

DETEKSI KANKER KOLOREKTAL (KANKER USUS BESAR) MENGGUNAKAN METODE GRAY LEVEL COOCCURENCE MATRIX DAN K-NEAREST NEIGHBOR BERBASIS PENGOLAHAN CITRA

(Detection of Colorectal Cancer (Colon Cancer) Using Methods Gray Level Cooccurrence Matrix and K-Nearest Neighbor Based Image Processing)

Firda Masitha¹, Ratri Dwi Atmaja, S., T.M.T.², Unang Sunarya, S.T., M.T.³
¹²³Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

¹Firdamasitha0@gmail.com

²ratriidwiatmaja@telkomuniversity.ac.id,

³unangsunarva@telkomuniversity.ac.id

Abstrak

Kanker kolorektal atau dapat disebut juga dengan kanker usus besar merupakan jenis kanker ganas yang terjadi ketika sel-sel abnormal tumbuh pada lapisan usus besar (kolon) atau *rectum* (bagian usus paling bawah sampai anus). Didunia kedokteran sering menjumpai beberapa jenis kanker usus besar seperti, *sarcoma*, *carcinoma*, dan *lymphoma*. Deteksi dan klasifikasi terhadap kanker sangatlah penting untuk pencegahan dan penanganan dini sehingga kanker tidak meningkat menjadi stadium yang lebih tinggi. Deteksi kanker yang dilakukan oleh dokter yaitu melalui tes darah atau sampel jaringan.

Tugas akhir ini bertujuan untuk memudahkan pendeteksian penyakit kanker usus besar berbasis pengolahan citra dengan bantuan *software* Matlab 2015a. Ekstraksi ciri yang digunakan adalah *Gray Level Cooccurrence Matrix* yaitu metode untuk membentuk fitur atau ciri yang tidak hanya didasarkan pada nilai pikselnya namun juga menggunakan penciri tekstur. Serta diklasifikasikan menggunakan *K-Nearest Neighbor*.

Proses pengujian citra dimulai dari ekstraksi ciri yaitu pengujian terhadap sudut orientasi dan jarak pada GLCM. Dilanjutkan dengan pengujian terhadap *pre-procesing* nya yaitu dengan menguji beberapa ukuran *resize*. Kemudian pada proses klasifikasi terdapat dua pengujian, yaitu jenis *distance* dan nilai variable K pada K-NN. Terdapat pengujian tambahan, yaitu pengujian terhadap kombinasi 1,2 dan 3 dari 4 parameter yaitu *contrast*, *correlation*, *energy* dan *homogeneity*. Sistem yang sudah diujikan mendapat akurasi yang paling baik, yaitu 75% terhadap jumlah data latih sebanyak 138 citra dan data uji sebanyak 60 citra yang diklasifikasikan menjadi tiga kelas, yaitu normal, kanker *carcinoma*, dan kanker *lymphoma*.

Kata Kunci : Kolorektal, *Gray Level Cooccurrence Matrix*, *K-Nearest Neighbor*, ekstraksi ciri, klasifikasi citra

Abstract

Colorectal cancer or may be referred to as colon cancer is an aggressive type of cancer that occurs when abnormal cells grow in the lining of the large intestine (colon) or the rectum (ileum to the anus). There are several types of cancer in medical world such as colon, sarcoma, carcinoma, and lymphoma. Detection and classification of cancer is important for the prevention and early treatment so that the cancer will not increased to the higher stage.. Detection of cancer by doctors is through blood or tissue samples.

This final project aims to facilitate the detection of colon cancer based on image processing using Matlab 2015a. Feature extraction that used in this studies is Gray Level Cooccurrence Matrix.it's a method for forming a feature that are not only based on the value of pixels but also using the identifier of the texture. As well as classified using K-Nearest Neighbor.

The testing process starts from the image feature extraction is test the orientation angle and distance at GLCM. Followed by testing is pre-procesing to test some resize. Then in the classification process, there are two tests, that kind of distance and the value of the variable K in K-NN. There is additional testing, the testing of the combination of 1,2 and 3 of the 4 parameters that is contrast, correlation, energy and homogeneity. The system has been tested and got the accuracy of 75% from 138 images of the training data and 60 images of the test data that are classified into three classes, normal, carcinoma cancer, and lymphoma cancer.

Keywords: Kolorektal, Gray Level Cooccurrence Matrix, K-Nearest Neighbor, feature extraction, image classification

1. Pendahuluan

Kanker merupakan salah satu penyebab utama mortalitas di dunia (sekitar 13% dari seluruh penyebab mortalitas), diperkirakan angka mortalitas sekitar 7,9 juta kematian pada tahun 2007 dan menurut WHO kanker

usus besar sejumlah 677.000 mortalitas/tahunnya [1]. Kanker usus besar (*colorectal cancer*) adalah keganasan yang menyerang usus besar (kolon) dan rektum (bagian usus paling bawah hingga anus) yang merupakan salah satu jenis kanker yang menduduki peringkat 3 terbanyak yang diderita di Indonesia baik perempuan maupun laki-laki pada rentang usia diatas 50 tahun [2].

Kanker yang terjadi pada usus besar seperti *Lymphoma* dan *Carcinoma* memiliki karakteristik yang berbeda sehingga dapat menjadi tanda bahwa seseorang mengidap kanker usus besar. Pengklasifikasian kanker usus besar sejak awal memungkinkan penyembuhan dan pencegahan kanker yang mudah dan murah bagi penderita. Karena itu, perlu dilakukan deteksi dini secara rutin untuk pencegahan dan pengobatan dari kanker usus besar. Metode deteksi dini kanker usus besar dapat dilakukan dengan metode manual, yaitu sel yang diletakkan diatas preparat dilihat melalui mikroskop untuk menentukan apakah sel tersebut terdapat kanker dengan melihat abnormalitasnya dengan membandingkan sel yang diamati dengan sel yang sehat. Pendiagnosaan dengan cara tersebut sangat berhubungan dengan penglihatan dokter, dimana kesalahan atau ketidak fokusan sangat mempengaruhi hasil diagnosa [2].

