
KLASIFIKASI CITRA TINGKAT KEMATANGAN BUAH ALPUKAT BERDASARKAN BENTUK WARNA MENGGUNAKAN METODE K-NEAREST NEIGHBO

Angga¹, Ahmad Syarif², Agung Ramadhanu³

Universitas Putra Indonesia “YPTK” Padang

Email: ¹angga19188@gmail.com, ²ahmadsyarif.skom@gmail.com,

³agung_ramadhanu@upiypk.ac.id

Abstract: *The development of information technology today has enabled fruit identification based on color characteristics through digital image processing. This process utilizes a camera to take pictures of the fruit, which are then processed using computer software with digital image processing techniques to determine the level of fruit ripeness efficiently. This study focuses on the use of a color feature extraction method based on RGB mean for the classification of the ripeness level of avocado fruit, which is one of the leading tropical fruits in Indonesia. Avocado, as a fruit that grows well in tropical areas such as Indonesia, requires the determination of the right level of ripeness to maximize shelf life and avoid damage. Therefore, this study applies the K-Nearest Neighbor (KNN) algorithm in the avocado ripeness classification process. The KNN algorithm, known for its ability to identify patterns by comparing feature distances, is used to group data based on color similarities and other characteristics. The purpose of this study is to evaluate the effectiveness of the KNN algorithm in predicting the level of avocado ripeness and to improve prediction accuracy through analysis of the fruit color structure. This research is expected to make a significant contribution to the development of fruit ripeness prediction methods and open up new opportunities in the application of data mining technology in agriculture, especially for avocado harvesting.*

Keyword: *Information technology, fruit identification, color feature extraction, K-Nearest Neighbor algorithm, avocado ripeness.*

Abstrak: Perkembangan teknologi informasi saat ini telah memungkinkan identifikasi buah berdasarkan ciri warna melalui pemrosesan citra digital. Proses ini memanfaatkan kamera untuk mengambil gambar buah, yang kemudian diolah menggunakan perangkat lunak komputer dengan teknik pengolahan citra digital untuk menentukan tingkat kematangan buah secara efisien. Penelitian ini fokus pada penggunaan metode ekstraksi fitur warna berbasis mean RGB untuk klasifikasi tingkat kematangan buah alpukat, yang merupakan salah satu buah tropis unggulan di Indonesia. Alpukat, sebagai buah yang tumbuh subur di daerah tropis seperti Indonesia, memerlukan penentuan tingkat kematangan yang tepat untuk memaksimalkan umur simpan dan menghindari kerusakan. Oleh karena itu, penelitian ini mengaplikasikan algoritma K-Nearest Neighbor (KNN) dalam proses klasifikasi kematangan alpukat. Algoritma KNN, yang dikenal karena kemampuannya dalam mengidentifikasi pola dengan membandingkan jarak fitur, digunakan untuk mengelompokkan data berdasarkan kemiripan warna dan ciri-ciri lainnya. Tujuan dari penelitian ini adalah untuk mengevaluasi efektivitas algoritma KNN dalam memprediksi tingkat kematangan alpukat dan meningkatkan akurasi prediksi melalui analisis struktur warna buah. Penelitian ini diharapkan dapat memberikan kontribusi signifikan terhadap pengembangan metode prediksi kematangan buah serta membuka peluang baru dalam penerapan teknologi data mining dalam bidang pertanian, khususnya untuk panen alpukat.

Kata kunci: Teknologi informasi, identifikasi buah, ekstraksi fitur warna, algoritma K-Nearest Neighbor, kematangan alpukat

PENDAHULUAN

Perkembangan teknologi informasi memungkinkan identifikasi buah berdasarkan ciri warna dengan bantuan komputer. Cara komputasi ini dilakukan dengan pengamatan visual tidak langsung, dengan menggunakan kamera sebagai pengolah citra dari gambar yang kemudian diolah dengan menggunakan perangkat lunak komputer. Teknik pengolahan citra digital digunakan untuk mempermudah dan mempercepat proses pengujian tingkat kematangan. Inputan pada proses klasifikasi adalah karakteristik nilai fitur citra yang didapatkan dari tahap ekstraksi fitur warna menggunakan mean RGB.

Komoditas pertanian sangat mudah berkembang dinegara tropis seperti Indonesia, sehingga produk yang dihasilkan berupa buah-buahan dikenal dengan buah-buahan tropis. Indonesia yang terletak di garis khatulistiwa menyebabkan Indonesia mendapatkan sinar matahari yang dibutuhkan oleh tanaman organik.

Alpukat merupakan tanaman yang dapat tumbuh subur di daerah tropis seperti Indonesia. Buah alpukat merupakan salah satu jenis buah yang digemari banyak orang karena selain rasanya yang enak, buah alpukat juga kaya antioksidan dan zat gizi seperti lemak yaitu 9,8 g/100 g daging buah.

Penentuan tingkat kematangan alpukat merupakan faktor yang penting mengingat tingkat kematangan berpengaruh terhadap umur simpan, dimana semakin tinggi tingkat kematangan alpukat maka umur simpannya akan semakin pendek karena semakin cepat rusak dan membusuk.

