

IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR DALAM KLASIFIKASI CITRA SAYUR

Adli Abdillah Nababan⁽¹⁾, Miftahul Jannah⁽²⁾

^{1,2} Bisnis Digital, STMIK Pelita Nusantara, Medan, Indonesia

e-mail: ¹adliabdillahnababan@gmail.com, ²miftahuljannah0077@gmail.com

ABSTRACT

Classification based on the color of vegetables is not always the only factor used to judge the type of vegetable, but it can be the first clue to visually assess the type of vegetable before deciding to sell or consume it. Farmers can use this technology to better monitor and control vegetation quality using sorting techniques designed to achieve optimal yields and better vegetation quality. This study is important enough to provide information to the public about the evaluation of vegetables based on color. With the system built, Pagar Merbau II Village can be advanced, especially in the utilization of farmers' vegetable crops. The color vegetable scoring system can be used to educate people to increase vegetable yield and create public awareness to consume fresh vegetables based on color. In this study, the technique of classifying vegetable quality based on color used a digital image processing approach to determine the type of vegetable using the K-Nearest Neighbor method to determine the type of vegetable. The results of the classification of vegetable types based on color for 40 samples of image testing data using the KNN method were 87,5%.

Keywords : *K-Nearest Neighbor, Computer Vision, Image Classification, Vegetables*

ABSTRAK

Klasifikasi berdasarkan warna sayuran tidak selalu menjadi satu-satunya faktor yang digunakan untuk menilai jenis sayuran, Tapi hal tersebut bisa menjadi petunjuk pertama untuk menilai jenis sayuran secara visual sebelum memutuskan untuk menjual atau mengkonsumsinya. Petani dapat menggunakan teknologi ini untuk memantau dan mengontrol kualitas vegetasi dengan lebih baik menggunakan teknik pemilahan yang dirancang untuk mencapai hasil optimal dan kualitas vegetasi yang lebih baik. Kajian ini cukup penting untuk memberikan informasi kepada masyarakat tentang penilaian sayuran berdasarkan warna. Dengan sistem yang dibangun, Desa Pagar Merbau II dapat dimajukan terutama dalam pemanfaatan hasil tanaman sayuran para petani. Sistem penilaian warna sayuran dapat digunakan untuk mengedukasi masyarakat untuk meningkatkan hasil sayuran dan menciptakan kesadaran umum untuk mengonsumsi sayuran segar berdasarkan warna. pada Penelitian ini, teknik Klasifikasi kualitas sayur berdasarkan warna menggunakan pendekatan pengolahan citra digital untuk menentukan jenis sayur menggunakan metode K-Nearest Neighbor untuk menentukan jenis sayur. pada penelitian ini jenis sayur yang digunakan adalah Bayam, Kol, Wortel dan Terong. Adapun hasil klasifikasi pada jenis sayuran berdasarkan warna terhadap 40 sampel data pengujian gambar menggunakan metode KNN sebesar 87,5%.

Kata Kunci: K-NearestNeighbor, Computer Vision, Klasifikasi Citra, Sayuran

1. Pendahuluan

Dalam era yang semakin maju ini, teknologi kecerdasan buatan telah menjadi bagian integral dari berbagai bidang kehidupan manusia. Salah satu cabang utama dalam kecerdasan buatan adalah *machine learning*, yang bertujuan untuk mengajarkan komputer untuk belajar dari data dan mengambil keputusan atau mengidentifikasi pola secara otomatis. Dalam konteks pertanian, klasifikasi jenis sayuran menjadi suatu aspek penting yang dapat membantu para petani, ahli pertanian, atau bahkan konsumen untuk mengenali jenis sayuran dengan lebih mudah dan akurat. Identifikasi jenis sayuran berdasarkan karakteristik fisiknya dapat menjadi tantangan, terutama ketika berhadapan dengan variasi bentuk, ukuran, dan warna yang berbeda-beda.

