

Klasifikasi Data Menggunakan Algoritma *K-Means Clustering* Dan *Naive Bayes Classifier* Berdasarkan Analisa Tekstur Metode *Gray Level Co-Occurrence Matrix* (GLCM)

Samsu Supriyatna¹

Universitas Pamulang
Sistem Informasi, Fakultas Ilmu Komputer, UNPAM, Tangerang Selatan
Jl. Raya Puspitek, No.10, Buaran, Serpong, Tangerang Selatan, Banten, Indonesia
e-mail : dosen02830@unpam.ac.id

ABSTRACT

Data classification in the system is widely used in various fields of information technology, data classification is obtained by analyzing objects based on the level of similarity and mapping the characteristics that are protected in the object based on the specified cluster. The use of classification in this study by use texture patterns contained in data by use K-Means Clustering algorithm and Naive Bayes Classifier. The K-Means Clustering algorithm is the process of grouping data based on the same characteristics, while Naive Bayes Classifier use probability and statistic calculations. This study aims to analyze data through texture patterns and the level of accuracy generated through K-Means Clustering algorithms and naïve bayes classifiers. Testing data on textures based on extract parameters with contrast and correlation classifications in GLCM (Gray-Level Co-Occurrence Matrix) methods and calculating the distance expressed in pixels with angle intervals of 0°, 45°, 90°, and 135°. Texture imagery is classified into three clusters by comparing accuracy values obtained. The accuracy value resulting from testing using the K-means Classifier algorithm is very good with an accuracy value of 100% and the accuracy value resulting from testing using the Naive Bayes Classifier method is very good with an accuracy value of 100%.

Keywords: *K-Means Clustering, Naive Bayes, and Classifier*

ABSTRAK

Klasifikasi data dalam sistem banyak digunakan dalam berbagai bidang teknologi informasi, klasifikasi data diperoleh dengan cara menganalisis objek berdasarkan tingkat kesamaan dan memetakan karakteristik yang terkandung dalam objek tersebut berdasarkan *cluster* yang ditentukan. Penggunaan klasifikasi dalam penelitian ini dengan memanfaatkan pola tekstur yang terdapat dalam sebuah data dengan memanfaatkan algoritma *K-Means Clustering* dan *Naive Bayes Classifier*. Algoritma *K-Means Clustering* merupakan proses mengelompokkan data berdasarkan karakteristik yang sama, sedangkan *Naive Bayes Classifier* memanfaatkan perhitungan *probabilitas* dan *statistic*. Penelitian ini bertujuan untuk menganalisis data melalui pola tekstur dan tingkat akurasi yang dihasilkan melalui algoritma *K-Means Clustering* dan algoritma *Naive Bayes Classifier*. Pengujian data pada tekstur berdasarkan parameter ekstrak dengan klasifikasi *contrast* dan *correlation* yang ada pada metode GLCM (*Gray-Level Co-Occurrence Matrix*) serta menghitung jarak yang dinyatakan dalam piksel dengan interval sudut 0°, 45°, 90°, dan 135°. Citra tekstur dilakukan klasifikasi kedalam tiga *cluster* dengan membandingkan nilai akurasi yang diperoleh. Nilai akurasi yang dihasilkan dari pengujian menggunakan algoritma *K-means Clasifier* sangat baik dengan nilai *accuracy* 100% dan nilai akurasi yang dihasilkan dari pengujian menggunakan metode *Naive Bayes Classifier* sangat baik dengan nilai *accuracy* 100%.

Kata Kunci: *K-Means Clustering; Naïve Bayes, dan Classifier*

1. PENDAHULUAN

Pesatnya perkembangan ilmu pengetahuan dan teknologi tidak terlepas dari sebuah penelitian (*Research*). Penelitian dilakukan bertujuan untuk menemukan hal baru atau memperbaiki sebuah penemuan yang sudah ada sebelumnya. Berbagai disiplin ilmu terus berinovasi untuk melakukan perubahan tidak terkecuali dengan inovasi pada teknologi informasi. Teknologi informasi menurut *Information Technology Association of America* (ITAA) merupakan suatu proses pengolahan data, penyimpanan data dan penyebaran informasi yang berupa gambar, *vocal*, teks dan numerik melalui sebuah mikro elektronika dengan basis telekomunikasi dan komputasi. (Purwanto & Goeritno, 2019) Teknologi informasi terdiri dari sebuah perangkat keras (*Hardware*) yang berfungsi sebagai objek untuk input, proses data, penyimpanan data, dan *output* data. Sedangkan perangkat lunak (*Software*) berfungsi sebagai *operation system* dimana melakukan sebuah perintah dan pengoperasian yang diminta oleh user (*Brainware*) melalui perangkat komputer (PC) atau *smartphone*. Perlu kita ketahui dalam sebuah *software* ketika melakukan *processing* terdapat suatu langkah-langkah penyelesaian dalam rangkaian instruksi yang biasa kita kenal dengan istilah algoritma (*Algorithm*). Algoritma (*Algorithm*) ditemukan oleh seorang ilmuwan muslim pada abad ke- 7 dan 8 Masehi yang bernama Abu Muhammad Ibn Musa Al-Khawarizmi atau biasa dikenal dengan sebutan Al-Khawarizmi yang lahir di kota Khawarizm (*Khanate of Khiva*), Uzbekistan. Algoritma terus banyak dikembangkan, dalam berbagai pemecahan masalah dari yang kompleks sampai dengan yang rumit. (Mulyadi, 2018) Adapun manfaat penemuan ini yang sudah kita nikmati diantaranya kecanggihan komputer, internet, penggunaan aplikasi seperti *facebook*, *twitter*, *whatsapp* dan sebagainya.