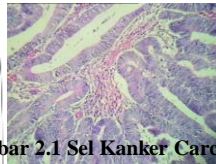
Oleh karena itu, pada tugas akhir ini dibuat suatu sistem yang mampu mendeteksi dini kanker usus besar dengan memanfaatkan citra hasil pemeriksaan. Ekstrasi fitur yang di gunakan adalah Gray Level Co-occurrence Matrix (GLCM), yang dimana dalam penelitian tersebut terdapat 4 fitur GLCM yaitu *contrast*, *correlation*, *energy*, dan *homogeneity*. Dan *K-Nearest Neighbor* (KNN) untuk pengklasifikasiannya. Klasifikasi pada sistem ini dibagi menjadi tiga kelas, yaitu normal, kanker *carcinoma*, dan kanker *lymphoma*.

2. Dasar Teori

2.1 Kanker Usus Besar

a. Carcinoma

Carcinoma adalah kanker yang dimulai di jaringan yang melapisi permukaan dalam atau luar tubuh, umumnya muncul dari sel-sel yang berasal dari lapisan kuman endodermal atau ectodermal selama embriogenesis. Lebih khususnya, *carcinoma* adalah jaringan tumor berasal dari sel epitel yang berubah atau rusak sehingga sel-sel tersebut menjadi berubah dan mulai menunjukkan sifat ganas abnormal. Kanker ini umumnya menyerang manusia yang sudah lanjut usia dan sangat jarang terjadi pada anak-anak [4].



Gambar 2.1 Sel Kanker Carcinoma

b. Lymphoma

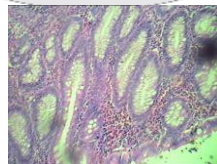
Lymphoma adalah kanker pada sel-sel limfatik dari sistem kekebalan tubuh. Biasanya, *lymphoma* hadir sebagai soloid tumor sel limfoid. Pengobatan mungkin melibatkan kemoterapi dan dalam beberapa kasus radioterapi dan/atau transplantasi sumsum tulang belakang, dan dapat disembuhkan tergantung pada histologi, jenis dan tahap penyakit.



Gambar 2.2 Sel Kanker Lymphoma

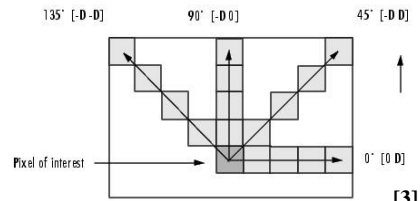
c. Usus Normal

Berikut adalah contoh gambar usus normal :



Gambar 2.3 Usus Normal

2.2 Gray Level Cooccurrence Matrix



Gambar 2.6 Arah orientasi sudut GLCM [3]

GLCM adalah matriks derajat keabuan yang merepresentasikan hubungan suatu derajat keabuan dengan derajat keabuan lain. GLCM merupakan dasar dari teknik tekstur Haralick, GLCM digunakan untuk analisis pasangan piksel yang bersebelahan tergantung dengan sudut yang digunakan. Apabila citra yang digunakan adalah citra biner maka yang digunakan adalah GLCM dua level. Setelah itu matriksnya dinormalisasi dengan menghitung peluang nilai piksel berdekatan dibagi dengan jumlah semua peluang nilai piksel berdekatan, sehingga hasil penjumlahan piksel dalam matriks tersebut 1 (Suresh dan Shunmuganathan 2012).

Adapun 4 persamaan tekstur fitur yang akan digunakan dalam perancangan system pencarian image menggunakan inputan Image, yaitu: [5]

a. *Contrast*

Kontras adalah perhitungan perbedaan intensitas antara piksel satu dan piksel yang berdekatan diseluruh gambar. Kontras bernilai nol untuk gambar yang konstan [7].

$$\sum_{i,j} |i - j|^2 P(i,j) \quad (2.1)$$

b. *Correlation*

Korelasi merupakan ukuran ketergantungan linier antar nilai aras keabuan dalam citra dihitung dengan menggunakan rumus:[7]

$$\sum_{i,j} \frac{(i - \mu_i)(j - \mu_j) P(i,j)}{\sigma_i \sigma_j} \quad (2.2)$$

c. *Homogeneity*

Homogenitas menunjukkan nilai distribusi terdekat antara elemen di GLCM dengan GLCM diagonal. Homogenitas bernilai satu untuk diagonal GLCM.[7]

$$\sum_{i,j} \frac{1}{1 + |i - j|} P(i,j) \quad (2.3)$$

d. *Energy*

Energy atau yang biasa disebut dengan *Angular Second Moment (ASM)* merupakan ukuran homogenitas citra dihitung dengan cara seperti berikut: [7]

$$\sum_{i,j} P(i,j)^2 \quad (2.4)$$

2.3 Proses Klasifikasi

Untuk melakukan proses deteksi, pada tugas akhir ini menggunakan *K- Nearest Neighbor (KNN)*. Algoritma *K-Nearest Neighbor (K-NN)* adalah metode yang digunakan untuk mengklasifikasikan suatu data. Pada data latih biasanya diambil lebih dari satu tetangga terdekat dengan data uji kemudian akan digunakan algoritma ini untuk ditentukan kelasnya. K-NN disebut juga teknik *lazy learning* yang proses klasifikasinya langsung berdasarkan pada data-data latih. Teknik *lazy learning* juga disebut dengan klasifikasi berdasarkan contoh atau klasifikasi berdasarkan kasus.