Salah satu metode klasifikasi yang telah terbukti efektif dalam berbagai aplikasi Metode K-Nearest Neighbor (KNN) adalah metode untuk mengklasifikasikan sekumpulan data berdasarkan pada sebagian besar kategori dan tujuannya adalah untuk mengklasifikasikan objek baru berdasarkan atribut dan sampel sampel dari data pelatihan [1]. Sehingga target keluaran yang diinginkan mendekati ketepatan dalam melakukan ujian belajar [2]. Dalam konteks identifikasi jenis sayuran, algoritma KNN dapat digunakan untuk memprediksi jenis sayuran dari contoh data baru berdasarkan kesamaan fitur dengan contoh-contoh data latih yang sudah terlabel.

Pendekatan KNN dalam klasifikasi jenis sayuran memberikan beberapa keunggulan, termasuk kemampuan untuk mengenali jenis sayuran yang belum pernah dilihat sebelumnya, toleransi terhadap data latih yang berubah-ubah, dan kemudahan dalam mengimplementasikan algoritma ini. Melalui penelitian ini,

diharapkan penggunaan algoritma KNN dalam identifikasi jenis sayuran dapat memberikan manfaat signifikan bagi para petani dan ahli pertanian dalam pengenalan jenis sayuran dengan lebih efisien dan akurat. Abdullah et al dalam penelitiannya "*Classification of Fruits Based on Shape and Color using Combined Nearest Mean Classifiers*" mendapatkan hasil bahwa sebesar 100 % [3]. Sedangkan Purwa Hasan et al dalam penelitiannya *Analysis K-Nearest Neighbor Method In Classification Of Vegetable Quality Based On Color* dengan hasil menunjukkan metode K-Nearest Neighbor di klasifikasi data memiliki akurasi persentase yang baik bila menggunakan data acak memiliki persentase 100%. [4].

Penelitian yang dilakukan oleh Didik berjudul Identifikasi Kesegaran Ikan Nila Berdasarkan Warna Insang Menggunakan Algoritma K-Nearest Neighbor (K-NN), dimana hasil penelitian ini digunakan untuk menghitung nilai warna insang ikan nila untuk menentukan kesegarannya. Berdasarkan hasil pengujian dari 30 percobaan dengan rating 96%. [5].

Sebuah studi oleh Mochamad Aryo et al. "Menggunakan pemrosesan citra dan klasifikasi K-Nearest neighbor pada citra telur" dengan akurasi algoritme 100% [6].

Subairi et al. melakukan penelitian "Implementasi Metode K-Nearest-Neighbor Untuk Mengenali Pola Tekstur Gambar Saliva Untuk Deteksi Ovulasi" dengan akurasi metode KNN adalah 93,3%. [7].

Dalam sebuah studi oleh Alifa Puter et al. berjudul "Pengenalan citra jenis bunga menggunakan algoritma KNN dengan ekstraksi warna HSV dan tekstur GLCM dengan akurasi tertinggi 71% menggunakan nilai K-7", pengenalan bunga berhasil dilakukan [8].

Kajian oleh Joan Angelina Widians dkk, "Klasifikasi Jenis Bawang Merah

Menggunakan Metode K-Nearest Neighbor Berdasarkan Ekstraksi Fitur Bentuk dan Tekstur” dengan akurasi 83,56 [9].

2. Metode

Metode penelitian yang dilakukan, terhadap penelitian akan dijelaskan dalam beberapa proses dan data yang akan digunakan adalah citra sayuran yang diuraikan dalam beberapa proses dan data



Gambar 1. Metode Penelitian

A. Citra

Citra adalah gambar yang menyerupai atau meniru suatu objek. Gambar dibagi menjadi dua jenis: gambar analog dan gambar digital. Gambar analog adalah gambar yang tidak dapat diproses oleh komputer. Ada beberapa citra digital yang umum digunakan yaitu citra biner, citra grayscale dan citra berwarna [10]. Pengolahan citra digunakan untuk mengolah informasi yang terkandung dalam citra untuk pengenalan objek secara otomatis. [11] Proses pengolahan citra diawali dengan konversi citra RGB menjadi grayscale, yaitu mengubah citra berwarna menjadi citra grayscale atau

yang digunakan adalah citra sayuran. Adapun tahapan dari penelitian ini dapat dilihat pada gambar berikut :

atau pengecilan dimensi citra menjadi dimensi yang telah ditentukan sehingga tidak ada perbedaan dimensi dari matriks data citra dan memudahkan langkah selanjutnya.

