

Identifikasi Keberadaan Tumor Pada Citra Mammografi Menggunakan Metode Run Length

Imam Santoso
Achmad Hidayatno
Andrio Ghara Pratama

Abstract : Breast cancer is one of the most cancer disease among women until now. This cancer was formed by abnormal cells in breast tissue. For early breast cancer detection the mammography is used by the radiologist. In mammography the breast tissue is scanned by X-rays then a resulting image called mammogram produce. The radiologist manually examine the mammogram identifying which area of scanned breast tissue that could have cancer suspect. In medical term the cancer appear as a mass or microcalcification. Because of the quality of mammograms and small of cancer area in early stadium the radiologist sometime have the problem to decide if there is a cancer perform as a mass or microcalcification.

According to mammograms that can be viewed as texture image, the image processing especially texture analysis computerizely helped the radiologist to classify the texturize cancer area. In this research one of the texture analysis methods named run length method use to get the features. With this features, k-NN (k-Nearest Neighbour) classifier then decide whether the suspect area is cancer according to the mass or microcalcification appeareance or just normal tissue.

As the result there is 72% correct identification among the mammograms data that have been analyzed texturizely with run length method.

Kanker payudara merupakan jenis kanker yang paling umum diderita oleh wanita saat ini. Penyebabnya belum dapat ditemukan secara pasti, serta diduga berkaitan dengan gen pada inti sel jaringan payudara itu sendiri yang mengontrol pertumbuhannya. Meski hanya disebabkan oleh sifat genetik, kemungkinannya hanya 5%-10 % sel kanker yang diturunkan dari ayah atau ibu. Pada umumnya kanker terjadi karena kelainan genetik yang disebabkan oleh factor penuaan atau gaya hidup.

Mammografi adalah hasil pemeriksaan radiologis khusus menggunakan sinar X dosis rendah untuk mengidentifikasi adanya kanker pada jaringan payudara, bahkan sebelum adanya perubahan yang terlihat pada payudara atau

benjolan yang dirasakan pasien. Mammografi dianggap sebagai senjata yang paling efektif untuk mengidentifikasi dan mendeteksi adanya kanker pada payudara, hal ini disebabkan tingkat akurasi yang mencapai hampir 80%-90% dari semua kasus kanker payudara. Mammografi tidak mencegah atau bahkan mengobati, namun dapat mengurangi resiko terjadinya kematian dengan mengidentifikasi keberadaan tumor pada jaringan payudara dalam tingkat yang masih dapat ditangani dengan lebih mudah.

Dalam penelitian ini akan digunakan salah satu dari metode analisis tekstur yaitu, metode *Run Length*, sebagai dasar pengembangan untuk mengidentifikasi dan

Imam Santoso, Achmad Hidayatno (immstso, achmad@elektro.ft.undip.ac.id), adalah dosen di Jurusan Teknik Elektro Fakultas Teknik Universitas Diponegoro (Undip) Semarang Jl. Prof. Sudharto, S.H. Tembalang Semarang 50275

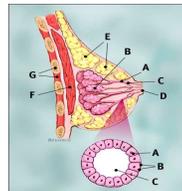
Andrio G.P. adalah mahasiswa di Jurusan Teknik Elektro Fakultas Teknik Universitas Diponegoro (Undip) Semarang Jl. Prof. Sudharto, S.H. Tembalang Semarang 50275

mendeteksi area yang merupakan potensi kanker dari citra mammogram.

Tujuan dari penelitian ini ialah untuk membuat sebuah sistem yang dapat mengidentifikasi dan mendeteksi adanya area pada citra mammografi yang merupakan potensi dari kelainan atau kanker.

Teori Kanker Payudara

Kanker payudara merupakan jenis kanker yang paling umum diderita oleh wanita saat ini. Penyakit ini terjadi dimana sel-sel tidak normal (kanker) terbentuk pada jaringan payudara. Diperkirakan satu diantara sepuluh sampai duabelas wanita harus berhadapan dengan kanker payudara (Timp 2006). Kanker payudara juga terjadi pada pria, meskipun sangat jarang ditemukan. Usia rata-rata wanita ketika terdiagnosis mengidap kanker payudara adalah 64 tahun, sementara sepertiga dari seluruh wanita terdiagnosis menderita penyakit ini saat berusia dibawah 50 tahun.



