

# SISTEM DETEKSI COVID-19 BERDASARKAN CITRA X-RAY MENGGUNAKAN METODE GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM) DAN K-NEAREST NEIGHBOUR (K-NN)

## DETECTION SYSTEM OF COVID-19 BASED ON X-RAY IMAGES USING METHOD GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM) AND K-NEAREST NEIGHBOUR (K-NN)

Dimas Setriyo Wahyudi<sup>1</sup>, Raditiana Patmasari, S.T., M.T.<sup>2</sup>, R Yunendah Nur Fu'adah, S.T., M.T.<sup>3</sup>

<sup>1,2,3</sup> Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

Jln. Telekomunikasi No. 1 Terusan Buah Batu Bandung 40257 Indonesia

[dimassetriyo@telkomuniversity.ac.id](mailto:dimassetriyo@telkomuniversity.ac.id)<sup>1</sup>, [raditiana@telkomuniversity.ac.id](mailto:raditiana@telkomuniversity.ac.id)<sup>2</sup>, [yunendah@telkomuniversity.ac.id](mailto:yunendah@telkomuniversity.ac.id)<sup>3</sup>

### Abstrak

*Coronavirus Disease (COVID-19)* yang disebabkan oleh virus *Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2)* adalah *coronavirus* jenis baru. COVID-19 memiliki tingkat penularan yang tinggi dan mampu menjadi pandemik dalam kurun waktu yang singkat. Proses penanganan yang lamban akan mengakibatkan komplikasi hingga kematian. Diagnosis dari COVID-19 dapat dilakukan melalui *screening test*. Pada umumnya deteksi COVID-19 menggunakan *real-time reverse transcription-polymerase chain reaction (rRT-PCR)*. Pada Tugas Akhir ini digunakan data latih dan data uji berupa citra digital *Chest X-Ray (CXR)* untuk kondisi normal, *pneumonia* dan positif COVID-19, *Gray-Level Co-occurrence Matrix (GLCM)*, serta *K-Nearest Neighbor (K-NN)*. Dari hasil pengujian, sistem mampu melakukan deteksi covid-19 melalui citra *X-Ray* dengan tingkat nilai akurasi mencapai 85,6%. Akurasi tersebut diupayakan melalui pengujian 180 citra sebagai pengoptimalan sistem yang menggunakan metode GLCM dan KNN untuk pengolahan citra di dalamnya. Dengan pengujian terhadap parameter pada GLCM seperti jarak ( $d$ ), sudut ( $\theta$ ) dan ciri statistik orde dua (*energy, contrast, correlation, variance, homogeneity dan entropy*) serta parameter pada nilai  $k$  (1,3,5 dan 7) dan persamaan jarak (*Euclidean, City-Block, Chebychef, Minkowski*) pada tahap klasifikasi KNN menjadi dasar pengujian dari performa sistem.

Kata kunci : COVID-19, *Gray-Level Co-Occurrence Matrix*, *K-Nearest Neighbour*.

### Abstract

*Coronavirus Disease (COVID-19)* yang disebabkan oleh virus *Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2)* adalah *coronavirus* jenis baru. COVID-19 memiliki tingkat penularan yang tinggi dan mampu menjadi pandemik dalam kurun waktu yang singkat. Proses penanganan yang lamban akan mengakibatkan komplikasi hingga kematian. Diagnosis dari COVID-19 dapat dilakukan melalui *screening test*. Pada umumnya deteksi COVID-19 menggunakan *real-time reverse transcription-polymerase chain reaction (rRT-PCR)*. Pada Tugas Akhir ini digunakan data latih dan data uji berupa citra digital *Chest X-Ray (CXR)* untuk kondisi normal, *pneumonia* dan positif COVID-19, *Gray-Level Co-occurrence Matrix (GLCM)*, serta *K-Nearest Neighbor (K-NN)*. Dari hasil pengujian, sistem mampu melakukan deteksi covid-19 melalui citra *X-Ray* dengan tingkat nilai akurasi mencapai 85,6%. Akurasi tersebut diupayakan melalui pengujian 180 citra sebagai pengoptimalan sistem yang menggunakan metode GLCM dan KNN untuk pengolahan citra di dalamnya. Dengan pengujian terhadap parameter pada GLCM seperti jarak ( $d$ ), sudut ( $\theta$ ) dan ciri statistik orde dua (*energy, contrast, correlation, variance, homogeneity dan entropy*) serta parameter pada nilai  $k$  (1,3,5 dan 7) dan persamaan jarak (*Euclidean, City-Block, Chebychef, Minkowski*) pada tahap klasifikasi KNN menjadi dasar pengujian dari performa sistem.

Keywords: COVID-19, *Gray-Level Co-Occurrence Matrix*, *K-Nearest Neighbour*.

### 1. Pendahuluan

Pada Desember 2019, terjadi kasus *pneumonia* misterius di Wuhan, Provinsi Hubei, China. Dari sampel yang diteliti menunjukkan etiologi *coronavirus* baru, virus ini dapat ditularkan dari manusia ke manusia dan telah menyebar secara luas dalam kurun waktu yang singkat. Sumber penularan pada kasus pertama dikaitkan dengan pasar hewan di Wuhan. Pasien yang dirawat memiliki *Acute Respiratory Distress Syndrome (ARDS)*. Untuk sementara waktu penyakit ini dinamakan sebagai 2019 novel *coronavirus (2019-nCov)*, kemudian World Health

Organization (WHO) mengumumkan nama baru pada 11 Februari 2020 yaitu *Coronavirus Disease (COVID-19)* yang disebabkan oleh virus *Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2)* [1].