skala abu-abu. [12]. Citra yang digunakan dalam penelitian ini adalah gambar sayuran Bayam, Kol, Wortel dan Terong

B. Resizing

Resizing adalah proses menyamaratakan koordinat pikse sebesar 400 x400 piksel, agar proses komputasi tidak teralu besar. tahap ini dilakukan proses pembesaran

C. Ekstraksi Citra

Ekstraksi citra dilakukan untuk mendapatkan intensitas warna *Red* (R), *Green* (G) dan *Blue* (B) pada citra sayuran, yang selanjutnya akan diubah kedalam warna abu-abu (*Grayscale*). Pada tahap ini berfungsi untuk melakukan penajaman objek tulisan pada citra agar memudahkan sistem untuk digunakan pada proses-proses berikutnya. Proses awal, citra berwarna RGB (Red, Green, Blue) diubah menjadi format citra keabuan (grayscale). untuk Rumus Grayscale [13]:

$$(0.2989 * R) + (0.5870 * G) + (0.1140 * B) \quad (1)$$

D. Klasifikasi dengan KNN

Dalam proses klasifikasi, sebelum melakukan prediksi perlu dilakukan proses pembelajaran terlebih dahulu. Proses pembelajaran membutuhkan data. Data yang dibutuhkan selama proses klasifikasi terdiri dari dua jenis, yaitu:

1. Data latih atau data latih adalah data yang digunakan dalam proses pembelajaran dalam proses klasifikasi.

2. Data uji atau data pengujian adalah data yang digunakan dalam proses prediksi dalam proses klasifikasi[14]. Jarak Euclidean hanyalah jumlah dari perbedaan intensitas piksel dan,

akibatnya, deformasi kecil dapat menghasilkan jarak Euclidean yang besar. Makalah ini mengusulkan jarak Euclidean baru, yang kami sebut *Euclidean Distance Image* (IMED). Tidak seperti yang tradisional, IMED memperhitungkan hubungan spasial piksel. Berdasarkan tiga

3. Hasil dan Pembahasan

Hasil dari penelitian ini diawali dengan data yang digunakan dalam penelitian. Pengumpulan data penelitian tidak boleh sembarangan, tergantung dari teknik pengumpulan data yang telah ditetapkan, ada beberapa tahapan pengumpulan data. Tujuan dari langkah-langkah pengumpulan data survei digunakan untuk proses entri gambar. Di bawah ini adalah contoh input data citra yang telah dilakukan resizing 400 x 400 Piksel :

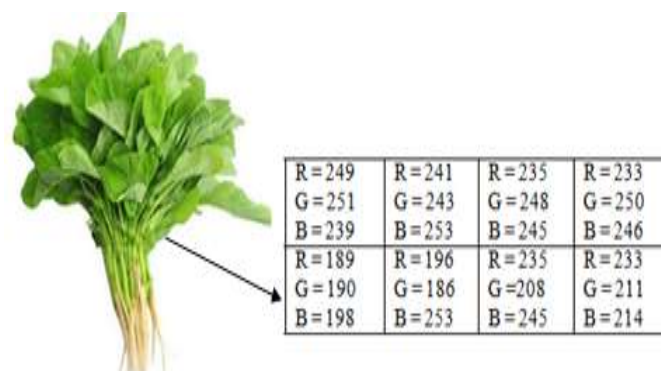


Gambar 2. Sampel data Penelitian

Sehingga jika nilai intensitas warna di konversi dalam bentuk matriks seperti berikut :

properti yang (bisa dibilang) merupakan metrik gambar yang secara intuitif dapat memuaskan, kami menunjukkan bahwa IMED adalah satu-satunya jarak Euclidean yang memiliki properti ini. Jarak Euclidean adalah jarak antara titik-titik dalam garis lurus. Metode jarak ini menggunakan teorema Pythagoras dan merupakan perhitungan jarak yang paling sering digunakan dalam proses pembelajaran mesin. Rumus *Euclidean Distance* adalah hasil dari akar kuadrat dari perbedaan antara dua vektor.