Data mining, sebagai ilmu yang menggunakan teknik statistik dan kecerdasan buatan, menawarkan pendekatan potensial untuk mengungkap pola dan hubungan tersembunyi dalam kumpulan data kredit. Salah satunya adalah algoritma K-Nearest Neighbor, yang berfokus pada konsep tetangga terdekat, dalam mengidentifikasi pola

kompleks yang sulit diakses dengan metode analisis konvensional.

Algoritme k-nearest neighbour adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran digambarkan ke ruang berdimensi banyak dengan tiap-tiap dimensi mewakili tiap ciri/fitur dari data. Klasifikasi data baru dilakukan dengan mencari label terdekat. Label terbanyak yang muncul menjadi label data baru. Bila $k=1$, data baru dilabeli dengan label tetangga terdekat. Jarak yang biasa dipakai adalah jarak Euklides. Algoritma KNN Regressor adalah algoritma dengan mengelompokkan data berdasarkan letak ketetanggaannya, pengelompokan data tersebut bergantung pada jumlah nilai k. Nilai RMSE dihasilkan untuk mengukur tingkat model yang sudah dibuat.

Melalui penelitian ini, kami berharap dapat memberikan kontribusi yang signifikan terhadap pengembangan metode prediksi tingkat kematangan buah alpukat berdasarkan bentuk struktur warna yang dihasilkan oleh buah alpukat. Hasil penerapan algoritma K-NN diharapkan dapat meningkatkan akurasi prediksi. Selain itu, penelitian ini juga diharapkan dapat membuka peluang pengembangan lebih lanjut penerapan teknologi data mining dibidang pertanian dan panen buah alpukat yang belum matang berdasarkan bentuk warna.

Data Mining

Data mining adalah metode untuk menganalisis pola dan karakteristik di masa depan serta untuk mengumpulkan informasi tak terduga yang belum pernah terlihat sebelumnya dari database yang besar. Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara berbeda dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. Data mining merupakan bidang dari beberapa

bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistic, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi.

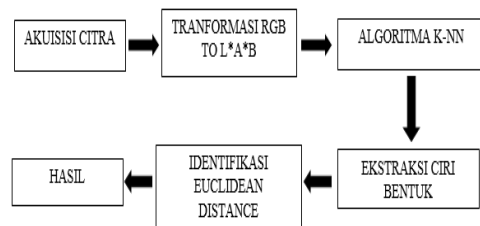
Algoritma K-Nearest Neighbor

Algoritma KNN yakni merupakan sebuah proses supervised yang memiliki arti bahwa pada proses KNN ini memerlukan informasi training guna dilakukannya klasifikasi terhadap suatu objek yang memiliki jarak terdekat.[9]

K-NN adalah teknik untuk mengklasifikasikan dataset yang telah diklasifikasi sebelumnya. Keakuratan algoritma K-NN sangat dipengaruhi oleh fitur yang berbeda ketika nilainya tidak sesuai dengan nilai perkiraan. Beberapa studi yang menggunakan algoritma K-NN menangani hampir secara eksklusif pemilihan fitur dan pembobotan untuk meningkatkan efisiensi algoritma dalam klasifikasi.

METODE

Adapun metodologi penelitian ini terdapat beberapa tahapan yang dilakukan untuk mendapatkan hasil berdasarkan penerapan menggunakan metode K-NN.



Gambar 1. Tahapan Penelitian

Akuisisi Citra

Akuisisi citra, dimana diambil citra tingkat kematangan alpukat yang digunakan sebagai data uji. Proses pengambilan citra melalui pengambilan mandiri secara langsung, dimana citra di foto dengan tingkat cahaya yang sama.

Data dari akuisisi citra yang digunakan adalah 2 jenis alpukat yang sering ditemukan dan mudah didapatkan, diantaranya: alpukat hijau dan alpukat coklat digunakan untuk menentukan tingkat kematangan buah alpukat.

Transformasi $L^*a^*b^*$

Transformasi $L^*a^*b^*$ Seperti hanya L^* , sumbu vertical L^* menyatakan gelap/terang Lightness, yang jangkau ukurannya dari 0 sampai 100. Sumbu yang lain adalah horizontal terdiri atas sumbu a^* dan b^* yang bersilangan tegak lurus satu dengan yang lain. Sumbu a^* dan sumbu b^* berpotongan di titik netral (abu-abu, hitam atau putih). Mereka menunjukkan prinsip bahwa warna merah tidak dapat sekali gus hijau, atau biru sekali gus kuning, mereka berlawanan; kalau dicampur menjadi netral.

Algoritma K-NN

Algoritma KNN yaitu mencari jarak paling terdekat diantara data yang terevaluasi dengan sejumlah K tetangga (*neighbor*) paling dekat dengan data uji. KNN bekerja dengan cara membandingkan data uji dan data training. KNN mencari jarak data training yang paling mendekati dengan data uji. Perhitungan jarak menggunakan rumus Euclidean Distance.