Pada Tugas Akhir ini digunakan data latih dan data uji berupa citra digital CXR untuk kondisi normal dan positif COVID-19, kemudian GLCM, serta K-NN. Metode GLCM adalah metode ekstraksi fitur untuk melihat tekstur berbasis matriks dan statistikal yang digunakan untuk mendapat fitur dari setiap piksel citra dimana fitur-fitur tersebut akan digunakan pada tahap klasifikasi. Sedangkan K-NN adalah algoritma

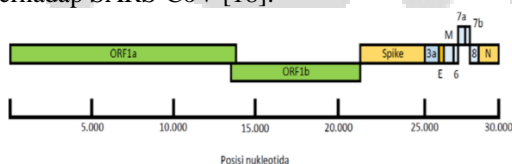
analisis klasifikasi yang digunakan untuk mengklasifikasikan data uji berada pada kondisi normal dan COVID-19. Pemilihan GLCM berkaitan dengan analisis tekstur yang mempermudah dalam analisis gambar medis, dan K-NN memiliki kelebihan ciri dalam pendekatan yang sederhana namun mampu bekerja dengan baik ketika digabungkan dengan metode pengumpulan data klasik lainnya, terutama dalam skala besar.

**2. Dasar Teori**

**2.1 Coronavirus Disease (COVID-19)**

*Coronavirus* adalah virus RNA, yaitu jenis virus yang mengandung *ribonucleic acid* (RNA) yang merupakan material genetik. Penyakit yang disebabkan oleh virus RNA yang telah teridentifikasi selama ini antara lain adalah pilek, influenza, SARS, COVID-19, Hepatitis C, Hepatitis E, *West Nile fever*, Ebola, Rabies, Polio dan Campak. Virus RNA COVID-19 memiliki ukuran partikel 120-160 nm. Virus ini utamanya menginfeksi hewan, termasuk diantaranya adalah kelelawar dan unta. Sebelum terjadinya wabah COVID-19, ada 6 jenis *coronavirus* yang dapat menginfeksi manusia, yaitu *alphacoronavirus* 229E, *alphacoronavirus* NL63, *betacoronavirus* OC43, *betacoronavirus* HKU1, *Severe Acute Respiratory Illness Coronavirus* (SARS-CoV), dan *Middle East Respiratory Syndrome Coronavirus* (MERS-CoV) [7].

*Coronavirus* yang menjadi etiologi COVID-19 termasuk dalam genus *betacoronavirus*. Struktur genom virus ini memiliki pola seperti *coronavirus* pada umumnya Gambar 1. Sekuens SARS-CoV-2 memiliki kemiripan dengan *coronavirus* yang diisolasi pada kelelawar, sehingga muncul hipotesis bahwa SARS-CoV-2 berasal dari kelelawar yang kemudian bermutasi dan menginfeksi manusia [15]. Pada kasus COVID-19, trenggiling diduga sebagai reservoir perantara. *Strain coronavirus* pada trenggiling adalah yang mirip genomnya dengan *coronavirus* kelelawar (90,5%) dan SARS-CoV-2 (91%) [17]. Genom SARS-CoV-2 sendiri memiliki homologi 89% terhadap *coronavirus* kelelawar ZXC21 dan 82% terhadap SARS-CoV [18].



Gambar 1 Struktur genom *coronavirus* [11].

**2.2 Gray-Level Co-occurrence Matrix (GLCM)**

Metoda GLCM bekerja dengan memanfaatkan kesamaan nilai (kookurensi) dari nilai piksel yang diobservasi dengan tetangganya dalam jarak dan

orientasi sudut pengamatan tertentu, dilanjutkan dengan menentukan fitur dari matriks citra tersebut. Jarak dinyatakan dalam piksel dan orientasi dinyatakan dalam derajat. Orientasi sudut ditentukan dengan interval sudut 45°, yaitu 0°, 45°, 90°, dan 135°. Sedangkan jarak antar piksel biasanya ditetapkan minimal sebesar 1 piksel.

Langkah-langkah untuk mencari fitur GLCM secara berurutan adalah sebagai berikut [20,21,22,23]:

1. Mencari Matriks Kookurensi.

a. Matriks asal.

Setiap piksel dari sebuah citra selalu memiliki nilai, nilai tersebut lalu direpresentasikan dalam bentuk sebuah matriks. Gambar 2 sebagai ilustrasi matriks asal sebagai berikut.

I =

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

Gambar 0 Matriks asal (matriks I).

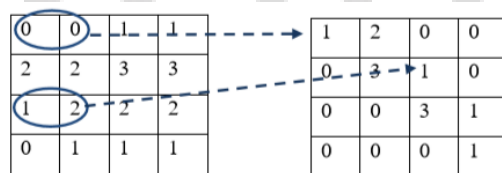
b. Membuat area kerja matriks.

