

Metode K-Nearest Neighbor dan Ekstraksi Fitur GLCM untuk Mengklasifikasikan Biji Kopi Robusta dan Arabika Lokal

Cahaya Jatmoko¹, Daurat Sinaga²

^{1,2} Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

e-mail: jatmoko14@dsn.dinus.ac.id, dauratsinaga@dsn.dinus.ac.id

Abstract

Some well-known coffee products are Robusta Coffee and Arabica Coffee. Utilizing technology to help identify how the appearance of coffee beans differs is one of the issues that needs to be tested, in fact, identification using image data can also be carried out optimally. The identification process is carried out by utilizing digital image processing techniques. In this study, the K-Nearest Neighbor (K-NN) Algorithm and Gray Level Co-occurrence Matrix (GLCM) Feature Extraction have been used to classify Robusta and Arabica coffee beans. An image dataset of 194 images of coffee beans is used which is a combination of 97 image data from Robusta coffee beans and 97 image data from Arabica coffee beans. Of the entire dataset, 174 data were used as training data, and the remaining 20 data were used as test data. The highest accuracy resulting from this experiment is 95% provided that the pixel distance = 1, the K value = 1, and the angle size = 45°.

Keywords: Coffee Beans, Robusta, Arabica, K-Nearest Neighbor, Gray Level Co-occurrence Matrix.

Abstrak

Beberapa produk kopi yang cukup terkenal adalah Kopi Robusta dan Kopi Arabika. Memanfaatkan teknologi untuk membantu mengidentifikasi bagaimana perbedaan penampilan biji kopi menjadi salah satu isu yang perlu di uji coba, bahkan, pengidentifikasian menggunakan data citra juga dapat dilakukan dengan optimal. Proses pengidentifikasian tersebut dilakukan dengan cara memanfaatkan teknik pengolahan citra digital. Pada penelitian ini, telah digunakan Algoritma K-Nearest Neighbor (K-NN) dan Ekstraksi Fitur Gray Level Co-occurrence Matrix (GLCM) untuk melakukan proses pengklasifikasian biji kopi Robusta dan Arabika. Digunakan dataset gambar sebanyak 194 gambar biji kopi yang merupakan gabungan dari 97 data citra dari biji kopi Robusta, dan 97 data citra dari biji kopi Arabika. Dari keseluruhan dataset tersebut, digunakan sebanyak 174 data sebagai data latih, dan 20 data sisanya sebagai data uji. Akurasi tertinggi yang dihasilkan dari eksperimen ini sebesar 95% dengan ketentuan jarak piksel = 1, nilai K = 1, dan besar sudut = 45°.

Kata Kunci: Biji Kopi, Robusta, Arabika, K-Nearest Neighbor, Gray Level Co-occurrence Matrix.

1. PENDAHULUAN

Di pasaran dunia, tanaman kopi memiliki nilai ekonomis yang cukup tinggi pada komoditas ekspor. Indonesia sendiri merupakan salah satu negara utama yang mengekspor kopi [ginting, 2019]. Studi kasus yang dilakukan di Kab. Batang, didapatkan informasi bahwa banyak masyarakat di Kab. Batang memproduksi kopi, terutama kopi robusta dan arabika lokal. Namun, masih banyak orang yang kesulitan membedakan antara kopi robusta dan arabika dalam penampilan visualnya. Dengan bantuan teknologi, dimanfaatkan metode klasifikasi dari pengimplementasian Algoritma K-Nearest Neighbor dan Ekstraksi Fitur GLCM untuk membantu orang-orang membedakan penampilan visual dari kopi robusta dengan kopi arabika.

Klasifikasi adalah metode pengelompokan suatu objek ke dalam kelas sejenisnya. Sebelum dilakukannya proses klasifikasi, dilakukan proses analisis terlebih dahulu terhadap sekumpulan data, yang kemudian setelah ditemukan suatu pola dari proses analisis tersebut akan dijadikan patokan sebagai proses klasifikasi dari data uji. Dari proses tersebut, klasifikasi termasuk ke dalam Supervised Learning [1]–[3]. Tahap yang dilakukan klasifikasi ada dua, pertama tahap pembelajaran dan kedua tahap klasifikasi. Tahap pembelajaran merupakan proses pengimplementasian data latih menggunakan algoritma klasifikasi. Sedangkan tahap klasifikasi, data uji akan diklasifikasikan bersamaan dengan data latih, yang kemudian dicari nilai akurasi sehingga akan didapatkan hasil kesimpulan dari nilai akurasi yang didapat [4]–[6].