Klasifikasi K-NN mempunyai dua langkah, yaitu :

- Menentukan tetangga – tetangga terdekat dari data tersebut.
- Menentukan kelas dari masing – masing tetangga terdekat tersebut.

Dalam tugas akhir ini pada klasifikasi K-Nearest Neighbor, perhitungan jarak terdekat yang digunakan adalah *Euclidean Distance*, *City Block*, *Cosine*, dan *Minkowski*.

Berikut ini adalah rumus secara sistematis dari klasifikasi KNN :

1. *Euclidean Distance*, dengan rumus

$$l_2(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.5)$$

◆=1

Dimana $L_2(X,Y)$ = jarak dua gambar dengan *Euclidean Distance*.

X_i = nilai pada citra latih Y_i = nilai pada citra uji d = jumlah data

2. *City block* atau *manhattan distance*, dengan rumus :

$$l1(x, y) = \sum |x_i - y_i| \quad (2.6)$$

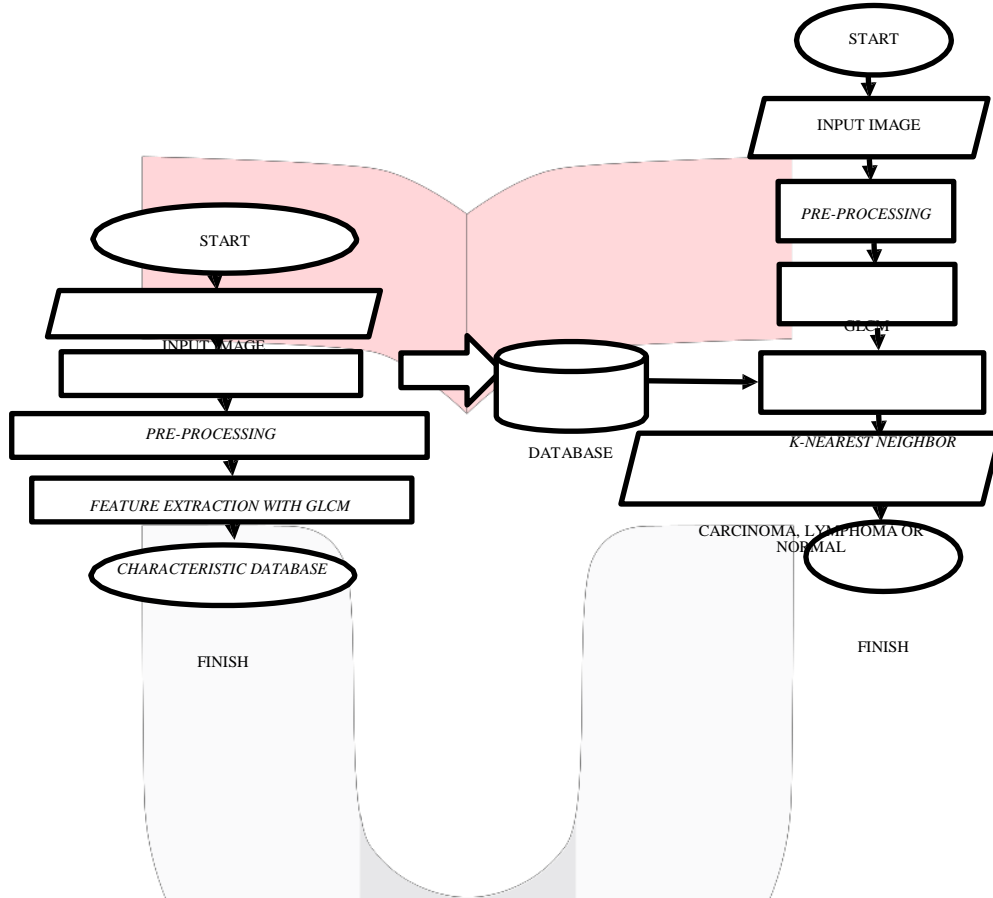
3. Cosine, dengan rumus :

$$\cos(\theta) = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2} \sqrt{\sum y_i^2}} \quad (2.7)$$

4. Minkowski, dengan rumus :

$$\|x - y\|_q = (\sum |x_i - y_i|^q)^{\frac{1}{q}} \quad (2.8)$$

Dimana $q \geq 1$ adalah parameter yang bisa diseleksi. Apabila $q=1$, maka jarak tersebut menjadi jarak *Manhattan*.
 Sedangkan $q=2$, jarak tersebut menjadi jarak *Euclidean*.
 untuk



Gambar 3.1 Flowchart Sistem Secara Umum

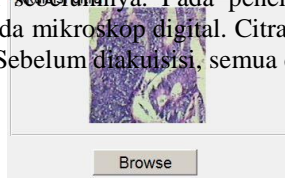
3. Perancangan Sistem dan Simulasi

Pada awalnya citra mengalami proses digitalisasi melalui kamera digital didapatkan citra berwarna berukuran 768 x 576 pixel yang berformat *.jpeg. Sebelum di *preprocessing* citra dicrop menjadi berukuran 300x300 piksel menggunakan paint.

3.1 Proses Pre-processing

1. Akuisisi Citra

Citra usus besar diperoleh dari penelitian sebelumnya. Pada penelitian sebelumnya pengambilan citra menggunakan kamera digital yang telah terpasang pada mikroskop digital. Citra yang dihasilkan yaitu citra berwarna dengan format *.jpeg dan berukuran 768x576 piksel. Sebelum diakuisisi, semua data citra diubah ukurannya menjadi 300x300 dengan aplikasi *paint*.