Berdasarkan Gambar 2 matriks I mempunyai 4 derajat keabuan, dengan penempatan nilai piksel referensi random. Sehingga pada area kerja matriks berjumlah 4 seperti pada Gambar 3.

Nilai piksel tetangga	0	1	2	3
Nilai piksel referensi				
0	0,0	0,1	0,2	0,3
1	1,0	1,1	1,2	1,3
2	2,0	2,1	2,2	2,3
3	3,0	3,1	3,2	3,3

Gambar 3 Area kerja dari matriks I.

c. Menentukan hubungan antara piksel yang diobservasi dengan piksel tetangga. Dengan menentukan nilai sudut ( $\theta$ ) dan jarak ( $d$ ). Hubungan spasial untuk  $d=1$  dan  $\theta=0^\circ$  pada matriks I dapat dilihat pada Gambar 4.



(a)

(b)

**Gambar 4** Hubungan spasial antar piksel dan matriks kookurensi.

Menghitung jumlah kookurensi dan mengisikannya pada area kerja matriks. Angka 1 pada (0, 0) pada Gambar 4 (b) menunjukkan jumlah hubungan pasangan (0, 0) pada matriks asal berjumlah 1. Begitupula angka 1 pada (1, 2) pada Gambar 4 (b) menunjukkan jumlah hubungan pasangan (1, 2) pada matriks asal berjumlah 1. Setelah semua jumlah kookurensi dihitung maka diperoleh matriks kookurensi seperti pada Gambar 4 (b).

d. Menjumlahkan matriks kookurensi dengan transposenya untuk menjadikannya simetris. Matriks kookurensi yang diperoleh kemudian ditambahkan dengan matriks transposenya untuk menjadikannya simetris terhadap sumbu diagonal. Berikut ini adalah (i, j) dari matriks asal ditambahkan dengan transposenya, dan hasilnya simetris, seperti pada Gambar 5.

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 0 & 0 \\ 2 & 6 & 1 & 0 \\ 0 & 1 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

I + I<sup>t</sup> = I simetris

**Gambar 5** Matriks kookurensi simetris.

2. Normalisasi matriks untuk mengubahnya ke bentuk probabilitas.

Matriks yang telah simetris selanjutnya dinormalisasi, setiap nilai diubah kedalam bentuk probabilitas. Nilai untuk masing-masing sel pada matriks kookurensi simetris dibagi dengan jumlah seluruh nilai keseluruhan matriks. Matriks yang telah dinormalisasi diperlihatkan pada Gambar 6. Nilai 0, 0833 pada (0, 0) diperoleh dari 2 dibagi jumlah seluruh nilai piksel yaitu 24.

0,0833	0,0833	0	0
0,0833	0,2500	0,0420	0
0	0,0420	0,2500	0,0042
0	0	0,0042	0,0833

**Gambar 6** Matriks kookurensi simetris ternormalisasi.

3. Mencari nilai fitur masing-masing orientasi. Berdasarkan dari setiap matriks kookurensi yang telah ternormalisasi, Haralick mengusulkan 14 fitur statistik yang dikenal dengan Fitur Haralick. Fitur-fitur tersebut merupakan hasil pengukuran statistik yang bisa di dapat dari satu matriks kookurensi dari sebuah citra.

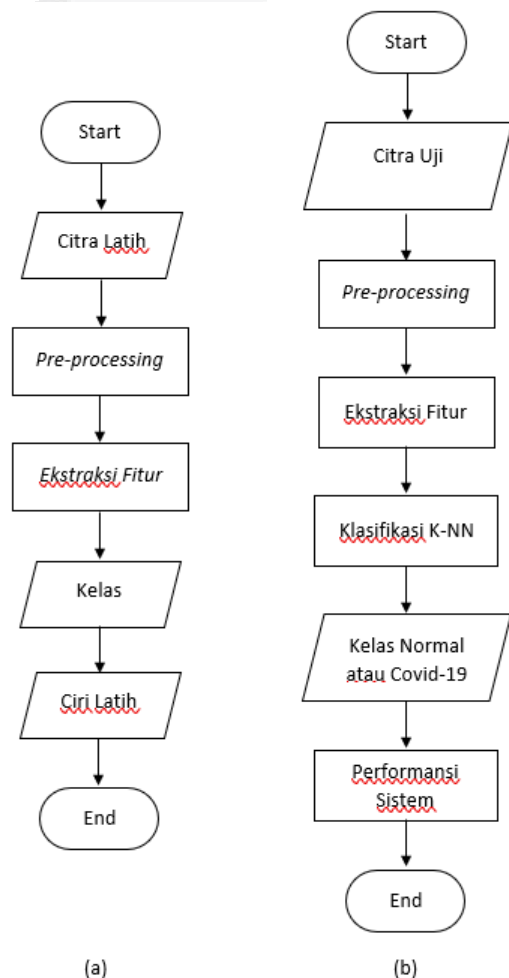
**2.5 K-Nearest Neighbor (KNN)**

*K-Nearest Neighbour* (K-NN) adalah sebuah algoritma analisis klasifikasi dengan metode prediksi, yang secara umum menentukan dugaan kelas atau kelompok dari suatu objek baru

bedasarkan karakteristik yang dimilikinya. Suatu objek tergolongkan dalam kelas yang sama dengan kelas dari mayoritas tetangganya. Prinsip kerja lanjut dari K-NN adalah mencari jarak terdekat antara data uji dengan sejumlah *K* tetangga terdekatnya. Nilai jarak merupakan pendekatan yang nantinya digunakan untuk menentukan kesamaan dua vektor fitur, yaitu fitur latih dan fitur uji. Pengukuran jarak pada metode K-NN dilakukan dengan beberapa metode, diantaranya:

1. Jarak Euclidean 
$$j(a, b) = \sqrt{\sum_{k=1}^d (a_k - b_k)^2}$$
 (1)
2. Jarak City-Block 
$$j(a, b) = \sum_{k=1}^d |a_k - b_k|$$
 (2)
3. Jarak Chebychef 
$$j(a, b) = \max_{k=1 \rightarrow N} (|a_k - b_k|)$$
 (3)
4. Jarak Minkowski 
$$j(a, b) = \sqrt[p]{\sum_{k=1}^d |a_k - b_k|^p}$$
 (4)

**3. Perancangan Sistem**



**Gambar 7** Diagram alir sistem

### 3.1 Input Citra

Citra paru-paru yang digunakan sebagai input sistem dalam tugas akhir ini berupa data sekunder yang didapat melalui situs *www.kaggle.com* yang dipilih berdasarkan kelas normal, pneumonia dan COVID-19. Citra tersimpan dengan format JPG file. Jumlah keseluruhan sampel adalah 180 buah, terbagi menjadi 2 untuk citra latih dan citra uji.

### 3.2 Pemrosesan Awal

Pemrosesan awal atau *Preprocessing* bertujuan untuk mengolah citra agar mendapatkan ciri atau informasi yang dibutuhkan. Pemrosesan awal meliputi:

#### 1. *Cropping* Citra

Proses pemotongan *Region of Interest* (RoI) dari citra input, dalam hal ini bagian sekitaran dada/paru-paru.

#### 2. Konversi RGB ke Grayscale

Citra paru-paru yang akan diproses adalah citra *graylevel*. Oleh karena itu, input citra yang masih dalam bentuk citra RGB ataupun lainnya perlu diubah tipenya menjadi citra *grayscale*.

#### 3. *Resize* Citra

Proses perubahan ukuran citra menjadi ukuran tertentu sebelum proses ekstrasi ciri. Ketiga proses ini dilakukan pada data input citra latih maupun citra uji.

### 3.3 Ekstrasi Fitur

Ekstraksi fitur merupakan tahapan pengumpulan informasi fitur dari citra. Pada Tugas Akhir ini metode ekstraksi fitur yang digunakan adalah GLCM. Adapun tahapan dalam ekstrasi ciri sebagai berikut:

#### 1. Pembuatan Matriks Ko-okurensi

Matriks ko-okurensi merupakan matriks yang berisi informasi frekuensi kemunculan pasangan piksel dalam sebuah citra. Pembuatan matriks sebanyak 12 orientasi matriks.

#### 2. Normalisasi Matriks Ko-okurensi

Normalisasi terhadap matriks ko-okurensi berfungsi untuk menyederhanakan nilai dari matriks ko-okurensi.

#### 3. Perhitungan Fitur Statistik

Setelah matriks ko-okurensi ternormalisasi diperoleh tahap selanjutnya adalah mengekstraksi fitur statistik dari matriks ko-okurensi ternormalisasi. Fitur statistik yang diperoleh adalah sebanyak 6 fitur dari 14 fitur Haralick.

### 3.4 Klasifikasi

Tahap klasifikasi adalah tahapan untuk mengklasifikasikan citra uji ke dalam kondisi normal atau COVID-19. Proses pengklasifikasian diawali dengan ekstraksi fitur uji kemudian fitur tersebut dibandingkan dengan fitur latih yang ada pada database fitur. Pada Tugas Akhir ini, digunakan metode K-NN sebagai metode klasifikasi dengan nilai  $K=1,3,5$ , dan 7 serta perhitungan jarak menggunakan persamaan *Euclidean*, *Minkowski*, *Chebychev*, dan *Cityblock*.

### 3.5 Akurasi Sistem

Akurasi sistem adalah ukuran ketepatan sistem dalam mengenali input yang diberikan sehingga menghasilkan output yang benar. Akurasi sistem merupakan perbandingan antara data uji yang dikenali dan benar (*True Positive*) terhadap jumlah data uji. Secara matematis akurasi sistem ditunjukkan pada Persamaan 5.

$$\text{Akurasi} = \frac{\sum TP}{\sum \text{Data Uji}} \times 100\% \quad (5)$$

Dengan,

TP

:

*True Positive*

#### 4. Analisis

##### 4.1 Pengujian Parameter Jarak dan Arah pada GLCM

Berikut adalah hasil pengujian parameter jarak dan arah pada GLCM. Parameter yang digunakan adalah  $d = 1, 2, 3, 4, 5$  dan 6;  $\theta = 0^\circ, 45^\circ, 90^\circ$  dan  $135^\circ$ ; parameter ciri statistik orde dua (*Energy*, *Contrast*, *Correlation*, *Variance*, *Homogeneity*, *Entropy*) dan parameter KNN  $k = 1$ ; *distance Euclidean*.