Pada eksperimen yang dilakukan oleh Arthur Ahmad Fauzi, Fitri Utamingrum, dan Fatwa Ramdani pada tahun 2020 [7], mengimplementasikan Algoritma K-Nearest Neighbor dan Ekstraksi Fitur Gray Level Co-occurrence Matrix dan Local Binary Pattern pada Teknologi UGV atau Autonomous Ground Vehicle. Teknologi UGV merupakan pengimplementasian dari proses klasifikasi yang dilakukan di jalan, salah satu contoh kejadian yang diklasifikasi adalah seberapa cepat kecepatan kendaraan sesuai dengan kondisi jalan yang ada. Eksperimen ini dilakukan dengan tujuan dapat membantu pengendara mengatur kecepatan kendaraannya sesuai dengan kondisi yang ada. Eksperimen ini melakukan pencarian fitur yang paling diskriminan tetapi tetap menjaga penggunaan jumlah fitur seminimal mungkin supaya proses komputasi dapat dilakukan dengan cepat dan tetap didapatkan hasil klasifikasi yang akurat. Akurasi yang didapatkan dari eksperimen ini sebesar 98% dengan kecepatan proses klasifikasi gambar sebanyak 12 gambar per detik. Pada eksperimen yang dilakukan oleh R. A. Saputra dkk pada tahun

2020 [8], mengimplementasikan Algoritma K-Nearest Neighbor dan Ekstraksi Fitur Gray Level Co-occurrence Matrix untuk melakukan proses klasifikasi penyakit daun padi. Penelitian tersebut dilakukan dengan tujuan dapat mengurangi resiko gagal panen dengan cara melakukan deteksi dini terhadap kesehatan daun padi dengan memanfaatkan teknologi. Digunakan sebanyak 120 data citra yang didapatkan dari Repositori UCI. Nilai ciri dari Ekstraksi Fitur GLCM yang digunakan pada penelitian ini adalah nilai kontras, energi, entropi, homogenitas, dan korelasi. Sementara untuk Algoritma K-NN, penggunaan nilai K dibatasi hanya dari 1 hingga 20. Hasil akurasi tertinggi yang didapatkan dari eksperimen ini sebesar 65,83% yang didapat dari nilai K=11, lalu nilai akurasi Kappa sebesar 0,485.

Penelitian yang dilakukan oleh Dinar Putra Pamungkas pada tahun 2019 [9] membahas tentang pengimplementasian Metode KNN dan Ekstraksi Fitur GLCM untuk mengidentifikasi jenis bunga anggrek. Bagian unik dari bunga anggrek yang akan dijadikan sebagai objek klasifikasi pada penelitian ini adalah bagian bibi bunga atau Labellum dari bunga anggrek. Tahapan penelitian ini diawali dengan mengubah dan menyamakan ukuran awal citra, kemudian mengkonversikan citra RGBnya menjadi citra Grayscale, menghitung Median Filter, dan dilanjutkan dengan proses ekstraksi fitur menggunakan GLCM, dan terakhir dilakukan proses identifikasi menggunakan KNN. Tingkat keberhasilan yang didapatkan dari hasil penelitian ini adalah dengan didaparkannya hasil akurasi tertinggi sebesar 80% dengan rata-rata akurasi sebesar 77%. Pada penelitian ini pula nilai K cukup berpengaruh pada hasil yang didapatkan, semakin besar nilai K yang digunakan, semakin kecil hasil akurasi yang didapatkan.

Penelitian yang dilakukan oleh Muhammad Ainnur Rizal pada tahun 2019 [10], mengimplementasikan Metode K-NN untuk mengklasifikasikan mutu dari biji kopi. Proses klasifikasi dilakukan berdasarkan dari warna dan tekstur biji kopi. Tujuan dari dilakukannya penelitian tersebut adalah untuk mengidentifikasi dini terhadap kecacatan biji kopi yang akan berpengaruh pada kualitas biji kopi ketika akan di ekspor. Kualitas dari biji kopi dan penilaian terhadap kecacatan yang dimiliki biji kopi disesuaikan dengan ketetapan standar mutu oleh Badan Standar Nasional. Digunakan sebanyak 60 data latih dan 15 data uji. Hasil akurasi tertinggi yang didapatkan adalah sebesar 66,67% yang didapatkan dari nilai K=1, 5, dan 7, dengan rata-rata nilai akurasi sebesar 65%.