Gambar 3.6 Hasil Akuisisi Citra

2. *Resize*

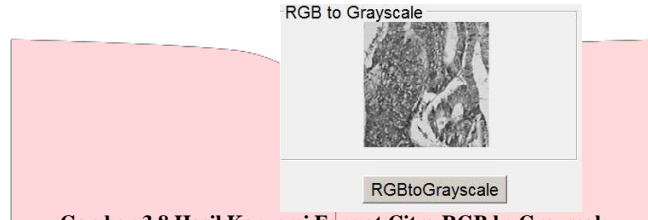
Setelah akuisisi citra , selanjutnya adalah pengujian pada beberapa ukuran piksel yang diturunkan setiap 20%, yaitu 300x300, 240x240, 192x192, dan 154x154.



Gambar 3.7 Hasil Resize

3. Grayscale

Tahap terakhir pada *pre-processing* yaitu perubahan format citra dari citra RGB menjadi citra dengan level keabuan atau dikenal dengan citra *grayscale*. Proses ini bertujuan mempermudah pengolahan informasi ke tahap selanjutnya.



Gambar 3.8 Hasil Konversi Format Citra RGB ke Grayscale

3.2 Ekstraksi Ciri Menggunakan GLCM

Ekstraksi ciri dalam kasus ini berfungsi sebagai klasifikasi dan interpretasi citra awal. Dalam ekstraksi ciri ini sendiri di ambil ciri-ciri khusus setiap gambar sehingga mempermudah dalam proses selanjutnya.[6] GLCM ini sendiri merupakan tahapan ekstraksi ciri yang dimana memakai sudut 0°, 45°, 90°, dan 135° dalam mengambil ciri-ciri dan pengujian pada sistem.

3.3 Klasifikasi Menggunakan K-NN

Setelah dilakukan proses ekstraksi ciri, selanjutnya masuk ke dalam proses klasifikasi ciri dengan K-Nearest Neighbor (KNN). Klasifikasi ciri ini bertujuan untuk mengklasifikasikan citra usus sehingga dapat mengeluarkan output yang tepat sesuai target yang ditentukan. Klasifikasi ini pada dasarnya adalah pengelompokan menjadi 3 kelas yaitu kanker *carcinoma*, kanker *lymphoma*, dan normal. Pada tahap latih dan uji hasil ekstraksi ciri disimpan dalam sebuah *database* yang kemudian dihitung jarak terdekat berdasarkan jarak antara data yang akan diuji dengan hasil tahap latih sebelumnya.

Berikut tahapan klasifikasi K-NN:

1. Menentukan jenis *distance* yang paling baik digunakan antara *Euclidean*, *cityblock*, *cosine*, atau *minkowski*.
2. Menentukan variable K yang paling baik antara 1, 3, 5, 9
3. Menggunakan jenis *distance* yang paling baik sesuai hasil pengujian
4. Menghitung jarak titik data uji dengan titik data *database* terdekat
5. Menentukan jumlah titik dari masing-masing kelas *database* terdekat yang terbanyak tergantung nilai variable K
6. Mengeluarkan hasil klasifikasi

Nilai variable K dan *distance* yang menghasilkan nilai akurasi yang paling tinggilah yang paling baik.

3.4 Akurasi

Akurasi merupakan ukuran ketepatan sistem dalam mengenali masukan yang diberikan sehingga menghasilkan keluaran yang benar. Secara matematis dapat dirumuskan sebagai berikut :

$$\frac{\text{Jumlah Data Benar}}{\text{Jumlah Data}} = 100\% \quad (3.1)$$

4. Analisis dan Hasil Pengujian

Masing-masing pengujian memiliki hasil yang direpresentasikan dalam bentuk tabel dan kesimpulan dalam bentuk grafik.

4.1 Analisis Hasil Pengujian

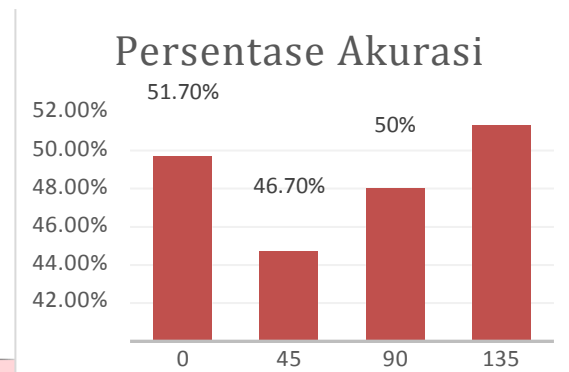
4.1.1 Analisis Hasil Pengujian Skenario 1

Pengujian dilakukan dengan mengubah derajat pada ekstraksi ciri menggunakan *gray level cooccurrence matrix*. Derajat yang digunakan adalah 0°, 45°, 90°, dan 135°.

