasi dilanjutkan dengan mengetahui tipe kanker. Perbedaan terlihat jelas pada tekstur jaringan paru sehat dan jaringan paru yang memiliki sel kanker.

### **Ekstraksi Ciri**

Ekstraksi ciri (*feature extraction*) merupakan bagian fundamental dari analisis citra. Fitur adalah karakteristik unik dari suatu objek. Analisis bentuk merupakan salah satu metode pemisahan fitur. Karakteristik fitur yang baik dapat memenuhi syarat, seperti:

1. Membedakan suatu objek dengan yang lainnya
2. Memperhatikan kompleksitas komputasi
3. Tidak terikat (*independence*), bersifat invarian terhadap berbagai transformasi (rotasi, penskalaan, pergeseran, dan lain sebagainya).
4. Jumlahnya sedikit [6]

### **Transformasi Wavelet**

Transformasi Wavelet akan menkonversi suatu sinyal ke dalam sederetan Wavelet. Gelombang singkat tersebut merupakan fungsi basis yang terletak pada waktu berbeda. Wavelet merupakan fungsi matematika yang memotong-motong data menjadi kumpulan-kumpulan frekuensi yang berbeda, sehingga masing-masing komponen tersebut dapat dipelajari dengan menggunakan skala resolusi yang berbeda. Transformasi *wavelet* terdiri atas Transformasi Wavelet Kontinu atau *Continuous Wavelet Transform* (CWT) dan Transformasi Wavelet Diskrit atau DWT. Perhitungan skala

dan pergeseran dalam CWT dapat dilakukan secara kontinu, sedangkan DWT hanya dilakukan pada sekelompok skala tertentu. Wavelet Daubechies adalah salah satu keluarga wavelet orthogonal. Wavelet Daubechies disimbolkan dengan dbN dengan N adalah angka indeks dari 2 sampai 20. Persamaan skala atau persamaan dilatasi (pergeseran) merupakan fungsi skala yang mengalami kontraksi (meregang) dan pergeseran [7]. Wavelet Daubechies merupakan pengembangan dari wavelet Haar. Wavelet Daubechies orde satu (db1) dengan panjang filter 2 sama dengan wavelet Haar. Daubechies 2 (db2) adalah wavelet Daubechies dengan banyak filter 6 dan seterusnya [8].

#### Gray Level Co-Occurance Matrix (GLCM)

GLCM merupakan metode analisis tektur yang banyak digunakan terutama untuk tektur tidak teratur. Metode GLCM dapat meningkatkan rincian suatu citra dan di representasikan. Pada eksraksi ciri tektur metode GLCM merupakan matriks orde 2, sebuah matriks yang terdiri dari jumlah baris dan kolom yang sama [9]. Proses ekstraksi ciri GLCM melalui 4 arah sudut dengan interval  $45^\circ$  yaitu  $0^\circ, 45^\circ, 90^\circ, 135^\circ$ , sedangkan jarak antar piksel sebesar 1 piksel. Matriks kookurensi merupakan matriks bujursangkar dengan jumlah elemen sebanyak kuadrat jumlah level intensitas piksel pada citra. Matriks kookurensi merupakan matriks bujursangkar dengan jumlah elemen sebanyak kuadrat jumlah level intensitas piksel pada citra. Setiap titik  $(p, q)$  pada matriks kookurensi berorientasi  $\theta$  berisi peluang kejadian piksel bernilai  $p$  bertetangga dengan piksel bernilai  $q$  pada jarak  $d$  serta orientasi  $\theta$  dan  $(180 - \theta)$  [10].

Ekstraksi ciri dengan analisis tektur dilakukan dengan mengambil ciri dari citra *Grayscale* untuk membedakan obyek satu dengan obyek lainnya obyek diekstrak berdasarkan parameter berikut yaitu berupa entropi, kontras, energi, homogenitas, dan korelasi.