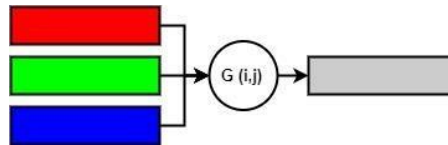
Penelitian yang dilakukan oleh Dedy Ikhsan, Ema Utami, dan Ferry Wahyu Wibowo pada tahun 2020 [11], membahas tentang pengimplementasian Algoritma K-NN untuk mengklasifikasikan mutu dari Greenbean Kopi Arabika Lanang berdasarkan bentuknya. Digunakan empat kelas untuk proses klasifikasi, yaitu Lanang Arabica utuh, Lanang Arabika pecah, Arabika utuh, dan Arabika pecah. Digunakan sebanyak 120 data citra, dimana 80 data citra sebagai data latih dan 40 data citra sisanya sebagai data uji. Dari hasil penelitian ini, didapatkan kesimpulan bahwa Algoritma KNN mampu mengklasifikasikan ke-empat kelas dengan nilai akurasi 63,5%. Selain itu, akurasi yang didapat dari hasil klasifikasi pada kelas Arabika utuh sebesar 93% dan pada kelas Lanang Arabika utuh sebesar 97%. Namun, penelitian ini masih menghasilkan hasil yang kurang baik pada kelas kopi pecah.

2. METODE PENELITIAN

2.1. RGB menjadi Grayscale

Proses konversi citra yang sebelumnya merupakan citra RGB menjadi citra Grayscale merupakan proses reduksi dimensi. Reduksi dimensi yang dimaksud adalah dari tiga dimensi yang ada (R, G, B) direduksi menjadi satu dimensi (grayscale), satu dimensi tersebut kemudian dimanipulasi sehingga bermanfaat meringankan beban komputasi dalam proses klasifikasi nantinya [12], [13]. Proses grayscale ini juga memiliki tujuan lain, yaitu untuk mengintegrasikan data sehingga dapat digunakan dengan optimal pada proses ekstraksi fitur GLCM. Intensitas setiap piksel grayscale adalah, pada nilai 0 akan berubah menjadi hitam, sedangkan pada nilai 255 akan menjadi putih [14]. Rumus yang digunakan ditunjukkan pada persamaan (1). Proses sederhana yang semula berupa RGB sehingga dapat menjadi citra Grayscale ditunjukkan pada Gambar 1. Dimana setiap objek berwarna yang ditampilkan pada Gambar 1 merepresentasikan Red, Green, Blue, dan Gray.

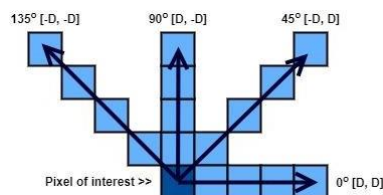
$$G(i, j) = \frac{B(i,j)+G(i,j)+R(i,j)}{3} \quad (1)$$



Gambar 1. Ilustrasi dari proses konversi citra RGB menjadi Grayscale

2.2. Gray Level Co-occurrence (GLCM)

Pada tahun 1973, Haralick mengusulkan suatu teknik lanjutan, yaitu Ekstraksi Fitur Gray Level Co- occurrence Matrix. Nilai (i) pada GLCM merupakan nilai piksel referensi, sedangkan nilai (j) dikenal sebagai tetangga. Setiap elemen (i,j) pada matriks merupakan elemen yang menandakan banyaknya kemunculan pasangan piksel dengan nilai i dan nilai j yang berada pada jarak yang cukup dekat. Hubungan spasial antara dua nilai piksel yang bertetangga atau berjarak dekat dapat ditentukan dengan beberapa cara dari penggunaan offset dan sudut yang berbeda. Keempat sudut yang berlaku pada GLCM adalah 0o, 45o, 90o, dan 135o [15]–[18]. Gambar 2 menampilkan bagaimana arah sudut yang berlaku pada GLCM.



Gambar 2. Arah sudut GLCM untuk proses ekstraksi fitur

GLCM melakukan proses pengelompokkan berdasarkan warna. Setelah proses pengelompokkan warna selesai, selanjutnya Menyusun dan membentuk Co-occurrence Matrix. Co-occurrence Matrix dibuat supaya dapat dilakukan proses perhitungan penjumlahan dengan piksel tetangga atau yang berdekatan. Setelahnya, dilakukan proses normalisasi terhadap Matrix GLCM yang dilakukan dengan cara melakukan pembagian antara elemen matriks dengan jumlah seluruh elemen pembentuknya (N). Setelah nilai hasil dari proses normalisasi didapatkan, dilanjutkan dengan mencari nilai probabilitas dengan menggunakan empat fitur perhitungan yang ada pada GLCM dan digunakan pada penelitian ini, yaitu nilai Kontras, Korelasi, Energi, dan Homogenitas [1], [19], [20]. Nilai kontras merupakan nilai dari hasil perhitungan perbedaan intensitas warna keabuan antara satu piksel dengan piksel yang lainnya secara berdekatan (tetangga). Rumus kontras ditunjukkan pada persamaan (2), dimana x adalah nilai baris pada matriks, y adalah nilai kolom pada matriks, dan $p(x,y)$ adalah nilai dari elemen kookurensi baris (x) dan kolom (y).