Tabel 4.1 Tabel Data Benar Setiap Sudut

	NORMAL	LYMPHOMA	CARCINOMA
0°	8	14	9
45°	8	15	5
90°	10	13	7

135°	9	16	7
------	---	----	---



Gambar 4.1 Gambar Persentase Akurasi Setiap Sudut

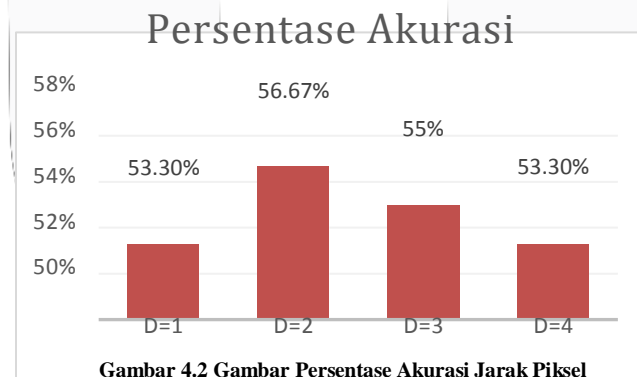
Dapat dilihat pada tabel 4.1 dan gambar 4.1, bahwa perubahan sudut pada ekstraksi ciri *gray level cooccurrence matrix* akan mempengaruhi akurasi sistem pendeteksian meskipun perbedaan dari ke empat percobaan tidak terlalu besar. Hasil akurasi tertinggi terdapat pada 135° dengan nilai akurasi sebesar 53.33%, dan pada ketiga derajat lainnya mengalami penurunan akurasi yaitu 51.7%, 50%, dan 46.7%. Hal ini disebabkan karena sudut orientasi menentukan arah hubungan tetangga dari piksel-piksel referensi dan menentukan nilai parameter pada setiap citra.

4.1.2 Analisis Hasil Pengujian Skenario 2

Pada pengujian skenario kedua dilakukan dengan mengubah jarak antar pikselnya yaitu antara d=1, d=2, d=3 dan d=4 dari sudut yang digunakan yaitu 135° dimana sudut 135° adalah yang mendapatkan akurasi paling baik daripada sudut 0°, 45° dan 90°.

Tabel 4.2 Tabel Data Benar Setiap Jarak Piksel

	NORMAL	LYMPHOMA	CARCINOMA
D=1	9	16	9
D=2	9	16	9
D=3	9	17	7
D=4	7	17	7



Gambar 4.2 Gambar Persentase Akurasi Jarak Piksel

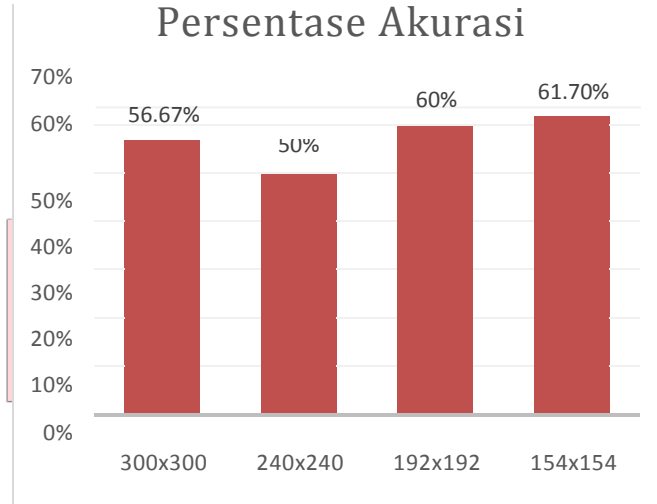
Pada gambar 4.2 dapat dilihat bahwa jarak pada piksel memberikan peningkatan pada optimasi akurasi. Dan yang mendapatkan akurasi terbaik dari skenario kedua ini adalah dengan jarak 2 antar pikselnya dengan nilai akurasi sebesar 56.67% walaupun tidak berbeda jauh dengan nilai akurasi pada jarak lainnya. Pada percobaan ini membuktikan bahwa jarak piksel yang paling dekat yaitu 1, tidak menjamin nilai akurasi yang tinggi. Karena D=2 mendapatkan akurasi yang paling tinggi, maka D=2 akan digunakan menjadi acuan pada skenario selanjutnya.

4.1.3 Analisis Hasil Pengujian Skenario 3

Pada pengujian skenario ketiga dengan menggunakan sudut dengan akurasi terbaik pada skenario pertama yaitu 135° dari skenario yang pertama dan menggunakan jarak antar pikselnya adalah 2 dari skenario kedua. Dilakukan pengujian terhadap pengaruh *resize* dalam *pre-processing* dengan merubah ukuran piksel citra setiap 20% penurunan yaitu menjadi 300x300, 240x240, 192x192, dan 154x154 piksel.

Tabel 4.3 Tabel Data Benar Setiap Ukuran *Resize*

	NORMAL	LYMPHOMA	CARCINOMA
300x300	9	16	9
240x240	10	14	6
192x192	9	17	10
154x154	9	19	5



Gambar 4.3 Gambar Persentase Akurasi Setiap Ukuran *Resize*

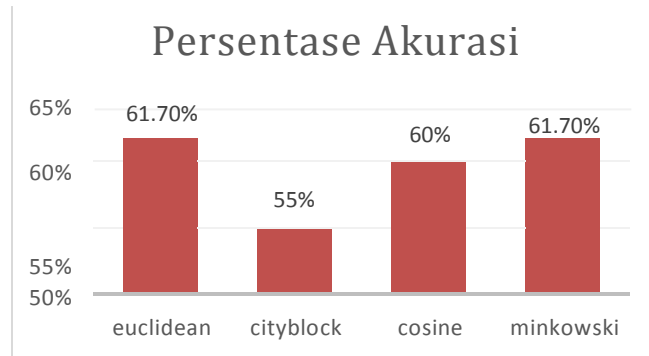
Gambar 4.3 memperlihatkan bahwa *resize* pada *pre-processing* juga menentukan hasil ciri yang diperoleh pada suatu citra. Pengaruh ukuran *resize* ini berkaitan dengan bagaimana ciri yang diambil dapat mencakup keseluruhan informasi yang terdapat dalam setiap citra. Ukuran *resize* 154x154 paling optimal dibandingkan dengan ukuran *resize* lainnya. akurasi terbaik didapat pada ukuran citra yang paling kecil pikselnya yaitu mendapat akurasi sebesar 61.7%. Hal ini dikarenakan dengan merubah ukuran piksel setiap citra dapat merubah nilai-nilai parameter yang digunakan untuk pembeda setiap kelasnya. Semakin kecil ukuran piksel citra semakin besar perbedaan nilai parameter setiap kelasnya. Maka dari itu untuk penelitian selanjutnya akan digunakan *resize* 154x154.