Berbagai bentuk algoritma yang telah ada, dalam kesempatan ini fokus penelitian terhadap pola tekstur citra digital dalam pengenalan tekstur dengan metode *Gray Level Co-occurrence Matrix* (GLCM). Penggunaan citra digital dalam kebutuhan akan pekerjaan eksperimental dan ekstensif untuk menetapkan kelayakan suatu data banyak digunakan. Dalam penelitian ini dilakukan identifikasi pada citra digital dengan tekstur yang berbeda-beda, dengan implementasi penggunaan metode GLCM pada *software* Matlab R2017b. Analisa citra digital dengan menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM) memanfaatkan sebuah pengukuran melalui proses ekstraksi (*feature extraction*). Indikator pola citra yang digunakan yaitu *contrast*, *corelation*, *energy*, and *homogeneity* dengan *pixel distance* yang telah ditetapkan. (Auliasari & Kertaningtyas, 2018)

Penelitian terhadap citra digital berdasarkan tekstur dengan menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM) ini diharapkan akan mampu mendapatkan hasil dan informasi citra digital yang akurat. Dan dilakukan pengujian lanjutan terkait efektivitas klasifikasi data atau citra digital (gambar) tipe file .jpg atau .gif dengan klasifikasi data yaitu algoritma *K-Means Clustering* dan *Algoritma Naive Bayes Classifier* pada *software* Matlab R2017b.

Sebelumnya pernah dilakukan juga penelitian terkait klasifikasi data menggunakan algoritma *K-Means Clustering* dan *Algoritma Naive Bayes* yang dilakukan oleh Karina Auliasari dan Mariza Kertaningtyas pada tahun 2018, pada penelitian itu dilakukan klasifikasi data pada tekstur batik dan brodatz dengan parameter *contrast and energy*. Adapun hasilnya untuk tekstur batik dengan algoritma *K-Means Clustering* dan *Naive Bayes* yaitu dengan waktu 7,2 ms dan 12 ms, sedangkan pengujian pola tekstur Brodatz dengan hasil catatan waktu 35 ms dan 42,6 ms. (Auliasari & Kertaningtyas, 2018)

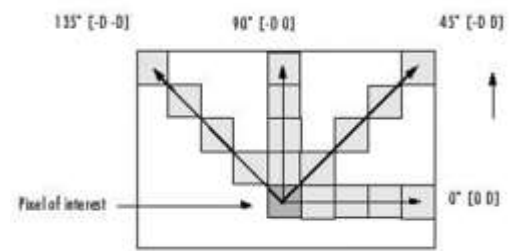
Penelitian lainnya terkait dengan penggunaan algoritma *K-Means Clustering* yaitu penelitian yang dilakukan oleh Tahta Herdian Andika dan Nuari Sivi Anisa pada tahun 2020 yang dimuat dalam *Aisyah Journal of Informatics and Electrical Engineering* dengan identifikasi terhadap citra daun. Proses analisis yang dilakukan hanya pada 5 *sampling* saja dan hasilnya masih menunjukkan adanya kesalahan klasifikasi data pada citra daun rambutan berdasarkan nilai *metric* dan *eccentricity* yang digunakan. (Andika & Anisa, 2020)

Berdasarkan literasi diatas penelitian ini menjadikan pengembangan dari penelitian sebelumnya terkait Analisa tekstur dengan metode GLCM dan penerapan klasifikasi data dengan algoritma *K-Means Clustering* dan *Naïve Bayes Classifier*. Adapun hal yang membedakan dalam penelitian ini yaitu pertama Analisa detail hasil perhitungan dari citra digital berupa nilai dari *contrast, corelation, energy, and homogeneity*. Dan dilanjutkan pada klasifikasi data dengan algoritma *K-Means Clustering* dan *Naïve Bayes Classifier* dengan indikator pola teksur dalam GLCM yaitu *contrast and correlation*.

2. METODOLOGI

2.1 Gray Level Co-occurrence Matrix (GLCM)

Gray Level Co-occurrence Matrix (GLCM) merupakan metode analisis tekstur pada citra. Parameter yang dihitung dalam GLCM berdasarkan parameter jarak dan orientasi sudut (θ). Dalam perhitungan berdasarkan piksel yaitu dengan menghitung banyaknya *distance reference* dan *distance neighbor* (piksel tetangga) berdasarkan sudut (θ) dan jarak d . Sedangkan perhitungan orientasi direpresentasikan dalam derajat dengan interval sudut $0^\circ, 45^\circ, 90^\circ$ dan 135° untuk mendeteksi *co-occurrence* (Ullu, et al., 2022).