$$\sum_{x,y} (x - y)^2 p(x, y) \quad (2)$$

Nilai Korelasi dapat menunjukkan ukuran dari ketergantungan linear derajat grayscale gambar, nilai tersebut akan menunjukkan adanya struktur linear yang ada didalam gambar digital. Rumus korelasi ditunjukkan pada persamaan (3), dimana μ_x , μ_y adalah nilai rata-rata pada elemen baris dan kolom dan

$\sigma_x \sigma_y$ adalah nilai standar deviasi pada baris dan kolom. Nilai Energi dapat menunjukkan seberapa tinggi nilai kemiripan dari nilai energi (piksel lain) yang juga sama tinggi. Rumus energi ditunjukkan pada persamaan (4). Nilai Homogenitas dapat menunjukkan jumlah level grayscale pada sebuah citra, petunjuk tersebut ditandai jika nilai pikselnya sama maka nilai homogenitas menjadi tinggi. Rumus homogenitas ditunjukkan pada persamaan (5).

$$\sum_{x,y} (x - \mu_x)(y - \mu_y) p(x, y) \quad (3)$$

$$\frac{\sigma_x \sigma_y}{\sum_{x,y} p(x - y)^2} \quad (4)$$

$$\frac{p(x, y)}{\sum_{x,y} 1 + |x - y|} \quad (5)$$

2.3. K-Nearest Neighbor (K-NN)

K-Nearest Neighbor merupakan algoritma dalam kategori Supervised Learning [9], [21], [22]. Supervised Learning melakukan pembelajaran dengan cara menghubungkan antara pola data yang sudah ada dengan data baru sehingga didapatkan pola data yang baru [Johan, 20]. Proses klasifikasi didasarkan pada data pembelajaran yang memiliki jarak terdekat, yang dimana jarak terdekat dapat diartikan memiliki persamaan ciri antar objek yang banyak [rachmawanto, 2018 (ref. seminar 1)]. Penggunaan nilai K pada Algoritma K-NN dapat memberikan hasil terbaik tergantung pada data yang digunakan. Secara umum, semakin tinggi nilai K yang digunakan pada proses klasifikasi, jumlah pengurangan efek noise juga akan semakin tinggi. Namun hal ini tidak selalu memberikan hasil terbaik [Wijaya, 2020]. Perhitungan jarak yang akan digunakan adalah Euclidean Distance, yang dimana rumuskan ditunjukkan pada persamaan (6). Dimana d adalah jarak Euclidean Distance, x adalah nilai citra dari data latih, y adalah nilai citra dari data uji, i adalah jumlah ketetanggaan, dan n adalah jumlah dari banyaknya citra yang digunakan.

$$d(x, y) = \sqrt{\sum n(xi - yi)^2} \quad (6)$$

Penggunaan Algoritma ini bertujuan untuk mengklasifikasikan objek penelitian sesuai dengan ciri dan data dari sampel yang ada. Objek penelitian direpresentasikan dengan titik, dengan diberi titik uji, dapat didapatkan sejumlah titik latih yang berada berdekatan atau bertetangga dengan titik tersebut. Kelebihan yang dimiliki K-NN adalah algoritma tersebut dapat menjadi lebih kuat dan menghasilkan keberhasilan yang tinggi dengan syarat jumlah data training yang digunakan banyak [23], [24]. Dalam melakukan prosesnya, Algoritma K-NN berfokus pada nilai yang memiliki jarak terdekat antara nilai satu dengan yang lainnya dan kemudian dilakukan pengelompokan dengan tujuan mencari kesamaan.