4.1.4 Analisis Hasil Pengujian Skenario 4

Pengujian dengan metode klasifikasi K-NN ini meliputi pengujian berdasarkan parameter-parameter terbaik yang didapatkan dari skenario-skenario sebelumnya. Untuk pengujian ini, pada proses *pre-processing* digunakan *resize* 154x154 karena pada skenario kedua telah dijelaskan bahwa akurasi paling baik didapat dengan menggunakan ukuran citra dengan piksel terkecil. Pada proses ekstraksi cirinya menggunakan sudut orientasi 135° *gray level cooccurrence matrix* karena pada skenario pertama, dibandingkan sudut yang lain sudut tersebut memperoleh akurasi yang paling tinggi dan menggunakan jarak antar pikselnya yaitu 2 yang didapat pada skenario kedua. Pada pengujian skenario ketiga ini dilakukan perubahan pada jenis *distance* K-NN yaitu *Euclidean*, *cityblock*, *cosine*, dan *minkowski*. Hasil pengujian yang telah dilakukan dapat dilihat pada tabel 4.4

Tabel 4.4 Tabel Data Benar Setiap Jenis Jarak

	NORMAL	LYMPHOMA	CARCINOMA
Euclidean	9	19	9
Cityblock	7	19	7
Cosine	8	18	10
Minkowski	9	19	9



Gambar 4.4 Gambar Persentase Akurasi Setiap Jenis Distance K-NN

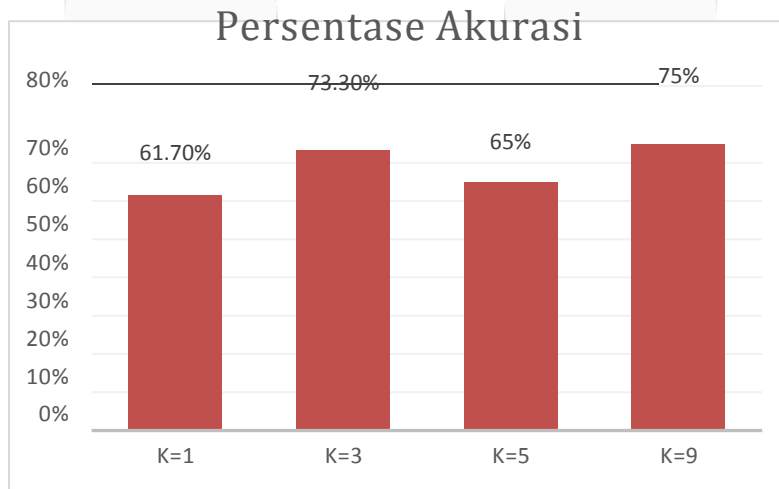
Grafik 4.4 menunjukkan bahwa tidak besar dampak akurasi yang didapat dengan melakukan perbedaan pada jenis *distance* pada K-NN. Pada grafik ditunjukkan bahwa akurasi terbaik pada skenario ketiga adalah sebesar 61.7%, yang pada hasil percobaan ketiga akurasi terbesar juga sebesar 61.7%. yang kesimpulannya bahwa jenis *distance* tidak banyak mempengaruhi optimasi akurasi sistem.

4.1.5 Analisis Hasil Pengujian Skenario 5

Pada pengujian kedua pada proses klasifikasi ini, nilai variable K akan diubah-ubah untuk melihat variable K mana yang menghasilkan akurasi terbaik dengan jenis *distance Euclidean* jenis yang menghasilkan nilai akurasi paling tinggi pada pengujian keempat, dengan sudut 135° dan jarak antar piksel yaitu 2 pada ekstraksi cirinya, dan resize 154x154 pada *pre-processing*nya. Hasil pengujian yang telah dilakukan pada skenario kelima dapat dilihat pada tabel 4.5.

Tabel 4.5 Tabel Data Benar Setiap Variabel K

	NORMAL	LYMPHOMA	CARCINOMA
K=1	9	19	9
K=3	15	19	10
K=5	14	17	8
K=9	16	18	11



Gambar 4.5 Gambar Persentase Akurasi Setiap Variabel K

Pada gambar 4.5 diatas menunjukkan hasil simulasi sistem dalam (%) yang dilakukan dengan perubahan nilai variable K pada K-NN. Terlihat bahwa penggunaan variable K sangat mempengaruhi hasil akurasi. Akurasi terbesar didapat ketika nilai K=9, yaitu sebesar 75%. Nilai variable K tersebut menghasilkan akurasi tertinggi karena ketika nilai variable K=9 maka nilai parameter setiap kelas semakin banyak yang dapat dibandingkan. Jika nilai variable K dibawah 9, maka akurasi cenderung semakin menurun hal ini disebabkan oleh nilai parameter yang hampir sama pada setiap kelasnya. Maka dari itu bila nilai variable K nya lebih dari satu akan membuat proses pengklasifikasiannya semakin jelas.