Gambar 1 Arah Kookurensi GLCM (Waliyansyah, 2020)

Dalam menghitung GLCM dilakukan dengan beberapa tahap diantaranya yaitu membentuk matriks GLCM dari pasangan piksel yang berjajar sesuai dengan arah $0^\circ, 45^\circ, 90^\circ$ atau 135° , melakukan penjumlahan dengan matriks GLCM dengan nilai transposnya, membagi setiap dari elemen matriks dengan jumlah pasangan piksel, melakukan normalisasi matriks GLCM dengan membagi setiap elemen matriks dengan jumlah pasangan piksel, dan ekstraksi ciri, (Widodo, et al., 2018) yaitu:

$$\text{Contrast} : \sum_{i_1} \sum_{i_2} (i_1 - i_2)^2 p(i_1, i_2) \quad (1)$$

$$\text{Homogeneity} : \sum_{i_1} \sum_{i_2} \frac{p(i_1, i_2)}{1 + |i_1 - i_2|} \quad (2)$$

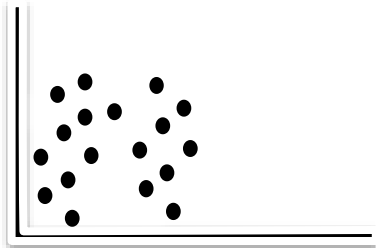
$$\text{Energy} : \sum_{i_1} \sum_{i_2} p(i_1, i_2) \quad (3)$$

$$\text{Entropy} : \sum_{i_1} \sum_{i_2} p(i_1, i_2) \log p(i_1, i_2) \quad (4)$$

2.2 Algoritma *K-Means Clustering*

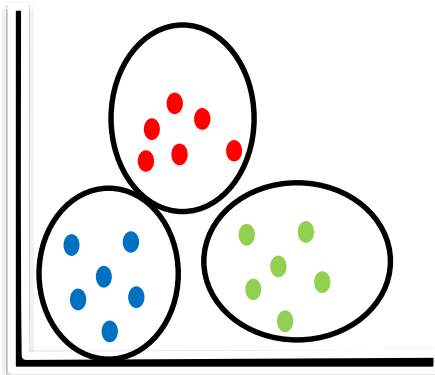
K-Means Clustering merupakan sebuah metode teknik pengelompokan *iteratif* dengan mempartisi data ke dalam sebuah *cluster* dengan karakteristik yang sama (Sari & Ayu, 2019). *K-Means Clustering* juga suatu metode penganalisaan data atau metode data *mining* yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Dengan kata lain, metode *K-Means Clustering* bertujuan untuk meminimalisasikan *objective function* yang di *setting* dalam proses

clustering dengan cara meminimalkan variasi antar data yang ada di dalam suatu *cluster* dan memaksimalkan variasi dengan data yang ada di *cluster* lainnya. Dalam *K-Means Clustering* terdapat dua pengelompokan proses diantaranya yaitu *hierarchical* dan *non hierarchical*. (Irfiani & Rani, 2018)



Gambar 2. Before K-Means

Data setelah dilakukan klasifikasi data menggunakan algoritma *K-means Clustering*.



Gambar 3 After K-Means

Data clustering menggunakan metode *K-Means Clustering* ini secara umum dilakukan dengan algoritma dasar sebagai berikut:

- a. Menentukan jumlah *cluster*.
- b. Melakukan alokasi data pada *cluster* secara acak (Tentukan posisi *Centroid*).

Rumus yang digunakan:

$$V = \frac{\sum_{i=1}^n x_i}{n} ; i = 1, 2, 3, \dots, n \quad (5)$$

Ket. : v = *centroid* pada *cluster*
 (Andika & Anisa, 2020)

x_1 = objek ke -i

n = banyaknya objek/jumlah objek yang menjadi anggota *cluster*.

- c. Menghitung jarak antara *centroid* dengan data yang ada pada setiap *cluster*.

Dalam menghitung jarak antara objek dengan *centroid* menggunakan *Euclidian Distance* dengan rumus sebagai berikut:

$$d(x, y) = ||x - y|| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} ; 1, 2, 3, \dots, n \quad (6)$$

Keterangan : x_i : objek x ke-i

y_i : daya y ke-i

n : banyaknya objek

(Andika & Anisa, 2020)

- d. Melakukan alokasi data pada *centroid* ke *cluster* terdekat.
- e. Apabila masih ada data yang berpindah *cluster* atau apabila perubahan nilai pada *objective function* yang digunakan di atas nilai *threshold* yang ditentukan kembali pada langkah ketiga.
- f. Akan tetapi jika nilai *centroid* tidak berubah dengan nilai sebelumnya maka algoritma berakhir. (Setiawan & Putra, 2018)