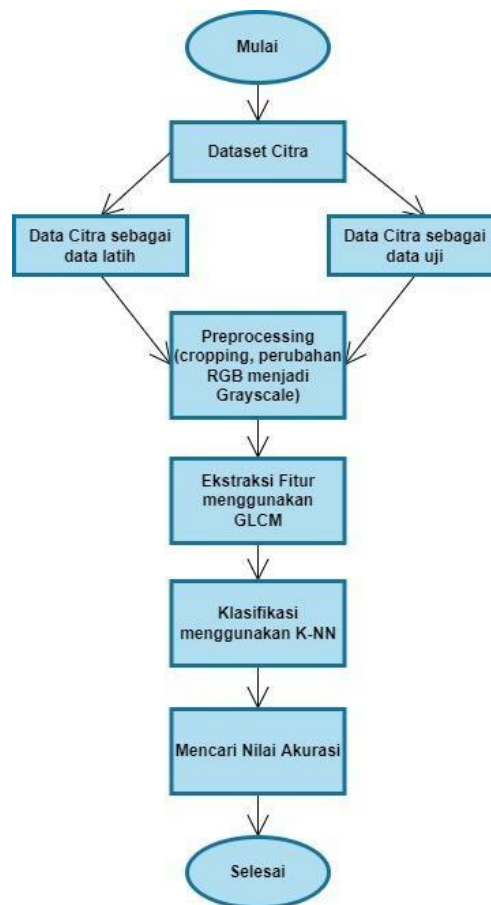
2.4. Kebutuhan Data

Digunakan sebanyak 197 data yang merupakan total dari seluruh data citra yang digunakan. Masing-masing data citra dari biji kopi robusta dan arabika berjumlah 97 data citra. Data yang akan digunakan sebagai data latih berjumlah 177 data, dan yang akan digunakan sebagai data uji sebanyak 20 data. Dari keseluruhan data citra yang sudah terkumpul, hal pertama yang dilakukan adalah dengan melakukan preprocessing data, yaitu proses cropping. Proses cropping yang dilakukan pada data citra awal bertujuan untuk menyamaratakan ukuran data citra antara yang satu dengan yang lainnya, hal ini dilakukan supaya dapat mengurangi eror pada saat proses klasifikasi. Tahap selanjutnya adalah melakukan proses perubahan range warna pada data citra dari RGB menjadi Grayscale. Setelah data citra menjadi data Grayscale, selanjutnya dilakukan ekstraksi fitur dengan menggunakan GLCM dan dilakukan perhitungan nilai Kontras, Korelasi, Energi, dan Homogenitas. Selain itu, sudut yang akan digunakan adalah 0°, 45°, 90°, dan 135° dengan jarak ketetanggaan ($d \leq 1$). Nilai K yang akan digunakan pada penelitian ini adalah $K = 1$, $K = 3$, $K = 5$, $K = 7$, dan $K = 9$. Tahap akhir, dilakukan perhitungan jarak menggunakan Euclidean Distance sehingga didapatkan hasil akurasi akhir. Alur proses ekstraksi fitur ditampilkan pada Gambar 3.



Gambar 3. Alur ekstraksi fitur menggunakan GLCM

2.5. Metode yang Diusulkan



Gambar 4. Alur Metode yang diusulkan

Digunakan data digital yang merupakan data gambar dari objek penelitian, yaitu biji kopi. Data gambar yang digunakan memiliki ekstensi .jpg. Sebelum melalui proses klasifikasi, data gambar terlebih dahulu dilakukan proses cropping dan kemudian dilakukan perubahan warna yang sebelumnya adalah RGB menjadi Grayscale. Selanjutnya, data gambar dilakukan proses ekstraksi dengan menggunakan Ekstraksi Fitur GLCM, baru setelahnya dilakukan proses klasifikasi menggunakan Algoritma K-NN. Langkah terakhir adalah mencari nilai akurasi untuk menentukan seberapa baik eksperimen yang dilakukan. Secara garis besar, alur dari metode yang diusulkan ditampilkan pada Gambar 4.

3. HASIL DAN PEMBAHASAN

3.1. Hasil Implementasi Ekstraksi Fitur GLCM pada Data Latih

Berdasarkan Tabel 1, seluruh kelas sudah dikelompokkan berdasarkan range nilainya dan disesuaikan dengan data citra aslinya