**Tabel 1** Hasil pengujian arah dan jarak pada GLCM.

d	$\theta$	Akurasi (%)	Jumlah Data Benar
1	0	75,6	68
	45	75,6	66
	90	62,2	56
	135	72,2	65
2	0	72,2	72
	45	73,3	66
	90	63,3	57
	135	73,3	66
3	0	81,1	73
	45	70	63
	90	63,3	57
	135	73,3	66
4	0	84,4	76
	45	67,8	61
	90	62,2	66
	135	65,6	59
5	0	80	72
	45	73,3	66
	90	64,4	58
	135	70	63
6	0	75,6	68
	45	73,3	66
	90	62,2	56
	135	71,1	64

Berdasarkan Tabel 1 akurasi terbesar didapatkan pada saat parameter jarak sebesar 1 *pixel* dengan arah  $0^0$ , yaitu 84.4%.

#### 4.2 Pengujian Ekstraksi Ciri Orde Dua GLCM

Pengujian pada tahap ini menggunakan parameter orde dua GLCM (*Energy, Contrast, Correlation, Variance, Homogeneity, Entropy*) dengan  $d = 4$ , level kuantisasi 8 dan parameter KNN  $k=1$ ; distance Euclidean.

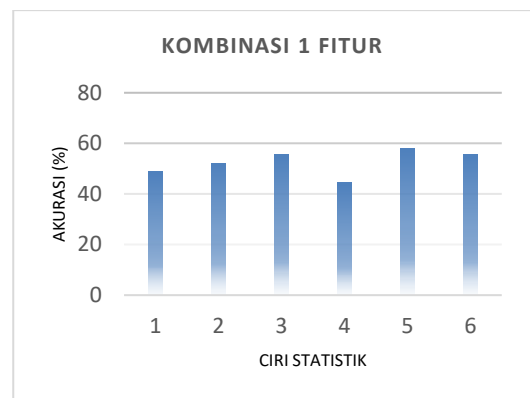
Pengujian sistem dilakukan dengan menganalisis pengaruh jumlah fitur terhadap parameter akurasi sistem. Pengujian sistem pertama-tama dilakukan terhadap 1 fitur tekstur statistik, hingga 5 fitur yang memberikan perbedaan antar kelasnya.

**Tabel 2** Fitur yang digunakan pada pengujian terhadap 6 fitur statistic.

No.	Fitur
-----	-------

1.	Energy
2.	Contrast
3.	Correlation
4.	Variance
5.	Homogeneity
6.	Entropy

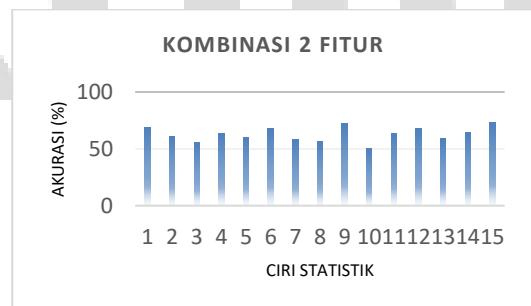
Pengujian selanjutnya dilakukan terhadap kombinasi 1 fitur tekstur statistik menggunakan *distance Euclidean*. Adapun hasil dari pengujian 1 fitur tekstur statistik menggunakan *distance Euclidean* dapat dilihat pada gambar 8.



**Gambar 8** Hasil pengujian kombinasi 1 fitur tekstur statistik menggunakan *distance Euclidean*.

Gambar 8 menunjukkan bahwa akurasi tertinggi yang dihasilkan pada pengujian kombinasi 1 fitur tekstur statistik menggunakan *distance Euclidean* adalah sebesar 57,8%, hasil tersebut diperoleh pada penggunaan  $K = 1$ , di mana kombinasi 1 fitur tersebut merupakan hasil ekstraksi dari matriks ko-okurensi dengan nilai  $d = 4$  dan  $\theta = 0^0$ .

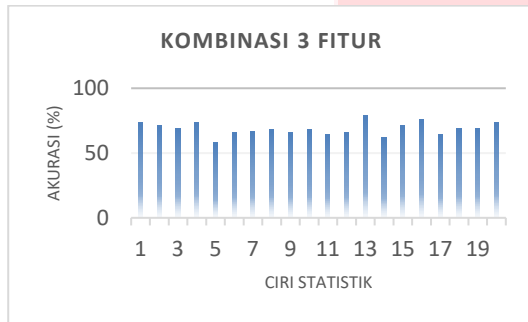
Pengujian berikutnya yang dilakukan terhadap kombinasi 2 fitur tekstur statistik menggunakan *distance Euclidean* memberikan hasil seperti yang ditunjukkan pada Gambar 9.