2.3 Algoritma *Naive Bayes Classifier*

Naive Bayes Classifier merupakan pengklasifikasi probabilitas sederhana berdasarkan pada teorema Bayes. Keuntungan *Naive Bayes Classifier* adalah hanya membutuhkan sejumlah kecil data pelatihan untuk memperkirakan parameter (sarana dan varian dari variabel) dalam proses klasifikasi. *Naive Bayes Classifier* merupakan metode yang bekerja berdasarkan perhitungan probabilitas dan statistic (Suharyanto & Zein, 2022). Klasifikasi data dilakukan dengan cara memprediksi probabilitas anggota kelas kedalam sebuah kelas tertentu. (Permana, et al., 2018) Probabilitas algoritma *Naive Bayes* dapat dirumuskan seperti dibawah ini:

$$p(C|F_1, \dots, F_n) \quad (7)$$

Penjelasan: C merupakan peubah kelas dependen yang berisi salah satu kelas dari bermacam-

macam kelas sedangkan F_1 sampai F_n adalah fitur masukan. (Auliasari & Kertaningtyas, 2018)

Semua klasifikasi *Naive Bayes* merupakan klasifikasi statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu *class*. Untuk klasifikasi *Bayes* sederhana dikenal sebagai *Naive Bayesian Classifier* yang dapat diasumsikan bahwa efek dari suatu nilai atribut sebuah kelas yang diberikan adalah bebas dari atribut lain. Asumsi ini disebut *class conditional independence* yang dibuat untuk memudahkan dalam perhitungan *Teorema bayes* merupakan dasar aturan dari *naive bayes* dengan asumsi “*naif*” tentang kemandirian bersyarat antara setiap pasang fitur yang diberikan nilai *variable*. Klasifikasi menggunakan *Naive Bayes* lebih cepat dibandingkan dengan metode lain, akan tetapi dikenal sebagai estimator yang kurang baik sehingga dalam penelitian ini melakukan perbandingan dengan metode *K-Means Clustering*. Untuk persamaan teorama *bayes* adalah:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (8)$$

Keterangan :

X : Data dengan *class* yang belum diketahui.

H : Hipotesis data merupakan suatu *class* spesifik.

P(H|X) : *Probabilitas* hipotesis H berdasar kondisi **X** (*posteriori probabilitas*).

P(H) : *Probabilitas* hipotesis H (*prior probabilitas*).

P(X|H) : *Probabilitas* X berdasarkan kondisi pada hipotesis H.

P(X) : *Probabilitas* X.

3. HASIL DAN PEMBAHASAN

Hasil serta pembahasan dalam penelitian ini mencakup proses, langkah-langkah serta pengujian pada tekstur citra digital *file extension* berupa .gif sebanyak 3 sample dengan metode *Gray Level Co-occurrence Matrix* (GLCM). *Output* yang dihasilkan dari pengujian dilakukan

perbandingan nilai dari masing-masing parameter, kedua yaitu pengujian klasifikasi data pada citra digital *file extension* berupa .gif sebanyak 66 *sample* dengan algoritma *K-means Clustering*, dan ketiga yaitu pengujian klasifikasi data pada citra digital *file extension* berupa .gif sebanyak 66 *sample* dengan algoritma *Naive Bayes Classifier*. Dalam proses pengujian klasifikasi data dilakukan dengan menggunakan *software Matlab R2017b* dengan *output* yang dihasilkan berupa grafik dan nilai akurasi dari masing-masing algoritma. Adapun paparan proses dalam penelitian ini adalah sebagai berikut:

3.1 Pengujian Gray Level Co-occurrence Matrix (GLCM)

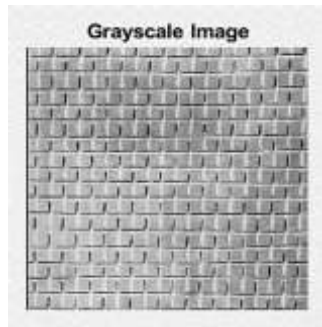
Dalam pengujian ini melakukan uji sample dengan menghitung *contrast, corelation, energy, and homogeneity* dengan piksel (*distance*) dengan skala 0° , 45° , 90° , and 135° . adalah sebagai berikut:

a. Pengukuran GLCM pada Citra Digital:

Untuk mengetahui nilai dari *contrast, corelation, energy, and homogeneity* pada citra digital dilakukan pengkodean menggunakan *software Matlab R2017b*.