Selanjutnya melakukan proses ekstraksi warna dari RGB ke Grayscale. Dari sebuah citra didapatkan intensitas piksel menggunakan perintah `Imtool` pada Matlab seperti gambar berikut:



Gambar 3. Contoh nilai piksel Bayam 1

	1	2	3	...	400
1	249	244	244	...	220
2	191	197	220	...	209
3	189	198	154	...	207
...
400	249	245	245	...	198

Tabel 1. Matriks Nilai RGB Bayam

	1	2	3	...	400
1	R = 249	R = 241	R = 235	...	R = 229
	G = 251	G = 243	G = 248		G = 212
	B = 239	B = 253	B = 245		B = 237
2	R = 189	R = 196	R = 235	...	R = 231
	G = 190	G = 186	G = 208		G = 210
	B = 198	B = 253	B = 245		B = 146
3	R = 188	R = 211	R = 225	...	R = 211
	G = 187	G = 183	G = 108		G = 220
	B = 198	B = 243	B = 205		B = 127
...
400	R = 254	R = 248	R = 227	...	R = 211
	G = 246	G = 242	G = 254		G = 202
	B = 251	B = 255	B = 243		B = 146

Selanjutnya melakukan konversi RGB ke Grayscale rumus berikut :

$$\text{Gray} = (0.2989 * R) + (0.5870 * G) + (0.1140 * B)$$

Perhitungan yang dilakukan terhadap masing-masing titik piksel adalah sebagai berikut :

$$G_{(1,1)} = (0.2989 * 249) + (0.5870 * 251) + (0.1140 * 239) = 249,009 = 249$$

$$G_{(1,2)} = (0.2989 * 189) + (0.5870 * 190) + (0.1140 * 198) = 190,594 = 191$$

$$G_{(1,3)} = (0.2989 * 188) + (0.5870 * 187) + (0.1140 * 198) = 188,532 = 189$$

$$G_{(2,1)} = (0.2989 * 241) + (0.5870 * 243) + (0.1140 * 253) = 243,518 = 244$$

$$G_{(2,2)} = (0.2989 * 196) + (0.5870 * 186) + (0.1140 * 253) = 196,608 = 197$$

$$G_{(2,3)} = (0.2989 * 211) + (0.5870 * 183) + (0.1140 * 243) = 198,191 = 198$$

$$G_{(3,1)} = (0.2989 * 235) + (0.5870 * 248) + (0.1140 * 245) = 243,748 = 244$$

$$G_{(3,2)} = (0.2989 * 235) + (0.5870 * 208) + (0.1140 * 245) = 220,268 = 220$$

$$G_{(3,3)} = (0.2989 * 225) + (0.5870 * 108) + (0.1140 * 205) = 154,019 = 154$$

Lakukan perhitungan di atas hingga $G_{(400 \times 400)}$ koordinat piksel yang tersedia, sehingga membentuk matriks grayscale baru yang digunakan untuk vektor perbandingan citra pada metode klasifikasi KNN. berikut tampilan hasil Grayscale:

Tabel 2. Matriks Nilai Grayscale Bayam 1

Ulangi langkah perhitungan Grayscale hingga 10 data latih sayur bayam, 10 Data latih Sayur Kol, 10 Data Sayur Wortel dan 10 Data Sayur Terong, sehingga akan didapatkan 40 Tabel grayscale untuk data Training.