- a. Entropi  
 $Entropi = -\sum_i \sum_j p(i,j) \log p(i,j)$  (1)
- b. Energi  
 $Energi = \sum_i \sum_j p^2(i,j)$  (2)
- c. Kontras  
 $Kontras = \sum_i \sum_j (i-j)^2 p(i,j)$  (3)
- d. Homogenitas  
 $Homogenitas = \sum_i \sum_j \frac{p(i,j)}{1+|i-j|}$  (4)
- e. Korelasi  
 $Korelasi = \sum_i \sum_j \frac{ijp_d(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$  (5)

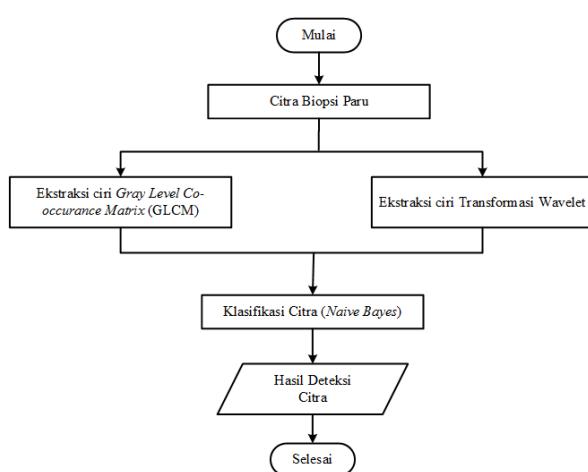
#### Klasifikasi Naive Bayes

Klasifikasi Naive Bayes merupakan pengklasifikasian statistik yang dapat digunakan untuk probabilitas keanggotaan suatu kelas. Naive bayes didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan *decision tree* dan *neural network*. Naive Bayes terbukti memiliki akurasi dengan data yang besar [11]. HMAP merupakan model penyederhanaan dari metode bayes, yang akan digunakan di dalam *machine learning* sebagai metode untuk mendapatkan hipotesis untuk suatu keputusan. HMAP (*Hypothesis Maximum Appropro Probability*) menyatakan hipotesa yang diambil berdasarkan nilai probabilitas kondisi prior yang dapat diketahui [12]. Dengan menggunakan algoritma bayes  $P(Y = y_i | X)$  dapat di representasikan menjadi:  $P(Y = y_i | X = x_k) = \frac{P(X = x_k | Y = y_i)P(Y = y_i)}{\sum_j P(X = x_k | Y = y_j)P(Y = y_j)}$  (6) dengan  $y_m$  merupakan nilai untuk  $Y$ ,  $x_k$  merupakan nilai vektor untuk  $Y$  dan penjumlahan pada penyebut menunjukkan seluruh nilai dari variabel acak  $Y$  [13].

#### METODE PENELITIAN

Langkah awal, melakukan proses akuisisi citra, dimana citra yang digunakan menggunakan data citra sekunder. Pada proses pengolahan citra menggunakan proses ekstraksi ciri dengan menggunakan tektur citra, sehingga digunakan metode wavelet dan GLCM. Metode

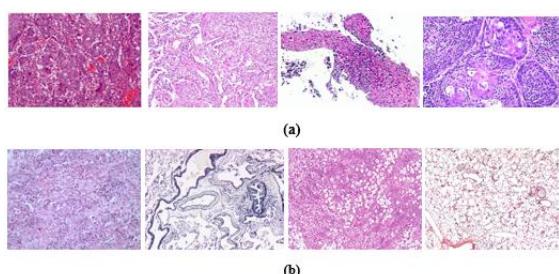
wavelet yang digunakan dalam penelitian menggunakan wavelet daubechies dengan dekomposisi level dan metode GLCM dengan jarak antar piksel sebesar 6 dengan menggunakan empat parameter (kontras, homogenitas, korelasi dan energi). Seluruh citra diubah menjadi *grayscale*, karena hanya dapat menganalisis satu kanal warna, kemudian seluruh citra dilakukan proses klasifikasi berdasarkan hasil ekstraksi ciri



**Gambar 1.** Diagram Proses Penelitian

## HASIL DAN PEMBAHASAN

Pada penelitian ini menggunakan citra sebanyak 35 citra yang terdiri atas 18 citra kanker dan 17 citra bukan kanker. Citra yang digunakan memiliki kanal RGB sehingga seluruh citra diubah menjadi satu kanal warna yaitu *grayscale*. Gambar 1 menunjukkan sampel citra kanker dan bukan kanker.