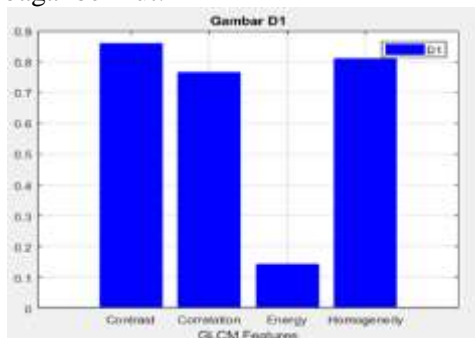
- 1) Pertama menyiapkan citra digital yang sudah dilakukan *pre-processing* dengan melakukan *resize* ukuran gambar, dan melakukan proses *grayscale* yaitu merubah citra keabuan.
- 2) Setelah itu pada tahap kedua melakukan pengujian citra dengan menggunakan nilai piksel (*distance*) -1 (minus satu).
- 3) Langkah selanjutnya melakukan pengujian pada arah *gray level co-occurrence matrix* (GLCM) terhadap akurasi hasil klasifikasi, dengan tujuan untuk mengetahui arah GLCM yang terbaik. Terdapat 4 arah GLCM yang digunakan pada pengujian ini yaitu, 0° , 45° , 90° dan 135° .
- 4) Adapun *sample* yang digunakan dalam pengujian GLCM ada 3 *sample*. engan hasil masing-masing dibawah ini:

b. Pengujian citra digital pada Gambar D1:



Gambar 4. Sample 1 (D1)

Didapatkan hasil dalam bentuk grafik sebagai berikut:



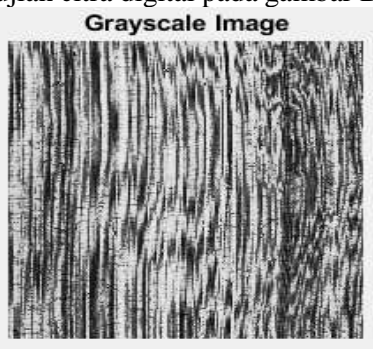
Gambar 5 Grafik hasil pengujian 1

Tampilan hasil perhitungan dengan relevansi sudut 0°, 45°, 90° dan 135°.

Tabel 1 Output Nilai Parameter *Sample 1*

Parameter	0°	45°	90°	135°	Average
Contraast	0,559	1,057	0,721	1,103	0,860
Correlatio n	0,848	0,711	0,803	0,699	0,765
Energy	0,154	0,134	0,154	0,133	0,144
Homogene ity	0,845	0,784	0,832	0,781	0,810

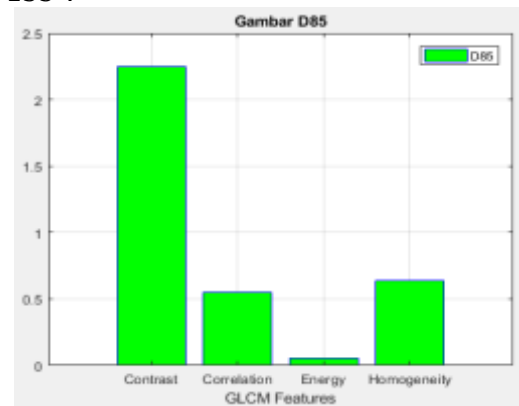
c. Pengujian citra digital pada gambar D85



Gambar 6. Sample 2 (D85)

Hasil dari pengujian tekstre pada citra digital sample 2, dengan menggunakan *software* Matlab 2017b mendapatkan beberapa *output* nilai dari masing-masing parameter dengan piksel (*distance*) -1

dengan *interval* sudut 0°, 45°, 90°, dan 135°.



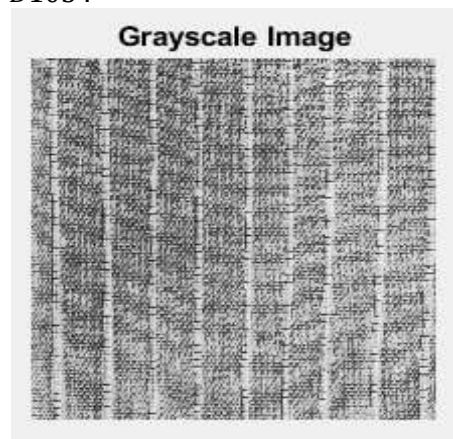
Gambar 7. Grafik pengujian 2

Bentuk rincian *output* nilai dari parameter *Contrast*, *Correlation*, *Energy*, *Homogeneity*.

Tabel 2 Output Nilai Parameter *Sample 2*

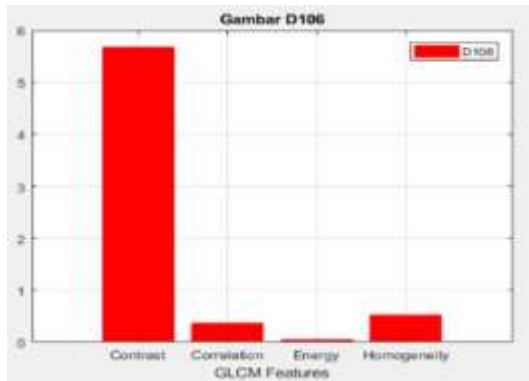
Parameter	0°	45°	90°	135°	Average
Contraast	1,488	2,531	1,837	3,136	2,248
Correlation	0,703	0,495	0,633	0,374	0,551
Energy	0,064	0,051	0,068	0,046	0,057
Homogeneity	0,685	0,613	0,693	0,560	0,638

d. Pengujian citra digital pada Gambar D105 :



Gambar 8. Sample Pengujian 3

Hasil dari pengujian tekstur pada citra digital *sample 2*, dengan menggunakan *software* Matlab 2017b mendapatkan beberapa *output* nilai dari masing-masing parameter dengan piksel (*distance*) -1 dengan *interval* sudut 0°, 45°, 90°, dan 135°.