Selanjutnya proses klasifikasi menggunakan KNN dengan menentukan $k=5$, misalnya terdapat sembarang nilai matriks grayscale citra uji berikut :

Tabel 3. Matriks Nilai Grayscale Citra Uji

	1	2	3	...	400
1	188	214	214	...	210
2	197	177	190	...	199
3	208	218	194	...	187
...
400	244	245	225	...	238

Untuk menguji nilai grayscale citra uji dikenali sebagai sayur apa, maka akan dilakukan perbandingan jarak piksel menggunakan teori jarak Ecludien Distance antara citra latih bayam 1 dan citra uji seperti berikut :

$$D = \sqrt{(249 - 188)^2 + (244 - 214)^2}$$

$$+ (244 - 214)^2 + (191 - 197)^2 + (197 - 177)^2 + (220 - 190)^2 + (189 - 208)^2 + (198 - 218)^2 + (154 - 195)^2 = 96,010$$

Lakukan perbandingan jarak terhadap 40 data citra latih, hingga mendapatkan tabel berikut :

Tabel 4 Nilai Jarak Gambar data Uji

No	Gambar	Distance
1	Sayur Bayam 1	68,05
2	Sayur Kol 1	79,54
3	Sayur Wortel 1	95,15
4	Sayur Terong 2	92,07
5	Sayur Bayam 2	66,06
6	Sayur Kol 2	67,05
7	Sayur Wortel 2	68,90
8	Sayur Terong 2	55,79
9	Sayur Bayam 3	63,98
10	Sayur Kol 3	64,42
11	Sayur Wortel 3	51,56
12	Sayur Terong 3	77,87
....
40	Sayur Terong 10	73,45

Selanjutnya menentukan nilai K, dalam penelitian nilai $K = 5$, maka akan diambil nilai jarak terkecil sebanyak 5 yang terdapat pada gambar, Sayur Bayam 2, sayur terong 2, Sayur Bayam 3, Sayur Kol 3 dan Sayur Bayam 2.

Setelah didapatkan Nilai jarak terkecil, maka langkah selanjutnya adalah melakukan Vote Majority, dimana citra uji akan di kelompokkan kedalam katagori paling banyak, karena katagori Sayur Bayam paling banyak, oleh karena itu

gambar citra Uji yang baru di kenal sebagai “Sayur Bayam”

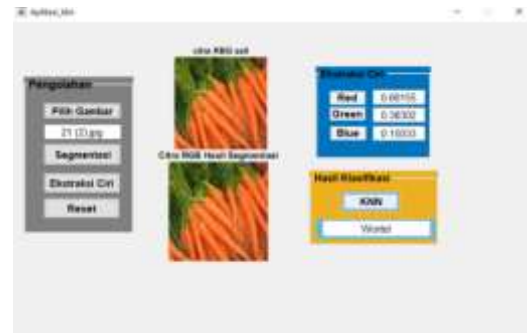
Adapun tampilan dari ap dapat menentukan Sayur yang sesuai dengan katagori warna sayur. berikut adalah tampilan aplikasi yang dibangun.



Gambar 4. Hasil Pengenalan Sayur Bayam



Gambar 5. Hasil Pengenalan Sayur Kol



Gambar 6. Hasil Pengenalan Sayur Wortel



Gambar 7. Hasil Pengenalan Sayur Terong

Berdasarkan pengujian yang dilakukan untuk 30 Data pengenalan gambar, untuk pengenalan gambar jenis sayur memberikan representasi hasil akurasi algoritma KNN dalam mendeteksi Pola adalah

$$\begin{aligned} \text{Persentase} &= \frac{\text{Jumlah True}}{\text{Total Angka}} \\ &= \frac{35}{40} * 100 \% \\ &= 87,5 \% \end{aligned}$$

4. Simpulan dan Saran

Kesimpulan sebagai berikut:

1. Sistem Pengenalan sayur berdasarkan warna memiliki tingkat akurasi Sebesar 87,5 % dalam mengenali 40 data gambar sayur

2. Penelitian menggunakan 80 total sampel data yang digunakan, 40 data untuk proses data training yang akan menjadi vektor acuan dalam proses pembelajaran kedekatan jarak tetangga untuk vektor gambar pengujian pada algoritma KNN dan 40 data untuk proses pengujian jenis sayu