**Gambar 2.** Sampel Citra (a) Sampel Citra Biopsi Kanker Paru [14] [15]  
(b) Sampel Biopsi Citra Paru Normal [16] [17]

Pada citra yang memiliki sel kanker jaringan paru akan terlihat tidak teratur dan citra paru tidak memiliki sel kanker akan terlihat lebih teratur. Proses ekstraksi ciri menggunakan metode transformasi wavelet, citra diekstraksi dengan menggunakan skala frekuensi dengan proses dekomposisi sebanyak 4 level, dengan terbagi atas 4 buah subband yaitu, koefisien aproksimasi (LL), koefisien detail horizontal (HL), koefisien detail vertikal (LH) dan koefisien detail diagonal (HH). Tabel 1 menunjukkan hasil ekstraksi ciri transformasi wavelet dekomposisi 4 dengan menggunakan transformasi wavelet daubechies.

**Tabel 1** Hasil Ekstraksi Ciri Transformasi Wavelet

Kategori	LL	HL	LH	HH
Kanker				
	2766,06	8,590	33,039	1,484

Kategori	LL	HL	LH	HH
Bukan Kanker				
	3255,58	10,458	14,281	1,199

Pada proses ekstraksi ciri dengan menggunakan *Gray Level Co-Occurrence Matrix* (GLCM), citra diekstraksi dengan 4 parameter, yaitu energi, kontras, korelasi, dan homogenitas. Masing-masing parameter mengekstraksi tingkat keteraturan pada citra, dengan menggunakan 4 arah sudut  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ , dengan menggunakan jarak antar piksel sebesar 6. Tabel 2 menunjukkan hasil ekstraksi dengan GLCM.

**Tabel 2.** Hasil Ekstraksi Ciri GLCM

Kategori	Korelasi	Kontras	Energi	Homo genitas
Kanker	6,2450	0,5633	0,0641	0,0352
Bukan Kanker	4,2541	0,5871	0,1052	0,1225

Berdasarkan kedua ekstraksi tersebut keempat parameter di kombinasikan  $4C_2$ , sehingga mendapatkan 6 pasangan masukan, karena pada proses klasifikasi naive bayes memerlukan dua

masukan, sehingga diperoleh tingkat akurasi yang ditunjukkan pada Tabel 3 dan Tabel 4.

**Tabel 3.** Hasil Tingkat Akurasi Klasifikasi Transformasi Wavelet level 4

Parameter Masukan	Akurasi Sistem (%)
Aproksimasi – Horizontal	62,85
Aproksimasi – Vertikal	62,85
<b>Aproksimasi – Diagonal</b>	<b>71,42</b>
Horizontal – Vertikal	62,85
Horizontal – Diagonal	60
Vertikal – Diagonal	62,85

**Tabel 4.** Hasil Tingkat Akurasi Klasifikasi Citra Mikroskopis Metode GLCM dengan Jarak Antar Piksel = 6

Parameter Masukan	Akurasi Sistem (%)
Kontras-Energi	77,14
<b>Kontras-Homogenitas</b>	<b>80</b>
Kontras-Korelasi	71,42
<b>Homogenitas-Korelasi</b>	<b>80</b>
Homogenitas-Energi	68,57
Korelasi-Energi	62,85

Berdasarkan Tabel 3 dan Tabel 4, menunjukkan bahwa metode GLCM memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan transformasi wavelet. Hal ini dikarenakan pada metode GLCM proses ekstraksi ciri yang dilakukan dari kualitas citra, tidak hanya melalui proses secara frekuensi. Parameter kombinasi sangat mempengaruhi tingkat akurasi klasifikasi yang dilakukan. Pada kombinasi kontras-homogenitas, homogenitas-korelasi kedua pasangan tersebut memiliki tingkat akurasi yang sama sebesar 80%. Kedua pasangan kombinasi tersebut memiliki nilai yang sama besar, dikarenakan jumlah kedua parameter tersebut banyak yang berada di dalam satu kelas target. Sedangkan, untuk transformasi wavelet akurasi yang diperoleh sebesar 71,42%.