Persentase akurasi tertinggi didapatkan ketika $K=9$, hal ini disebabkan karena sistem mampu mendeteksi dengan benar lebih dari setengah citra uji pada setiap kelasnya. Pada kelas normal sistem mampu mendeteksi 16 dari 20 citra dengan benar, pada kelas lymphoma sistem mampu mendeteksi 18 dari 20 citra dengan benar yang berarti mendekati sempurna, dan pada kelas carcinoma sistem mampu mendeteksi 11 dari 20 citra dengan benar. Bila dibandingkan dengan variable K yang lain, yang salah satu kelasnya tidak dapat mendeteksi citra lebih dari setengah citra uji.

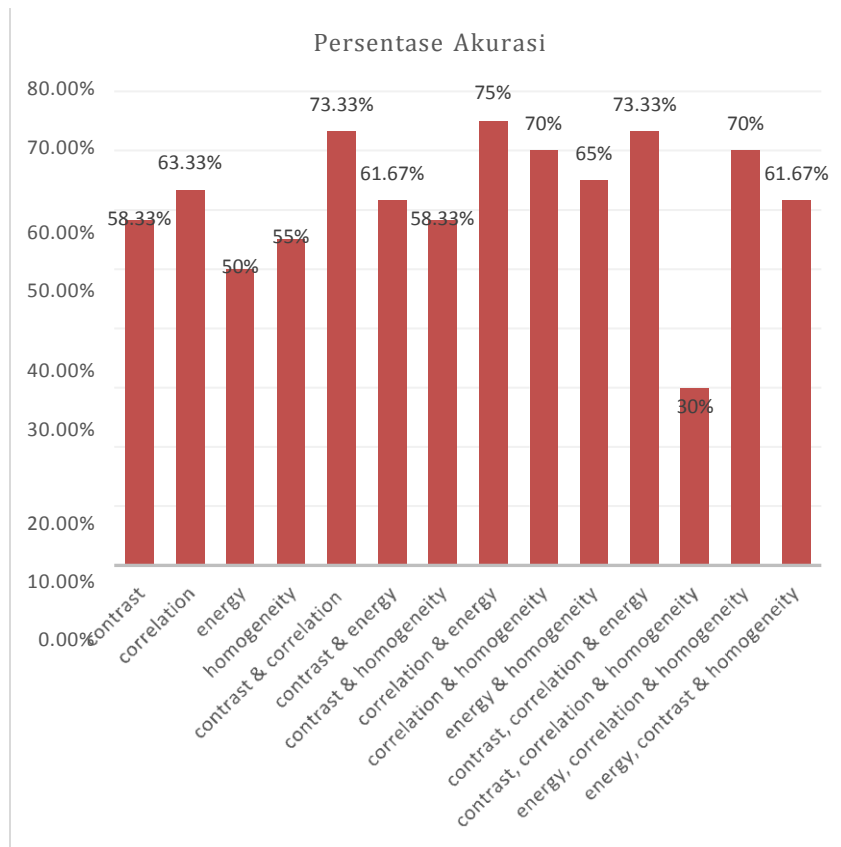
4.1.6 Analisis Hasil Pengujian Skenario 6

Hasil pengujian kelima dilakukan dengan membandingkan parameter yang digunakan, dengan menggunakan 1, kombinasi 2 dan kombinasi 3 dari 4 parameter yang digunakan, ditunjukkan oleh tabel dibawah ini dan huruf N pada tabel adalah NORMAL, huruf C adalah CARCINOMA dan L adalah LYMPHOMA :

Tabel 4.6 Data Benar Pada Skenario 6

	N	L	C
Menggunakan 1 parameter			
<i>Contrast</i>	10	9	16
<i>Correlation</i>	15	6	17
<i>Energy</i>	10	8	12
<i>Homogeneity</i>	10	7	16
Menggunakan kombinasi 2 parameter			
<i>Contrast & correlation</i>	15	18	11
<i>Contrast & energy</i>	9	17	11
<i>Contrast & homogeneity</i>	10	16	9
<i>Correlation & energy</i>	15	19	11
<i>Correlation & homogeneity</i>	14	19	9
<i>Energy & homogeneity</i>	12	17	10
Menggunakan kombinasi 3 parameter			
<i>Contrast, correlation & energy</i>	15	18	11
<i>Contrast, correlation & homogeneity</i>	1	0	19
<i>Energy, correlation & homogeneity</i>	14	19	9
<i>Energy, contrast & homogeneity</i>	10	17	10

Pada tabel 4.6 menunjukkan bahwa hasil data yang benar tidak menentu. Karena semua data dari ketiga kelas yang ada mempunyai kemiripan sehingga setiap parameter memiliki nilai yang hampir sama dengan setiap kelasnya. Dan pada tabel diatas dapat diambil kesimpulan bahwa satu parameter dengan akurasi terbaik mempengaruhi kenaikan hasil akurasi bila dikombinasikan dengan parameter lainnya.