Tabel 1. Hasil Implementasi

Citra Latih	Sudut	Ekstraksi Fitur GLCM				Kelas
		Contras	Correlation	Energy	Homogeneity	
Arabika (1).jpg	0°	0.027493	0.98936	0.38439	0.98627	Arabika
	45°	0.040342	0.9844	0.37634	0.98003	Arabika
	90°	0.030548	0.98818	0.38146	0.98477	Arabika
	135°	0.041766	0.98385	0.3751	0.97917	Arabika
Robusta (1).jpg	0°	0.0071331	0.9972	0.73329	0.99643	Robusta
	45°	0.010004	0.99607	0.73209	0.99501	Robusta
	90°	0.0072254	0.99716	0.73335	0.99639	Robusta
	135°	0.010046	0.99606	0.73228	0.99508	Robusta
Arabika (2).jpg	0°	0.0094114	0.99696	0.81814	0.99542	Arabika
	45°	0.012028	0.99612	0.81755	0.9942	Arabika
	90°	0.0078912	0.99745	0.81835	0.9961	Arabika
	135°	0.013106	0.99577	0.81744	0.99389	Arabika
Robusta (2).jpg	0°	0.011689	0.99622	0.64863	0.99416	Robusta
	45°	0.016133	0.99479	0.6461	0.99195	Robusta
	90°	0.011447	0.9963	0.64913	0.99429	Robusta
	135°	0.015959	0.99484	0.64665	0.99208	Robusta

3.2. Hasil Implementasi K-Nearest Neighbor

Berikut pada Tabel 2 akan ditampilkan sampel dari hasil perhitungan K-NN menggunakan nilai K=1 terhadap sudut 0°

Tabel 2. Sampel perhitungan terhadap nilai K=1 pada sudut 0°

Data Citra	Sudut	Ekstraksi Fitur GLCM				Kelas	Jarak Euclidean
		Contras	Correlation	Energy	Homogeneity		
Arabika(1), Robusta(97)	0°	0.027493	0.98936	0.38439	0.98627	Arabika	0.3994
Robusta(1), Robusta(97)	0°	0.0071331	6,925	0.73329	0.99643	Robusta	0.0521
Arabika(2), Robusta(97)	0°	0.0094114	0.99696	0.81814	0.99542	Arabika	0.0361
Robusta(1), Robusta(97)	0°	0.011689	0.99622	0.64863	0.99416	Robusta	0.1353

3.3. Hasil Pengujian

Berdasarkan Tabel 3 ditampilkan hasil dari tahap pengujian untuk sudut 0°, 45°, 90°, dan 135°. Data yang diberi warna merah merupakan hasil pengujian yang memberikan hasil klasifikasi salah. Setelah didapatkan keseluruhan hasil pengujian, selanjutnya dilakukan perhitungan hasil akurasi menggunakan Confusion Matrix pada setiap sudut. Dengan menggunakan persamaan (3).

$$\text{Confusion Matrix} = \frac{\text{jumlah citra uji benar} \times 100\%}{\text{jumlah total citra uji}} \quad (3)$$

Tabel 3. Hasil Pengujian menggunakan Citra Uji

Citra Uji	Hasil Aktual	Prediksi			
		Sudut 0°	Sudut 45°	Sudut 90°	Sudut 135°
Arabika (88).jpg	Arabika	Arabika	Arabika	Arabika	Arabika
Arabika (89).jpg	Arabika	Arabika	Arabika	Arabika	Arabika
Arabika (90).jpg	Arabika	Arabika	Arabika	Arabika	Arabika
Arabika (91).jpg	Arabika	Arabika	Arabika	Arabika	Arabika
Arabika (92).jpg	Arabika	Arabika	Arabika	Arabika	Arabika
Arabika (93).jpg	Arabika	Arabika	Arabika	Arabika	Arabika
Arabika (94).jpg	Arabika	Arabika	Arabika	Arabika	Arabika
Arabika (95).jpg	Arabika	Arabika	Arabika	Arabika	Arabika
Arabika (96).jpg	Arabika	Arabika	Arabika	Arabika	Arabika
Arabika (97).jpg	Arabika	Robusta	Arabika	Robusta	Robusta
Robusta (88).jpg	Robusta	Robusta	Robusta	Robusta	Robusta
Robusta (89).jpg	Robusta	Robusta	Robusta	Robusta	Robusta
Robusta (90).jpg	Robusta	Robusta	Robusta	Robusta	Robusta
Robusta (91).jpg	Robusta	Robusta	Robusta	Robusta	Robusta
Robusta (92).jpg	Robusta	Robusta	Robusta	Robusta	Robusta
Robusta (93).jpg	Robusta	Robusta	Robusta	Robusta	Robusta
Robusta (94).jpg	Robusta	Robusta	Robusta	Robusta	Robusta
Robusta (95).jpg	Robusta	Arabika	Arabika	Arabika	Arabika
Robusta (96).jpg	Robusta	Robusta	Robusta	Robusta	Robusta
Robusta (97).jpg	Robusta	Robusta	Robusta	Robusta	Robusta