## KESIMPULAN

Berdasarkan penelitian ini maka dapat disimpulkan tingkat akurasi transformasi wavelet sebesar 71,42% pada pasangan koefisiesn aproksimasi - koefisien detail diagonal, sedangkan pada GLCM memiliki tingkat akurasi sebesar 80% pada pasangan kontras - homogenitas dan homogenitas -korelasi.

Transformasi wavelet menggunakan 4 parameter subband dengan dekomposisi level 4, yaitu koefisien aproksimasi, koefisien detail horizontal, koefisies detail vertikal, dan koefisien detail diagonal, sedangkan pada GLCM menggunakan 4 parameter yaitu energi, korelasi, kontras dan homogenitas.

## DAFTAR PUSTAKA

- [1] Ferlay, J., Isabelle S, Rajesh D, Sultan E, Colin M, Marise R, Donald M, Parkin, D F, Freddie B, 2010, Estimates of Worldwide Burden of Cancer in 2008: GLOBOCAN 2008, *International Journal of Cancer*, Volume 127, pp. 2893-2917.
- [2] Taher, F., Werghi, N., Al-Ahmad, H. dan Donner, C., 2013, Extraction and Segmentation of Sputum Cells for Lung Cancer Early Diagnosis, *Algorithms*, Volume 6, pp. 513-531.
- [3] Kennedy, T.C, Miller. Y dan Prindiville,C., "creening for Lung Cancer Revisted and the Role of Sputum Cytology and Flourecence Bronchoscopy in a High-Risk Group, *Multimodality Approach to Lung Cancer*, pp. 72-79, 2000.
- [4] Saksono, H.T., Rizal, A. dan Usman, K., 2010, Pendekstian Kanker Paru-Paru Dengan Menggunakan Transformasi Wavelet Dan Metode Linear Discriminant

- Analysis, *Teknologi Elektro* , Volume 9, Issue 1, pp. 18-26.
- [5] Perhimpunan Dokter Paru Indonesia., 2003, *Kanker Paru "Pedoman Diagnosis dan Penatalaksanaan Asma Di Indonesia"*, Jakarta : Pusat Data dan Informasi Kementerian Kesehatan Republik Indonesia.
- [6] Putra, D., 2010, *Pengolahan Citra Digital*, Yogyakarta : Andi
- [7] Lamabelawa, M.I.J dan Malelak, Y., 2012, Sistem Temu Kembali Tenun Ikat NTT dengan Transformasi Wavelet, *Seminar Nasional Sains dan Teknik (SAINSTEK 2012)*, pp 157-163
- [8] Suma'inna dan Gugun.G., 2013, Implementasi Transformasi Wavelet Daubechies Pada Kompreksi Citra Digital, *CAUCHY*, Volume 2, Issue 4.
- [9] Albergtsen, F., 2008, Statistical Texture Measure Computed from Gray Level Matrices, *Technical Note*, pp. 1-14.
- [10] Fajri, H.R., Nugroho, H.A., Soesanti, Indah, 2015, Ekstraksi Ciri Berbasis Wavelet dan GLCM untuk Deteksi Dini Kanker Payudara Pada Citra Mammogram, *Seminar Nasional Sains dan Teknologi*.
- [11] Prasetyo, E., 2012, *Data Mining Konsep dan Aplikasi Menggunakan Matlab*, Yogyakarta : Andi
- [12] Rochmad, M., 2009, Identifikasi Kerusakan Pankreas Melalui Iridology Menggunakan Metode Bayes Untuk Pengenalan Diabetes Melitus, *Seminar Nasional Informatika*, Volume 1, Issue 1, pp. A33-A42.
- [13] Mitchell, T., 1997, *Machine Learning*, New Delhi : McGrawHill