3. Nilai k yang digunakan dalam algoritma KNN sebanyak 5 jarak ketetanggaan

Daftar Pustaka

- [1] Diah, K. T., Faqih, A., & Kusumoputro, B. (2019, November). Exploring the feature selection of the EEG signal time and frequency domain features for k-NN and weighted k-NN. In 2019 IEEE R10 Humanitarian Technology Conference (R10-HTC) (47129) (pp. 196-199). IEEE
- [2] Abu Alfeilat, H. A., Hassanat, A. B., Lasassmeh, O., Tarawneh, A. S., Alhasanat, M. B., Eyal Salman, H. S., & Prasath, V. S. (2019). Effects of distance measure choice on k-nearest neighbor classifier performance: a review. *Big data*, 7(4), 221-2480
- [3] A. Harjoko and O. Mahmod, "JURNAL RESTI Classification of Fruits Based on Shape and Color," vol. 5, no. 158, pp. 51-57, 2023.
- [4] P. H. Putra, M. S. Novelan, and M. Rizki, "Analysis K-Nearest Neighbor Method in Classification of Vegetable Quality Based on Color," *J. Appl. Eng. Technol. Sci.*, vol. 3, no. 2, pp. 126-132, 2022, doi: 10.37385/jaets.v3i2.763.
- [5] D. Remaldhi, D. Wahiddin, and ..., "Identifikasi Kesegaran Ikan Nila Berdasarkan Warna Insang Menggunakan Algoritma K-Nearest Neighbor (K-NN)," ... *Student J. ...*, vol. II, no. 1, pp. 197-202, 2021, [Online]. Available: <http://journal.ubpkarawang.ac.id/mahasiswa/index.php/ssj/article/download/247/175>
- [6] M. A. A. Kurniawan, E. Ermatita, and N. Falih, "Pemanfaatan Pengolahan Citra dan Klasifikasi K-Nearest Neighbor pada Citra Telur Ayam," *Inform. J. Ilmu Komput.*, vol. 16, no. 3, p. 164, 2020, doi: 10.52958/iftk.v16i3.2131.
- [7] Subairi, Rahmadwati, and E. Yudaningsy, "Implementasi Metode k-Nearest Neighbor pada Pengenalan Pola Tekstur Citra Saliva untuk Deteksi Ovulasi," *J. EECCIS*, vol. 12, no. 1, pp. 9-14, 2018.
- [8] A. Salsabila, R. Yunita, and C. Rozikin, "Identifikasi Citra Jenis Bunga menggunakan Algoritma KNN dengan Ekstraksi Warna HSV dan Tekstur GLCM," *Technomedia J.*, vol. 6, no. 1, pp. 124-137, 2021, doi: 10.33050/tmj.v6i1.1667.
- [9] J. A. Widians, H. S. Pakpahan, E. Budiman, H. Haviluddin, and M. Soleha, "Klasifikasi Jenis Bawang Menggunakan Metode K-Nearest Neighbor Berdasarkan Ekstraksi Fitur Bentuk dan Tekstur," *J. Rekayasa Teknol. Inf.*, vol. 3, no. 2, p. 139, 2019, doi: 10.30872/jurti.v3i2.3213.
- [10] Garini, P. A., & Sutarno. (2017). Pengenalan Angka Arab Timur Tulisan Tangan Menggunakan Zone Centroid Zone (ZCZ) dan Backpropagation. *Prosiding Annual Research Seminar 2017 Computer Science and ICT*, 3(1), 1-5.
- [11] Anggelin, A. F., Sanjaya, A., & Setiawan, A. B. (2018). riska Fitria Anggelina Pengenalan Pola Tulisan Huruf Jepang (Hiragana)

- Menggunakan Partisi Citra. *Generation Journal*, 2(1), 25. <https://doi.org/10.29407/gj.v2i1.12057>
- [12] Garini, P. A., & Sutarno. (2017). Pengenalan Angka Arab Timur Tulisan Tangan Menggunakan Zone Centroid Zone (ZCZ
- [13] Garini, P. A., & Sutarno. (2017). Pengenalan Angka Arab Timur Tulisan Tangan Menggunakan Zone Centroid Zone (ZCZ) dan Backpropagation. *Prosiding Annual Research Seminar 2017 Computer Science and ICT*, 3(1), 1–5.
- [14] Anggoro, K., Triyono, J., Raharjo, S., Informatika, S., & Industri, F. T. (2021). *Jurnal SCRIPT Vol . 9 No . 1 Juni 2021*