Gambar 9. Tampilan Grafik Pengujian 2

Bentuk rincian *output* nilai dari parameter *Contrast*, *Correlation*, *Energy*, *Homogeneity*.

Tabel 3. *Output* Parameter *Sample 3*

Parameter	0°	45°	90°	135°	Average
Contraast	6,903	7,302	1,259	7,207	2,668
Correlation	0,213	0,167	0,856	0,178	0,354
Energy	0,027	0,026	0,061	0,026	0,035
Homogeneity	0,461	0,450	0,721	0,452	0,521

e. Kesimpulan pengujian citra digital menggunakan GLCM

Dari penelitian terhadap tekstur citra digital yang telah dipaparkan diatas dapat disimpulkan bahwa penggunaan metode GLCM dalam membedakan citra digital melalui tekstur yang ada pada *sample 1*, *sample 2*, dan *sample 3* sangat efisien dan akurat. Rata-rata nilai dari citra yang telah dilakukan pengujian nilai pada parameter *Contrast* dan *Corelation* lebih cenderung tinggi diatas indikator *energy*. Maka dari itu parameter *Contrast* dan *Corelation* dijadikan sebuah indikator klasifikasi tekstur pada pengujian klasifikasi data menggunakan algoritma *K-Means Clustering* dan *Naive Bayes Classifier*.

Penelitian pada citra digital berdasarkan analisa tekstur menggunakan Metode *Gray Level Co-occurrence Matrix* (GLCM) dengan relevansi sudut 0°, 45°, 90° dan 135° dan skala piksel (*distance*) - 1 pada *sample 1*, *sample 2*, dan *sample 3* terdapat perbedaan nilai parameter yang dihasilkan. Adapun nilai yang dihasilkan cukup akurat, ketika coba merubah indikator skala piksel (*distance*) menjadi 1 dan 2 nilai yang dihasilkan pada parameter berubah akan tetapi perubahan

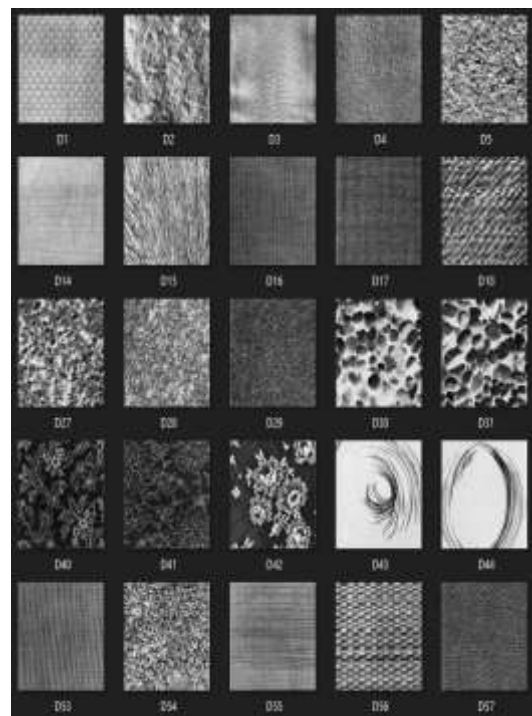
nilai relatif sama dari ketiga *sample* tersebut.

Tabel 4. Perbandingan Hasil Pengujian

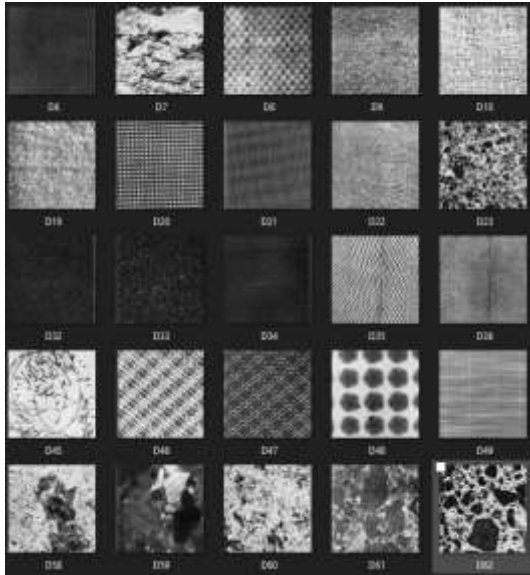
Citra	Average			
	Contrast	Corelation	Energy	Homogeneity
Sample 1	0,860	0,765	0,144	0,810
Sample 2	2,248	0,551	0,057	0,638
Sample 3	5,668	0,354	0,035	0,521