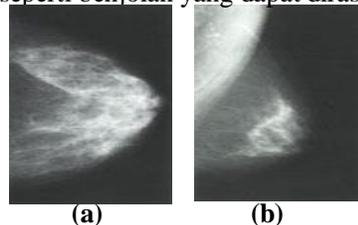
Gambar 1. Bagian-bagian payudara

Bagian-bagian payudara:

- A. puting Inset : A. sel-sel *duct* normal
- B. *duct* B. membran dasar
- C. *lobula* C. *lumen* (pusat *duct*)
- D. bagian *duct* yang menahan susu
- E. lemak
- F. otot pektoral
- G. dinding dada

Mammografi

Mammografi adalah pemeriksaan radiologi khusus menggunakan sinar-X dosis rendah untuk mendeteksi kelainan pada payudara, bahkan sebelum adanya gejala yang terlihat pada payudara seperti benjolan yang dapat dirasakan.

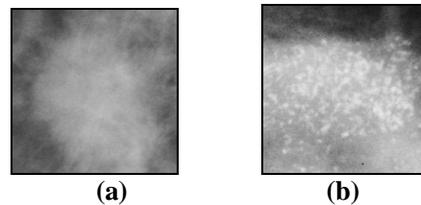


Gambar 2. (a) posisi *cranio-caudal*
(b) posisi *medio-lateral oblique*

Jenis Kanker Payudara

Secara mammografi, kanker payudara dikenali dengan keberadaan lesi massa atau biasa disebut massa, atau keberadaan mikrokalsifikasi.

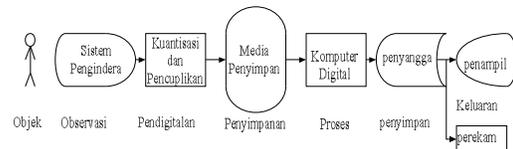
- Lesi massa. Sebagian besar tumor payudara, baik yang tidak ganas maupun berpotensi kanker tampak sebagai massa. Sebuah 'massa' adalah area terdapatnya lesi yang tampak dari dua proyeksi foto mammografi yang berbeda.
- Mikrokalsifikasi. Ciri lainnya dari kanker adalah keberadaan mikrokalsifikasi. Mikrokalsifikasi berbentuk seperti noda berukuran kecil dan terkadang berupa titik-titik, terdapat di dalam lobula atau ductal.



Gambar 3. (a) citra massa
(b) citra mikrokalsifikasi

Definisi Pengolahan Citra

Pengolahan Citra merupakan proses pengolahan dan analisis citra yang banyak melibatkan persepsi visual. Proses ini mempunyai ciri data masukan dan informasi keluaran yang berbentuk citra. Dalam definisi yang lebih luas, pengolahan citra digital juga mencakup semua data dua dimensi. Citra digital adalah barisan bilangan nyata maupun kompleks yang diwakili oleh bit-bit tertentu. Proses pengolahan citra digital dapat digambarkan dengan Gambar 2.4.



Gambar 2.4 Diagram alir pengolahan citra digital

Proses Pra-pengolahan

Pada Tahap ini diperlukan proses perbaikan citra untuk meningkatkan kualitas citra dalam memperlihatkan ciri-cirinya.

Klasifikasi Citra

Proses analisis tekstur menggunakan metode *run length* yang dilanjutkan dengan proses klasifikasi k-NN.

Perbaikan Kualitas Citra

Beberapa cara untuk mengurangi derau pada citra diantaranya menggunakan tapis linear, menggunakan tapis median serta menggunakan tapis adaptif. Tapis linear cukup efektif untuk mengurangi derau berupa bulir-bulir atau bercak dari citra yang berasal dari fotografi. Filter adaptif efektif untuk mengurangi derau Gaussian. Median filter lebih baik digunakan untuk mengurangi derau berupa bercak hitam putih, maupun noda (*speckle*), yang dikenal dengan derau *salt and pepper*.

Citra radiografi memiliki kontras yang rendah sehingga menyebabkan kesulitan pada proses pengolahannya. Tujuan dari perbaikan kontras adalah untuk memperbaiki kualitas citra radiografi.

Operasi morfologi adalah bagian pengolahan citra yang mengolah citra berdasarkan bentuk. Dalam operasi morfologi, nilai dari tiap piksel pada citra keluaran didasarkan pada perbandingan piksel yang bersesuaian pada citra masukan terhadap piksel-piksel tetangganya. Karena sifat tersebut operasi morfologi biasa digunakan dalam pengolahan citra berdasarkan analisis tekstur.