**Gambar 9** Hasil pengujian kombinasi 2 fitur tekstur statistik menggunakan *distance Euclidean*.

Gambar 9 menunjukkan bahwa pada pengujian terhadap kombinasi 2 fitur tekstur statistik menggunakan jarak Chebychev diperoleh akurasi

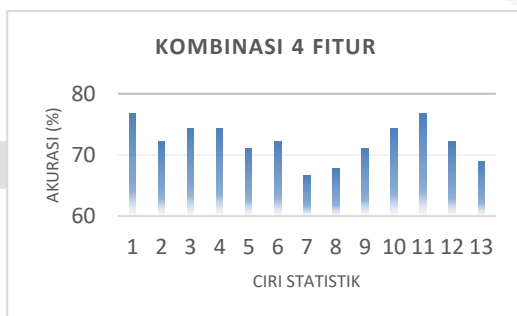
tertinggi yaitu 73,3%. Hasil tersebut diperoleh menggunakan  $K = 1$  di mana kombinasi 2 fitur tersebut merupakan hasil ekstraksi dari matriks ko-okurensi dengan nilai  $d = 4$  dan  $\theta = 0^\circ$ . Pengujian selanjutnya dilakukan terhadap kombinasi 3 fitur tekstur statistik menggunakan jarak Chebychev memberikan hasil seperti yang dapat dilihat pada Gambar 10.



**Gambar 10** Hasil pengujian kombinasi 3 fitur tekstur statistik menggunakan *distance Euclidean*.

Gambar 10 menunjukkan bahwa pada pengujian terhadap kombinasi 3 fitur tekstur statistik menggunakan *distance Euclidean* diperoleh akurasi tertinggi yaitu 78,9%. Hasil tersebut diperoleh menggunakan  $K = 1$  di mana kombinasi 3 fitur tersebut merupakan hasil ekstraksi dari matriks ko-okurensi dengan nilai  $d = 4$  dan  $\theta = 0^\circ$ .

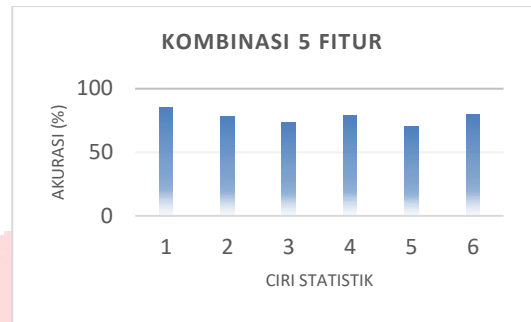
Pengujian selanjutnya dilakukan terhadap kombinasi 4 fitur tekstur statistik menggunakan *distance Euclidean* memberikan hasil seperti yang dapat dilihat pada Gambar 11.



**Gambar 11** Hasil pengujian kombinasi 4 fitur tekstur statistik menggunakan *distance Euclidean*.

Gambar 11 menunjukkan bahwa pada pengujian terhadap kombinasi 4 fitur tekstur statistik menggunakan *distance Euclidean* diperoleh akurasi tertinggi yaitu 76,7%. Hasil tersebut diperoleh menggunakan  $K = 1$  di mana kombinasi 4 fitur tersebut merupakan hasil ekstraksi dari matriks ko-okurensi dengan nilai  $d = 4$  dan  $\theta = 0^\circ$ .

Pengujian selanjutnya dilakukan terhadap kombinasi 5 fitur tekstur statistik menggunakan *distance Euclidean* memberikan hasil seperti yang dapat dilihat pada Gambar 12.



**Gambar 12** Hasil pengujian kombinasi 5 fitur tekstur statistik menggunakan *distance Euclidean*.

Berdasarkan Gambar 12, akurasi tertinggi yang dihasilkan pada pengujian terhadap 5 fitur menggunakan *distance Euclidean* adalah sebesar 85,6%, yaitu pada penggunaan  $K=5$  dengan fitur *Energy*, *Contrast*, *Correlation*, *Variance*, *Homogeneity*.

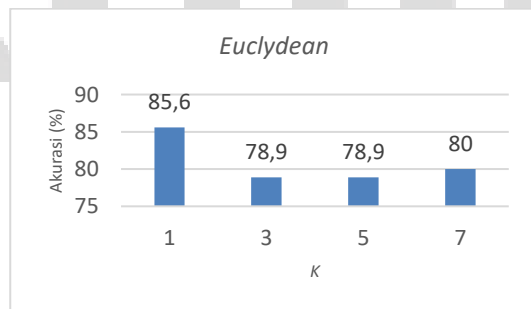
Berdasarkan seluruh pengujian kombinasi fitur yang dilakukan menggunakan *distance Euclidean*, diperoleh akurasi tertinggi sebesar 85,6%, dengan parameter  $d = 4$  dan  $\theta = 0^\circ$ . Penggunaan *distance Euclidean* memberikan hasil berupa nilai acuan awal terhadap kombinasi terbaik yang akan digunakan pada tahap pengujian berikutnya.

#### 4.3 Pengujian Parameter *Distance* dan Nilai $K$ pada KNN

Pada tahap ini akan diuji pengaruh Nilai  $K$  pada KNN. Parameter jarak yang digunakan adalah  $d$  (1, 2, 3, 4, 5 dan 6) arah  $0^\circ$  parameter ciri orde dua (*Energy*, *Contrast*, *Correlation*, *Variance*, *Homogeneity* dan *Entropy*) dan parameter KNN  $k$  (1,3,5,7), *distance* (*Euclidean*, *Minkowski*, *Chebychev*, *Cityblock*).

##### 4.3.1 Jarak *Euclidean*

Berikut ini adalah hasil pengujian 5 fitur statistik terhadap akurasi sistem menggunakan jarak *Euclidean* dengan  $K = 1, 3, 5$  dan  $7$ .