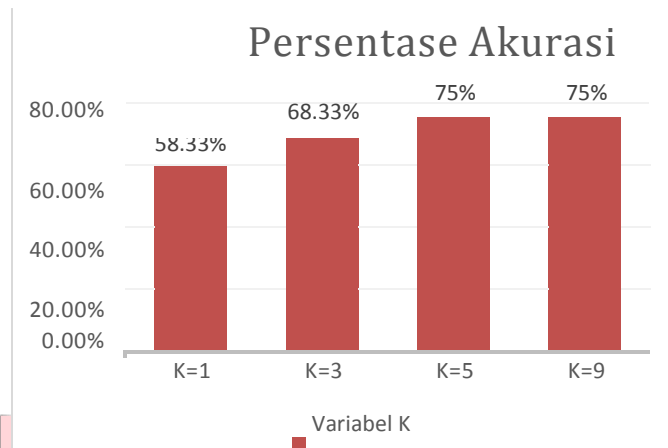


Gambar 4.6 Gambar Persentase Hasil Skenario 6

Dapat dilihat pada gambar 4.6 diperoleh persentase paling tinggi dari yang menggunakan kombinasi 2 parameter correlation dan energy, yaitu sebesar 75%. Yang menunjukkan bahwa parameter *correlation* dapat mempengaruhi naiknya nilai akurasi dan parameter yang mengandung ciri citra paling identik. Namun nilai akurasi yang didapat tidak lebih tinggi dari skenario sebelumnya dengan 4 parameter. Maka dari itu dilakukan pengujian tambahan yaitu dengan mengubah variable K 1,3,5 dan 9 dengan tetap menggunakan kombinasi 2 parameter tersebut. Hasilnya juga tidak lebih tinggi dengan pengujian pada skenario keempat melainkan sama, yaitu 75%. Didapat hasil sebagai berikut :

Tabel 4.7 Data Benar Setiap Variabel K dengan Kombinasi 2 Parameter

	NORMAL	LYMPHOMA	CARCINOMA
K=1	10	17	8
K=3	15	17	9
K=5	14	19	12
K=9	15	19	11



Gambar 4.7 Persentase Akurasi Variabel K dengan Kombinasi 2 Parameter

Pada pembuktian tersebut dapat disimpulkan bahwa dengan mengubah pemakaian parameter tidak membuat nilai akurasi semakin tinggi. Namun dapat mempersingkat pengujian dengan hanya menggunakan 2 kombinasi parameter saja. Dan setelah diuji juga dengan mengganti variable Knya 1, 3, 5, dan 9 pada variable K = 5 dan variable K = 9 mendapatkan nilai akurasi yang sama tingginya dengan yang menggunakan 4 parameter dengan variable K = 9 yaitu sebesar 75%.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisis yang telah dilakukan pada proses deteksi kanker *kolorectal* atau yang biasa disebut dengan kanker usus besar, maka dapat diambil kesimpulan sebagai berikut :

1. Sistem ini sudah mampu mendeteksi kanker usus besar dengan pembagian tiga kelas yaitu normal, kanker carcinoma, dan kanker lymphoma dengan akurasi maksimal sebesar 75%.
2. Perubahan sudut orientasi dan jarak piksel pada ekstraksi ciri *gray level cooccurrence matrix* akan mempengaruhi sistem pendeteksian. Hasil akurasi tertinggi didapat dengan menggunakan sudut 135° dengan hasil akurasi 53.33% dan $D = 2$ yang digabungkan dengan sudut orientasi 135° dengan hasil nilai akurasi 56.67%
3. Akurasi sistem juga dipengaruhi oleh besar atau kecilnya ukuran piksel pada citra. Semakin kecil ukuran piksel pada citra semakin besar akurasi yang didapatkan.
4. Dengan mengubah jenis *distance* tidak banyak mempengaruhi hasil akurasi pada sistem.
5. Tingkat akurasi tertinggi yang didapat dari hasil pengujian dengan menggunakan semua kombinasi output ciri *gray level cooccurrence matrix* menggunakan sudut orientasi 135° dan jarak piksel 2 berupa *contrast*, *correlation*, *energy*, dan *homogeneity*, *resize* pada *pre-processing* yaitu 154×154 , dan pada K-NN digunakan *distance* jenis *Euclidean* dengan nilai variable $K=9$.
6. Dengan melakukan pengujian terhadap kombinasi 1,2 dan 3 dari 4 parameter dan melakukan pengujian lanjutan dengan mengubah variabel K 1,3,5 dan 9 menggunakan kombinasi parameter yang mendapatkan nilai akurasi paling tinggi, yaitu sebesar 75% tidak dapat mengoptimalkan lagi nilai akurasi yang sudah didapat dari pengujian sebelumnya namun mendapat nilai akurasi yang sama 75%.

5.2 Saran

Adapun saran untuk pengembangan tugas akhir selanjutnya adalah:

1. Sistem dapat dikembangkan menggunakan metode ekstraksi ciri dan klasifikasi lain yang lebih baik serta bandingkan performansinya.
2. Uji parameter lain pada metode ekstraksi ciri dan klasifikasi
3. Sistem dibuat secara realtime
4. Memperbanyak parameter ciri yang diambil dari data

Referensi

- [1] Nugrahanto."Deteksi Gejala Hematuria Berdasarkan Jumlah Eritrosit pada Urine Berbasis Pengolahan Citra." Tugas Akhir Program Sarjana Institut Teknologi Telkom. Bandung,2012.
- [2] Zardi, Adilla. "Klasifikasi Kanker Usus Besar Berdasarkan Analisis Tekstur Dengan Deteksi Binary Large Object (BTOB)." Tugas Akhir Program Sarjana Universitas Telkom. Bandung, 2015

- [3] Z. A. Trimiarsih. "Klasifikasi Kanker Usus berdasarkan Citra Mikroskopik Patologi Menggunakan Contourlet Transform dan Support Vector Machine (SVM)." Tugas Akhir Program Sarjana Institut Teknologi Telkom. Bandung, 2013
- [5] Sianipar, R.H. 2013. *Pemograman MATLAB dalam Contoh dan Penerapan*. Bandung, Informatika.
- [6] Daya, Yudhi."Klasifikasi Kanker Usus Besar Menggunakan Ekstraksi Ciri Grey Level Co-occurrence Matrix dengan Metode Levenberg-Marquardt Algorithm." Tugas Akhir Sarjana Universitas Telkom. Bandung, 2013
- [7] <http://www.mathworks.com/help/images/ref/graycoprops.html>