3.2 Pengujian Klasifikasi data dengan menggunakan Algoritma *K-Means Clustering* dan *Naive Bayes Classifier*

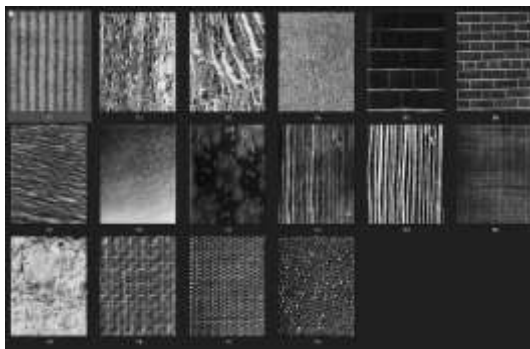
Fokus kedua dalam penelitian ini yaitu melakukan pengujian klasifikasi data citra digital menggunakan algoritma *K-Means Clustering* dan algoritma *Naive Bayes Classifier* terhadap citra digital yang telah dilakukan *grayscale* dengan *extensi file .gif* pada 66 *sample* dengan bahasa pemrograman dan *software* Matlab 2017b. Berdasarkan analisa tekstur Metode *Gray Level Co-occurrence Matrix* (GLCM). Dengan memanfaatkan parameter nilai *Contrast* dan *Corelation* pada jarak piksel (*distance*) -1 dan interval sudut 0°, 45°, 90°, dan 135°.



Gambar 10. *Sample*



Gambar 11. Sample



Gambar 12 Sample

Keluaran (*output*) yang dihasilkan setelah dilakukan pengujian terhadap 66 *sample* dengan menggunakan Algoritma *K-Means Clustering and Bayes Naive Classifier* pada *software* Matlab 2017b didapatkan nilai *accuracy* dari 5 *replication*.

Hasil pengujian data

Replicate 1,5 iterations, total sum of distances = 12.0438.

Replicate 2,1 iterations, total sum of distances = 12.2688.

Replicate 3,1 iterations, total sum of distances = 12.2688.

Replicate 4,1 iterations, total sum of distances = 11.7977.

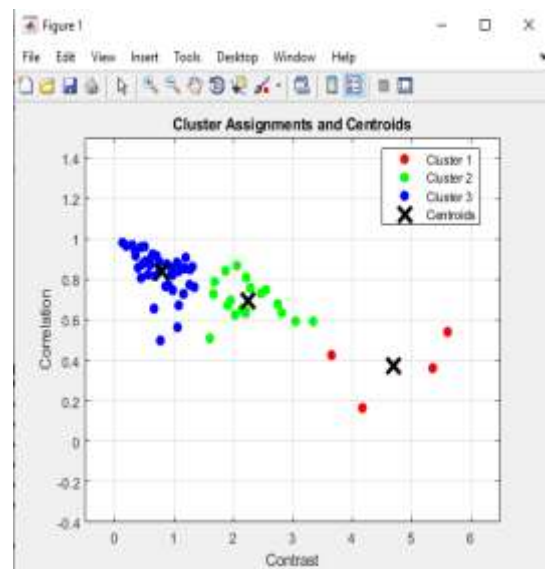
Replicate 5,2 iterations, total sum of distances = 11.7977.

Best total sum of distances = 11.7977
accuracy = 100

Tabel 5. *Output Best Total Sum Of Distances*

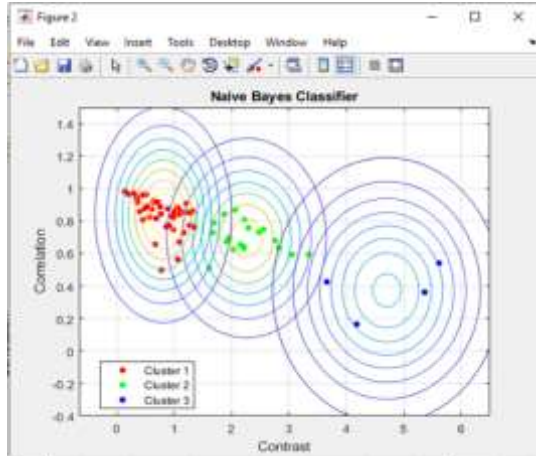
Replicate	Iterations	Sum Of Distances
1	5	12.0438.
2	1	12.2688
3	1	12.2688
4	1	11.7977
5	2	11.7977
Best total sum of distances		11.7977
Accuracy		100

Hasil uji data citra digital yang dilakukan menggunakan Algoritma *K-Means Clustering* terlihat data terbagi kedalam tiga *cluster* sesuai dengan nilai parameter *Contrast* dan *Corelation* pada jarak piksel (*distance*) -1 dan interval sudut 0°, 45°, 90°, dan 135°. Klasifikasi menyesuaikan dengan nilai parameter yang sama dan data terdekat yang membentuk sebuah *centroid*. Dan menunjukkan hasil dengan *accuracy* 100%.



Gambar 13. Algoritma *K-Means Clustering*

Hasil pengujian kalsifikasi data citra digital menggunakan algoritma *Naife Bayes Clasifier* data berhasil ter *cluster* dengan baik sesuai nilai piksel (*distance*) - 1 dari parameter *contrast* dan *correlation* menunjukkan *accuracy* 100%.