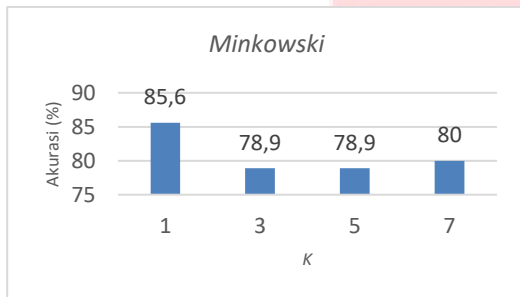
**Gambar 13** Hasil pengujian kombinasi 5 fitur tekstur statistik menggunakan jarak *Euclidean*.

Gambar 13 menunjukkan bahwa akurasi terbaik pada penggunaan 5 fitur statistik adalah 85,6% di mana hasil tersebut lebih baik dibandingkan

dengan penggunaan kombinasi fitur lainnya. Pada pengujian terhadap 5 fitur statistik penggunaan nilai  $K = 1$  memberikan hasil yang lebih baik dibandingkan dengan  $K = 3, 5$  dan  $7$ .

#### 4.3.2 Jarak Minkowski

Hasil pengujian 5 fitur tekstur statistik menggunakan nilai  $K = 1, 3, 5$  dan  $7$  menggunakan perhitungan jarak *Minkowski* dapat dilihat pada Gambar 14.

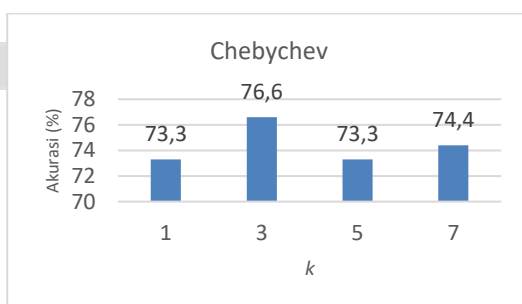


**Gambar 10** Hasil pengujian terhadap 5 fitur tekstur statistik menggunakan jarak Minkowski.

Berdasarkan Gambar 14 dapat dilihat bahwa akurasi tertinggi pada pengujian 5 fitur tekstur statistik menggunakan jarak Minkowski adalah 85,6%. Penggunaan nilai  $K = 1$  cenderung memberikan akurasi yang lebih tinggi dibandingkan dengan  $K = 3, 5$  dan  $7$ .

#### 4.3.3 Jarak Chebychev

Hasil pengujian 5 fitur tekstur statistik menggunakan nilai  $K = 1, 3, 5$  dan  $7$  menggunakan perhitungan jarak *Chebychev* dapat dilihat pada Gambar 15.



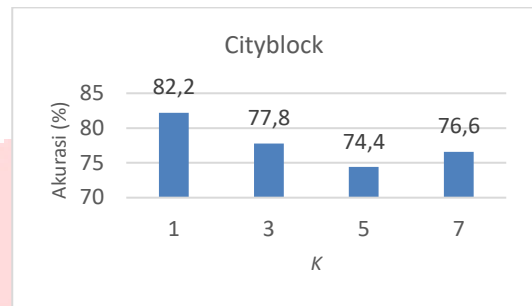
**Gambar 15** Hasil pengujian terhadap 5 fitur tekstur statistik menggunakan jarak Chebychev.

Berdasarkan Gambar 15 dapat dilihat bahwa akurasi tertinggi pada pengujian 5 fitur tekstur statistik menggunakan jarak *Chebychev* adalah 76,6%. Penggunaan nilai  $K = 3$  cenderung memberikan akurasi yang lebih tinggi dibandingkan dengan  $K = 1, 5$  dan  $7$ .

#### 4.3.4 Jarak Cityblock

Hasil pengujian 5 fitur tekstur statistik menggunakan nilai  $K = 1, 3, 5$  dan  $7$

menggunakan perhitungan jarak *Cityblock* dapat dilihat pada Gambar 16.



**Gambar 16** Hasil pengujian terhadap 5 fitur tekstur statistik menggunakan jarak Cityblock.

Berdasarkan Gambar 4.9 dapat dilihat bahwa akurasi tertinggi pada pengujian 5 fitur tekstur statistik menggunakan jarak *Cityblock* adalah 82,2%. Penggunaan nilai  $K = 1$  cenderung memberikan akurasi yang lebih tinggi dibandingkan dengan  $K = 3, 5$  dan  $7$ .

## 5 Saran dan Kesimpulan

### 5.1 Kesimpulan

Berdasarkan hasil pengujian dan perhitungan yang dilakukan pada penelitian tugas akhir ini, diajukan kesimpulan sebagai berikut :

1. Sistem ini mampu mengklasifikasikan kondisi pasien menggunakan metode GLCM dengan klasifikasi K-NN dengan akurasi terbesar adalah 85,6% dengan parameter  $d=4$ ; arah  $0^\circ$ ; fitur *Energy, Contrast, Correlation, Variance, Homogeneity* pada GLCM dan  $k=1$ ; *distance Euclidean* atau *Minkowski*.
2. Sistem tidak mampu menjadi pengganti *Screening test*, namun dapat sebagai diagnosis penguat/tambahan untuk memastikan kondisi pasien secara akurat.