Analisis Tekstur

Salah satu teknik dalam pengolahan citra digital adalah analisis tekstur. Teknik ini berkaitan dengan pengolahan parameter tekstur pada citra digital. Meskipun tidak ada definisi tekstur secara matematis yang dapat diterima, dapat dikatakan tekstur merupakan pola berulang pada piksel-piksel dalam citra. Pola yang dimaksud seperti bentuk, kedalaman, warna, kecerahan dan sebagainya. Tekstur dapat diamati dalam bentuk pola terstruktur pada permukaan suatu benda seperti kayu, kain, tanah, pasir, padang rumput, hutan, air, dan lain-lain. Tekstur alami umumnya mempunyai pola acak, sedangkan yang buatan seringkali berpola deterministik atau periodik.

Pencirian Tekstur

Ada dua pendekatan yang digunakan untuk pencirian tekstur dari suatu area, yaitu statistik dan struktural. Pendekatan statistik mempertimbangkan parameter tekstur, distribusi intensitas pada piksel-piksel, serta hubungan antar piksel bertetangga. Metode ini berdasar pada frekuensi spasial dan menghasilkan karakterisasi tekstur seperti halus, kasar, berbulir, dan lain-lain. teknik struktural

berkaitan dengan penyusunan bagian-bagian terkecil (primitif) suatu citra, contoh metode ini adalah model fractal.

GLRL (Grey Level Run Length Matriks)

Metode *Run Length* menggunakan pendekatan statistik. Metode ini menggunakan distribusi suatu *pixel* dengan intensitas yang sama secara berurutan dalam satu arah tertentu sebagai primitifnya. Masing-masing primitif didefinisikan atas arah, panjang dan level keabuan. Panjang dari primitif tekstur pada arah yang berbeda dapat digunakan untuk menggambarkan suatu tekstur.

Untuk melakukan ekstraksi ciri dengan metode *run length*, citra aras keabuan dengan matriks $f(x,y)$ harus ditransformasikan terlebih dahulu kedalam matriks *grey level run length* (GLRL), $B(a,r)$.

$$f(x, y) \xrightarrow{GLRL} B(a, r) \tag{1}$$

Elemen matriks dari GLRL $B(a,r)$ menghitung banyaknya primitif dengan panjang r dan level keabuan a . Jumlah dari primitif dapat diperoleh dengan persamaan :

$$K = \sum_{a=1}^L \sum_{r=1}^{Nr} B(a, r) \tag{2}$$

Dengan, L : banyaknya level keabuan dari citra

Nr : panjang maksimal dari primitif

K : jumlah primitif

M,N : dimensi citra

0	1	2	3
0	2	3	3
2	1	1	1
3	0	3	0

Adapun ciri dari tekstur dapat diperoleh dari persamaan-persamaan berikut ini :

1. Short Run Emphasis (SRE)

SRE mengukur distribusi *short run*. SRE sangat bergantung pada banyaknya *short run* dan diharapkan bernilai kecil pada tekstur halus dan bernilai besar pada tekstur kasar.

$$SRE = \frac{1}{K} \sum_{a=1}^L \sum_{r=1}^{Nr} \frac{B(a, r)}{r^2} \tag{3}$$

2. Long Run Emphasis (LRE)

LRE mengukur distribusi *long run*. LRE sangat bergantung pada banyaknya *long run* dan diharapkan bernilai besar pada tekstur halus dan bernilai kecil pada tekstur kasar.

$$LRE = \frac{1}{K} \sum_{a=1}^L \sum_{r=1}^{Nr} r^2 B(a,r) \tag{4}$$

1. Grey Level Uniformity (GLU)

GLU mengukur persamaan nilai derajat keabuan diseluruh citra dan diharapkan bernilai kecil jika nilai derajat keabuan serupa di seluruh citra.

$$GLU = \frac{1}{K} \sum_{a=1}^L \left(\sum_{r=1}^{Nr} B(a,r) \right)^2 \tag{5}$$

2. Run Length Uniformity (RLU)

RLU mengukur persamaan panjangnya *run* diseluruh citra dan diharapkan bernilai kecil jika panjangnya *run* serupa di seluruh citra.

$$RLU = \frac{1}{K} \sum_{r=1}^{Nr} \left(\sum_{a=1}^L B(a,r) \right)^2 \tag{6}$$