Gambar 14. Naife Bayes Classifier

4. KESIMPULAN

Dari rangkaian penelitian yang dilakukan pada objek data citra digital berdasarkan tekstur dengan metode *gray level co-occurrence matrix* (GLCM) pada tiga *sample*. Dan selanjutnya dilakukan klasifikasi data pada citra digital 66 *sample* menggunakan algoritma algoritma *K-Means Clustering* dan algoritma *naive bayes classifier* adalah Pengujian tekstur menggunakan metode *gray level co-occurrence matrix* (GLCM) dengan relevansi sudut 0° , 45° , 90° dan 135° dan skala (*distance*) -1 nilai yang dihasilkan menunjukkan performa yang baik dalam analisa tekstur. Pengujian pada citra digital berdasarkan analisa tekstur dengan metode *gray level co-occurrence matrix* (GLCM) penggunaan nilai skala (*distance*) -1,1 dan 2 berpengaruh pada nilai parameter *contrast*, *corelation*, *energy*, and *homogeneity* yang dihasilkan. Akan tetapi perubahan nilai relatif sama dari ketiga *sample* tersebut. Nilai akurasi yang dihasilkan dari pengujian menggunakan algoritma *K-means Clasifier* sangat baik dengan nilai *accuracy* 100%. Nilai akurasi yang dihasilkan dari pengujian menggunakan metode *Naive Bayes Classifier* sangat baik dengan nilai *accuracy* 100%.

DAFTAR PUSTAKA

Auliasari, K. & Kertaningtyas, M., 2018. Studi Komparasi Pola Tekstur Citra Digital Menggunakan Metode K-Means dan Naive Bayes. *Jurnal Informatika*,

Volume 18 No.02, Desember 2018, pp. 175-185.

Surya, R. A., Fadlil, A. & Yudhana, A., 2017. Ekstrasi Ciri Metode Gray Level Co-Occurrence Matrix (GLCM) dan Filter Gabor untuk klasifikasi Citra Batik Pekalongan. *Jurnal Pengembangan IT*, Volume 02, p. 26.

Widodo, R., Widodo, A. W. & Supriyanto, A., 2018. Memanfaatkan Ciri GLCM Citra Buah Jeruk Keprok (*Citrus reticulata* Blanco) untuk mengklasifikasi Mutu. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Volume 2, pp. 5769-5776.

Setiawan, K. N. & Putra, M. S., 2018. Klasifikasi Citra Mammogram Menggunakan Metode K-Means, GLCM, dan Support Vector Machine (SVM). *Merpati*, Volume 6, pp. 13-24.

Saleh, A., 2015. Penerapan Metode Klasifikasi Naive Bayes Dalam Prakira Penggunaan Listrik Rumah Tangga. *Citec Journal*, Volume 2, No.3, Mei 2015-Juni 2015, p. 2015.

Permana, A., Bhakti, R. M. H. & Prayogi, D., 2018. Segmentasi K-Means Clustering Pada Citra Menggunakan Ekstraksi Fitur Warna dan Tekstur. *Jurnal Ilmiah Intech, Information Technology Journal Of UMUS*, Volume 2, No.01, pp. 89-97.

Mulyadi, A., 2018. Pemikiran Al-Khawarizmi Dalam Meletakkan Dasar Pengembangan Ilmu Astronomi Islam. *International Journal Ihya' Ulum Al-Din*, Volume Vol 20 No 1 (2018), pp. 63-86.

Purwanto, E. H. & Goeritno, A., 2019. Pemanfaatan Teknologi Informasi untuk Pembaruan Peta Sarana Desa di Desa Pemagarsari, Parung, Bogor. *Jurnal Pengabdian Pada Masyarakat*, Volume 4, No.3, pp. 329-340.

Irfiani, E. & Rani, S. S., 2018. Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita. *Jurnal Sistem dan*

Teknologi Informasi, Volume Vol. 6, No. 4, pp. 165-172.

Waliyansyah, A. R., 2020. Identifikasi Jenis Biji Kedelai (Glycine Max L) Menggunakan Gray Level Coocurance Matrix (GlcM) Dan K-Means Clustering. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, Volume Vol. 7, No. 1, pp. 17-26.

Andika, T. H. & Anisa, N. S., 2020. Sistem Identifikasi Citra Daun Berbasis Segmentasi Dengan Menggunakan Metode K-Means Clustering. *Aisyah Journal of Informatics and Electrical Engineering*, 02(1), pp. 9-17.

Ullu, H. H. et al., 2022. Ekstraksi Fitur Berbasis Tekstur Pada Citra Tenun Timor Menggunakan Metode Gray Level Co-occurrence Matrix (GLCM). *Journal Of Information and Technology Unimor*, pp. 70-74.

Sari, D. W. & Ayu, K. G., 2019. Penentuan Kriteria Dalam Memilih Sekolah Dasar Dengan Menerapkan K-Means Clustering. *Sainstech ISTN*, pp. 24-28.

Suharyanto, E. & Zein, A., 2022. Analisis data minat calon mahasiswa universitas pamulang dengan menggunakan algoritma naive bayes classifier. *Sainstech ISTN*, pp. 70-76.