- a. Perhitungan Confusion Matrix pada sudut 0°

Tabel 4. Prediksi benar dan salah hasil terhadap sudut 0°

Hasil Aktual	Hasil Prediksi sudut 45	
	Arabika	Robusta
Arabika	9	1
Robusta	1	9

- b. Perhitungan Confusion Matrix pada sudut 45°

Tabel 5. Prediksi benar dan salah hasil terhadap sudut 45°

Hasil Aktual	Hasil Prediksi sudut 45	
	Arabika	Robusta
Arabika	10	0
Robusta	1	9

- c. Perhitungan Confusion Matrix pada sudut 90°

Tabel 6. Prediksi benar dan salah hasil terhadap sudut 90°

Hasil Aktual	Hasil Prediksi sudut 90°	
	Arabika	Robusta
	9	1
	1	9

$$\text{Confusion Matrix} = \frac{9 + 9}{20} \times 100\% = 90\%$$

- d. Perhitungan Confusion Matrix pada sudut 135°

Tabel 7. Prediksi benar dan salah hasil terhadap sudut 135°

Hasil Aktual	Hasil Prediksi sudut 90°	
	Arabika	Robusta
Arabika	9	1
Robusta	1	9

$$\text{Confusion Matrix} = \frac{9 + 9}{20} \times 100\% = 90\%$$

Tabel 8. Hasil perhitungan Confusion Matriks dengan nilai K=1

Confusion Matrix			
Sudut 0°	Sudut 45°	Sudut 90°	Sudut 135°
90%	95%	90%	90%

Dari hasil perhitungan dengan menggunakan rumus untuk setiap sudut yang digunakan, didapatkan hasil akurasi sebesar 90% untuk sudut 0°, 95% untuk sudut 45°, 90% untuk sudut 90°, dan 90% untuk sudut 135°. Dari hasil tersebut, sudut 45° memberikan hasil akurasi terbaik.

4. KESIMPULAN

Dari hasil eksperimen yang telah dilakukan, dengan menggunakan sebanyak 177 data citra latih dan 20 data citra uji menghasilkan akurasi sebesar 95% yang didapat dari sudut 45° dengan jarak piksel 1 dan nilai $K=1$. Data benar sejumlah 18 citra dari total keseluruhan 20 citra. Untuk keseluruhan akurasi yang sudah dihasilkan, nilai terendah yaitu 90%. Hal ini membuktikan bahwa menggunakan GLCM-KNN dapat dilakukan proses klasifikasi bijikopi dengan hasil akurasi tinggi.

5. SARAN

Pada penelitian selanjutnya, diharapkan hasil perhitungan dengan fitur ekstraksi lain misalnya menerapkan Linear Binary Pattern (LBP) dengan nilai radius 1,4, dan 8 sehingga dapat menganalisa hasil fitur ekstraksi yang dihasilkan.