3. Run Percentage (RPC)

RPC mengukur keserbasamaan dan distribusi *run* dari sebuah citra pada arah tertentu. RPC bernilai paling besar jika panjangnya *run* adalah 1 untuk semua derajat keabuan pada arah tertentu.

$$RPC = \frac{K}{\sum_{a=1}^L \sum_{r=1}^{Nr} r B(a,r)} = \frac{K}{MN} \tag{7}$$

Citra aras keabuan $f(x, y)$

0	1	2	3
0	2	3	3
2	1	1	1
3	0	3	0

Run Length $B(a,r)$

0^0	1	2	3	4
0	4	0	0	0
1	1	0	1	0
2	3	0	0	0
3	3	1	0	0

Gambar 2.5 Contoh Citra dan matriks *run length* pada arah 0°

Klasifikasi

Jumlah Kesalahan k-NN	Normal	Terdapat Massa	Terdapat at Mikrokalsifikasi	Jumlah Kesalahan
k = 1	5	9	1	15
k = 2	5	8	1	14
k = 3	5	9	1	15
k = 4	8	10	0	18
k = 5	7	7	0	14
Jumlah sampel	15	15	20	

Tujuan dari proses klasifikasi citra adalah untuk mendapatkan gambar atau peta tematik. Gambar tematik adalah gambar yang terdiri dari bagian-bagian yang menyatakan suatu objek atau tema. Setiap objek pada gambar tersebut mempunyai simbol yang unik yang dapat dinyatakan dengan warna atau pola tertentu. Sebagai contoh, citra hasil suatu foto instrumen biomedis dapat diklasifikasi menjadi bagian tulang, jaringan sehat, dan jaringan sakit. Proses klasifikasi citra dilakukan dengan memasukkan setiap piksel citra tersebut ke dalam suatu kategori objek yang sudah diketahui

Metode k-Nearest Neighbour (k-NN)

Metode k-NN menyediakan training set untuk mengarahkan atau memberi informasi atau pengetahuan dalam ciri statistik tentang kelas-kelas yang ada, dengan kata lain klasifikasi berdasarkan pada mayoritas kelas k terdekat dari objek yang dianalisis. Metode klasifikasi ini tidak menggunakan model sebagai dasar penyesuaian, dan hanya berdasarkan pada memori. Pada objek yang akan diklasifikasi, dicari sejumlah objek pada training set yang terdekat dengan objek yang akan diklasifikasikan. Metode ini sangat sederhana dan bersifat intuitif.

Jarak Euclidean

Metode k-NN dapat digunakan pada berbagai macam jenis data seperti data kuantitatif, data nominal, dan data ordinal. Rumus jarak yang digunakan dalam perhitungan berbeda untuk masing-masing jenis data. Menurut Kardi Teknomo dalam websitenya (<http://people.revoledu.com>), untuk perhitungan data kuantitatif, jarak yang umumnya digunakan adalah jarak Euclidean. Jarak Euclidean dapat dituliskan dalam persamaan:

$$d_{s,l} = \sqrt{\sum_k (s_k - l_k)^2} \tag{8}$$

dengan: k = jumlah variabel
 s = variabel sampel uji
 l = variabel sampel latih

Pengujian Pada Data Latih

Dari pengujian yang dilakukan, didapatkan nilai rata-rata untuk setiap ciri tekstur dari semua citra latih normal, massa maupun mikrokalsifikasi.

Tabel 1. Nilai rata-rata untuk setiap ciri tekstur

	Normal	Massa	Mikrokalsifikasi
SRE	0.614493	0.660385	0.754647
LRE	5741.956	431.8126	74.84809
GLU	9497.05	27230.79	44832.79
RLU	1172.254	1340.899	1646.533
RPC	0.146217	0.310277	0.511561

Pengujian Pada Data Uji

Pengujian dilakukan dengan menggunakan metode klasifikasi k-NN dan didapatkan jumlah percobaan yang salah sebagai berikut :

Dari Pengujian didapatkan rasio pengenalan untuk k bernilai 1, 2, 3, 4 dan 5 secara berurutan adalah 70%, 72%, 70%, 64% dan 72%. Sedangkan untuk rasio kesalahan untuk bernilai 1, 2, 3, 4 dan 5 secara berurutan adalah 30%, 28%, 30%, 36% dan 28%. Pada umumnya dengan semakin besarnya daerah pencarian ciri citra latih yang terdekat dengan sampel uji, maka tingkat kesalahan akan meningkat. Hal ini tidak muncul pada hasil pengujian disebabkan untuk menekan tingkat kesalahan saat semakin besarnya daerah pencarian, variasi nilai-nilai ciri tekstur sebagai acuan klasifikasi untuk tiap kategori diperbanyak jumlahnya yaitu dengan menggunakan jumlah citra latih yang besar. Citra latih berjumlah 105 citra masing-masing 35 ROI terdapat massa tumor, 35 ROI terdapat mikrokalsifikasi dan 35 ROI normal. Metode klasifikasi k-NN akan semakin baik dengan jumlah data latih yang semakin besar.