### 5.2 Saran

Sistem deteksi COVID-19 ini masih dapat dikembangkan, sehingga tingkat akurasi yang diperoleh lebih besar dan akurat. Oleh karena itu Adapun saran untuk pengembangan tugas akhir ini selanjutnya yaitu perlu :

1. Database dari citra paru-paru perlu semakin diperbanyak, meliputi dari pasien *pneumonia*, covid-19, normal, dan sembuh dari covid-19.
2. Menggunakan metode yang berbeda untuk mendapat hasil lebih maksimal dalam mendeteksi / mengklasifikasikan hasil ekstraksi ciri citra, sehingga dapat dibandingkan metode mana yang lebih baik.
3. Penggunaan metode ekstraksi ciri citra yang lebih baik.

## REFERENCE

- [1] Rothan HA, Byrareddy SN, "The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak," *europemc.org, Journal of Autoimmunity*, Feb. 26, 2020.
- [2] World Health Organization, "Situation Report – 42," *World Health Organization*, updated March. 2, 2020. [Online]. Available: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.
- [3] Kementerian Kesehatan Republik Indonesia, "COVID-19," Kementerian Kesehatan Republik Indonesia, Info Infeksi Emerging, 2020. [Online]. Available: <https://infeksiemerging.kemkes.go.id/>.
- [4] Z.Y. Zu, et al., "Coronavirus disease 2019 (COVID-19): a perspective from China," *pubs.rsna.org, Radiology*, Feb. 21, 2020.
- [5] E.E.D. Hemdan, M.A. Shouman and M.E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," *arxiv.org, Cornell University*, March 24, 2020.
- [6] L. Wang, A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images," *arxiv.org, Cornell University*, March 22, 2020.
- [7] Riedel S, Morse S, Mietzner T, Miller S. Jawetz, Melnick, & Adelberg's Medical Microbiology. 28th ed. New York: McGrawHill Education/Medical, p.617-22, 2019.
- [8] Huang C, et al., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *europemc.org, Lancet (London, England)*, Jan. 24, 2020.
- [9] Guan WJ, et al., "Clinical Characteristics of Coronavirus Disease 2019 in China," *europemc.org, New Engl J Med*, Feb. 28, 2020.
- [10] Chen J, et al., "Clinical progression of patients with COVID-19 in Shanghai, China," *europemc.org, J Infect*, March 19, 2020.
- [11] Wang D, et al., "Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China," *europemc.org, JAMA*, Feb. 7, 2020.
- [12] World Health Organization, "Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases," *Institutional Repository for Information Sharing*, Geneva: World Health Organization, March 2, 2020.
- [13] Yan G, et al., "Covert COVID-19 and false-positive dengue serology in Singapore," *europemc.org, Lancet Infect Dis.*, March 4, 2020.
- [14] Arentz M, et al., "Characteristics and Outcomes of 21 Critically Ill Patients With COVID-19 in Washington State," *jamanetwork.vom, JAMA*, March 19, 2020.
- [15] Han Y and Yang H, "The transmission and diagnosis of 2019 novel coronavirus infection disease (COVID-19): A Chinese perspective," *europemc.org, J Med Virol*, March 12, 2020.
- [16] Zhou P, et al., "A pneumonia outbreak associated with a new coronavirus of probable bat origin," *europemc.org, Nature*, Feb. 3, 2020.
- [17] Zhang T, Wu Q and Zhang Z., "Probable Pangolin Origin of SARSCoV-2 Associated with the COVID-19 Outbreak," *europemc.org, Curr Biol*, March 19, 2020.
- [18] Chan JF-W, et al., "Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan," *europemc.org, Emerg Microbes Infect*, Jan. 28, 2020.
- [19] Marchiori, Dennis M. "Clinical Imaging: with Skeletal, Chest, and Abdominal Pattern Differentials," Elsevier Mosby, 2014.
- [20] R Y.N. Fu'adah, "Deteksi Katarak Pada Citra Mata Digital Menggunakan Metoda Analisis Tekstur Statistik dan *K-Nearest Neighbor*," M.T. thesis, Institut Teknologi Bandung, Indonesia, 2014.
- [21] N. Charimmah, E. Lanovia, K. Usman and L. Novamizanti, "Deteksi Kantuk Melalui Citra Wajah Menggunakan Metode Gray Level Co-occurrence Matrix (GLCM) dan Klasifikasi Support Vector Machine (SVM)," in Seminar Nasional Teknik Elektro UIN Sunan Gunung Djati Bandung (SENTER 2019) , 2020.
- [22] Liu M, Zeng W, Wen Y, Zheng Y, Lv F, Xiao K. COVID-19 pneumonia: CT findings of 122 patients and differentiation from influenza pneumonia. *Eur Radiol*. 2020 Oct; 30(10):5463-5469. doi: 10.1007/s00330-020-06928-0. Epub 2020 May 12. PMID: 32399710; PMCID: PMC7216854.
- [23] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M. T. Islam, "Can AI help in screening Viral and



COVID-19 pneumonia?" IEEE Access,  
Vol. 8, 2020, pp. 132665 - 132676.



Telkom  
University