DAFTAR PUSTAKA

- P. N. Andono, E. H. Rachmawanto, N. S. Herman, and K. Kondo, "Orchid types classification using supervised learning algorithm based on feature and color extraction," *Bull. Electr. Eng. Informatics*, vol. 10, no. 5, pp. 2530–2538, Oct. 2021.
- A. N. Handayani, H. W. Herwanto, K. L. Chandrika, and K. Arai, "Recognition of Handwritten Javanese Script using Backpropagation with Zoning Feature Extraction," *Knowl. Eng. Data Sci.*, vol. 4, no. 2, p. 117, Dec. 2021.
- M. N. M. Ibrahim and M. Z. M. Yusoff, "Twitter sentiment classification using Naive Bayes based on trainer perception," in *2015 IEEE Conference on e-Learning, e-Management and e-Services (IC3e)*, 2015, pp. 187–189.
- G. Kaushal and R. Bala, "GLCM and KNN based Algorithm for Plant Disease Detection," *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, vol. 6, no. 7, pp. 5845–5852, 2017.
- N. Nafiah, "Klasifikasi Kematangan Buah Mangga Berdasarkan Citra HSV dengan KNN," *J. Elektron. List. dan Teknol. Inf. Terap.*, vol. 1, no. 2, pp. 1–4, 2019.
- J. Li, S. R. Schill, D. E. Knapp, and G. P. Asner, "Object-based mapping of coral reef habitats using planet dove satellites," *Remote Sens.*, vol. 11, no. 12, Jun. 2019.
- A. A. Fauzi, F. Utaminigrum, and F. Ramdani, "Road surface classification based on LBP and GLCM features using KNN classifier," *Bull. Electr. Eng. Informatics*, vol. 9, no. 4, pp. 1446–1453, 2020.
- R. A. Saputra, Suharyanto, S. Wasiyanti, D. F. Saefudin, A. Supriyatna, and A. Wibowo, "Rice Leaf Disease Image Classifications Using KNN Based on GLCM Feature Extraction," *J. Phys. Conf. Ser.*, vol. 1641, no. 1, pp. 1–6, 2020.
- D. P. Pamungkas, "Ekstraksi Citra menggunakan Metode GLCM dan KNN untuk Identifikasi Jenis Anggrek (Orchidaceae)," *Innov. (Innovation Res. Informatics)*, vol. 1, no. 2, pp. 51–56, 2019.
- M. A. Rizal, "Klasifikasi Mutu Biji Kopi Menggunakan Metode K – Nearest Neighbor Berdasarkan Warna Dan Tekstur," *Univ. Teknol. Yogyakarta*, vol. 1, no. 1, pp. 1–8, 2019.
- D. Ikhsan, E. Utami, and F. W. Wibowo, "Metode Klasifikasi Mutu Greenbean Kopi Arabika Lanang Dan Biasa Menggunakan K-Nearest Neighbor Berdasarkan Bentuk," *J. Ilm. SINUS*, vol. 18, no. 2, p. 1, 2020.
- L. B. Rachman and Basari, "Detection of Cholesterol Levels by Analyzing Iris Patterns using Backpropagation Neural Network," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 852, no. 1, p. 012157, Jul. 2020.
- L. B. Rachman and Basari, "Detection of cholesterol levels by analyzing iris patterns using backpropagation neural network," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 852, no. 1, 2020.
- C. Irawan, W. Listyaningsih, D. R. I. M. Setiadi, C. Atika Sari, and E. Hari Rachmawanto, "CBIR for Herbs Root Using Color Histogram and GLCM Based on K-Nearest Neighbor," *Proc. - 2018 Int. Semin. Appl. Technol. Inf. Commun. Creat. Technol. Hum. Life, iSemantic 2018*, no. 3, pp. 509–514, 2018.

- T. Y. Prahudaya and A. Harjoko, "Metode Klasifikasi Mutu Jambu Biji Menggunakan Knn Berdasarkan Fitur Warna Dan Tekstur," *J. Teknosains*, vol. 6, no. 2, p. 113, 2017.
- S. A. Bandy and A. H. Mir, "Statistical textural feature and deformable model based brain tumor segmentation and volume estimation," *Multimed. Tools Appl.*, vol. 76, no. 3, pp. 3809–3828, 2017.
- H. J. Motlak and S. I. Hakeem, "Detection and classification of breast cancer based-on Terahertz imaging technique using artificial neural network & K-nearest neighbor algorithm," *Int. J. Appl. Eng. Res.*, vol. 12, no. 21, pp. 10661–10668, 2017.
- S. Winiarti, F. I. Indikawati, A. Oktaviana, and H. Yuliansyah, "Consumable Fish Classification Using k- Nearest Neighbor," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 821, no. 1, pp. 1–11, 2020.
- I. Nurhaida, H. Wei, R. A. M. Zen, R. Manurung, and A. M. Arymurthy, "Texture fusion for batik motif retrieval system," *Int. J. Electr. Comput. Eng.*, vol. 6, no. 6, pp. 3174–3187, 2016.
- M. Daniel, J. Raharjo, and K. Usman, "Iris-based image processing for cholesterol level detection using gray level co-occurrence matrix and support vector machine," *Eng. J.*, vol. 24, no. 5, pp. 135–144, Sep. 2020.
- P. Moallem, A. Serajoddin, and H. Pourghassem, "Computer vision-based apple grading for golden delicious apples based on surface features," *Inf. Process. Agric.*, vol. 4, no. 1, pp. 33–40, 2017.
- I. U. W. Mulyono et al., "Parijoto Fruits Classification using K-Nearest Neighbor Based on Gray Level Co- Occurrence Matrix Texture Extraction," *J. Phys. Conf. Ser.*, vol. 1501, no. 1, 2020.
- E. Hari Rachmawanto, G. Rambu Anarqi, D. R. I. Moses Setiadi, and C. Atika Sari, "Handwriting Recognition Using Eccentricity and Metric Feature Extraction Based on K-Nearest Neighbors," *Proc. - 2018 Int. Semin. Appl. Technol. Inf. Commun. Creat. Technol. Hum. Life, iSemantic 2018*, pp. 411–416, 2018.
- S. Y. R. Riska and P. Subekti, "Klasifikasi Level Kematangan Buah Tomat Berdasarkan Fitur Warna Menggunakan Multi-Svm," *J. Ilm. Inform.*, vol. 1, no. 1, 2016.