Kesimpulan

Berdasarkan hasil pengujian dan pembahasan dapat disimpulkan beberapa hal sebagai berikut:

1. Proses perbaikan citra merupakan salah satu faktor penting didalam tugas akhir ini, hal ini dikarenakan citra mammografi asli memiliki kontras yang rendah dan akan menyebabkan kesulitan dalam proses analisis tekstur maupun pada proses klasifikasi.

2. Analisis tekstur menggunakan metode *run length* dapat digunakan untuk mendapatkan ciri tekstur yang akan dijadikan parameter dalam proses klasifikasi dengan k-NN.
3. Dibutuhkan data base yang besar untuk memudahkan program dalam proses klasifikasi.
4. Nilai SRE (*short run emphasis*) akan bernilai kecil pada tektur halus dan bernilai besar pada tektur kasar.
5. Sebaliknya nilai LRE (*long run emphasis*) akan bernilai besar pada tekstur halus dan bernilai kecil pada tekstur kasar.
6. GLU (*grey level uniformity*) mengukur persamaan nilai derajat keabuan diseluruh citra, sehingga akan bernilai kecil, jika nilai derajat keabuan diseluruh citra lebih serupa (tekstur halus).
7. RLU (*run length uniformity*) mengukur panjangnya *run* diseluruh citra, sehingga akan bernilai kecil, jika panjang *run* diseluruh citra serupa (tekstur halus).
8. Pada proses klasifikasi k-NN, rasio pengenalan untuk k bernilai 1, 2, 3, 4 dan 5 secara berurutan adalah 70%, 72%, 70%, 64% dan 72%. Sedangkan untuk rasio kesalahan untuk bernilai 1, 2, 3, 4 dan 5 secara berurutan adalah 30%, 28%, 30%, 36% dan 28%.

Saran

1. Proses perbaikan kualitas citra memegang peranan penting dalam mendapatkan ciri tekstur yang khas dari citra mammografi, sehingga diperlukan penelitian lebih lanjut untuk mendapatkan metode baru yang lebih cocok untuk digunakan pada citra mammografi.
2. Penelitian dapat dilanjutkan dengan metode analisis tekstur yang lain seperti metode GLCM (*grey level co-occurrence matrices*), metode auto korelasi dan metode *law*, kemudian dapat ditentukan metode mana yang lebih cocok untuk digunakan pada citra mammografi.
3. Metode klasifikasi k-NN akan semakin baik dengan penambahan data citra latih, sementara pengujian terhadap lebih banyak citra uji dapat dilakukan untuk mendapatkan nilai k yang memberikan hasil klasifikasi yang lebih baik.

DAFTAR RUJUKAN

- Teknomo, K., *Kardi Teknomo's Page – k Nearest Neighbors Tutorial*, <http://people.revoledu.com/kardi/tutorial/kNN>
- Timp, Sheila., *Analysis of Temporal Mammogram Pairs to Detect and Characterise Mass Lesions*, Groningen, 2006
http://webdoc.uhn.ru.nl/mono/t/timp_s/analoftem.pdf www.breastcancer.org
- The MathWorks. Inc , *Image Processing Toolbox 5.1*, Matlab 7.1 SP3
- Dharmais Cancer Hospital – National Cancer Center, <http://www.dharmais.co.id>
- Healthwise – National Cancer Institute, *Breast cancer: Treatment - Patient Information [NCI PDQ]*,
- Munir, R., *Pengolahan Citra Digital dengan Pendekatan Algoritmik*, Informatika, Bandung, 2004 [htm](#)
- Indriani, Mita, *Klasifikasi Tekstur Menggunakan Metode LDA dan kNN Berdasarkan Pencirian Metode Run-Length*, Tugas Akhir Jurusan Teknik Elektro Undip, 2007