Ekstraksi Ciri Bentuk

Ekstraksi ciri dilakukan pada citra sel alpukat hijau dan alpukat coklat yang telah berhasil disegmentasi. Tujuan dilakukannya proses ekstraksi ciri ini adalah untuk mendapatkan data numerik ciri masing-masing citra dalam parameter ciri citra, yaitu luas area, tepi area, kebundaran bentuk, dan rata-rata warna red, green, dan blue.

Identifikasi Euclidean Distance

Euclidean Distance merupakan teknik yang paling sederhana untuk menghitung jarak di antara dua vector lurus jarak garis antara dua titik, yang memeriksa akar perbedaan kuadrat antara

koordinat sepasang objek. Jarak Euclidean yang paling minimal menunjukkan jarak terdekat untuk proses pengenalnya.[15]

Berikut adalah rumus dari Euclidean Distance

$$P = (P_1, P_2, \dots, P_n)$$

$$Q = (q_1, q_2, \dots, q_n)$$

$$D =$$

$$= \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2}$$

$$= \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

Dimana Padalah citra P, Q adalah citra Q, d adalah ukuran jarak antara query citra P dan Q yang ada di dalam database. Sedangkan p merupakan feature vector pada citra P dan q merupakan feature vector pada citra.

HASIL DAN PEMBAHASAN

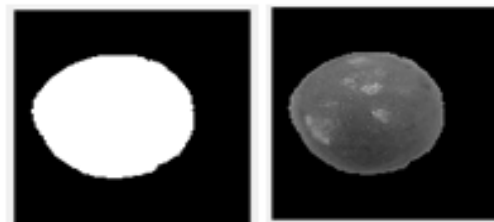
Pada klasifikasi tingkat kematangan buah alpukat berdasarkan bentuk warna, terlebih dahulu dilakukan proses pelatihan. Dataset yang dikumpulkan adalah 40 citra buah alpukat. Citra yang digunakan dalam pelatihan adalah 40 data citra atau 50% dari total citra. Kemudian, 40 data latih. Setelah dataset terkumpul, langkah selanjutnya adalah mempersiapkan pelatihan. Pelatihan dan pengujian dilakukan menggunakan aplikasi Matlab.

Tahap awal dalam dilakukan proses transformasi ruang warna dari citra Red, Green, Blue menjadi citra L*a*b. Hal ini dilakukan agar kandungan warna dapat diidentifikasi secara digital. Proses hasil transformasi dari citra RGB ke L*a*b pada aplikasi Matlab dapat dilihat pada Gambar 2 berikut ini



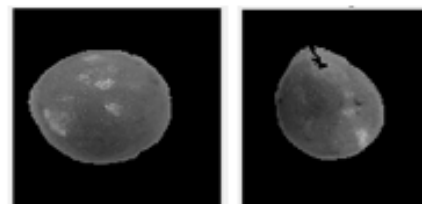
Gambar 2. (a) Citra RGB (b) Hasil Transformasi

Setelah citra di transformasi menjadi L*a*b, agar mudah dalam melakukan segmentasi citra selanjutnya citra diubah atau ditransformasi kedalam bentuk biner. Keluaran dari proses ini adalah citra biner, dimana objek yang diinginkan memiliki nilai 1 (warna putih), sedangkan nilai background adalah 0 (warna hitam). Citra hasil transformasi ke biner dipisahkan objeknya dengan background, sehingga citra yang diperoleh dapat dijadikan sebagai masking untuk melakukan proses.



Gambar 3. (a) Biner, (b) Transformasi Grayscale

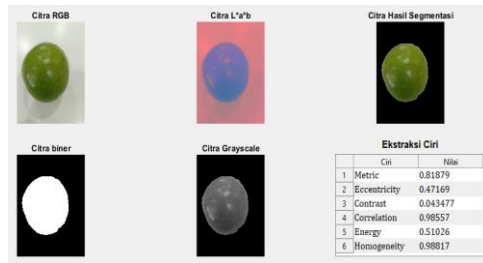
Langkah selanjutnya adalah segmentasi citra menggunakan K-Mean Clustering. Langkah ini dilakukan agar dapat mempartisi data menjadi beberapa region kluster. Hasil segmentasi citra dengan K-Mean Clustering dapat dilihat pada gambar 4 berikut ini.



Gambar 4. Hasil Segmentasi Citra Dengan K-Mean Clustering

Ekstraksi ciri menggunakan ekstraksi ciri bentuk dan tekstur. Fitur yang diekstraksi kemudian digunakan sebagai parameter atau nilai input untuk membedakan objek satu dengan yang lain pada tahap identifikasi. Untuk ekstraksi ciri bentuk, menggunakan parameter metric and eccentricity. Metric didapatkan dari perbandingan antara luas dan keliling suatu benda. Sedangkan eccentricity didapatkan dari perbandingan antara jarak fokus elips minor dengan fokus elips

mayor suatu benda. Untuk ekstraksi ciri tekstur menggunakan Gray-Level Co-Occurrence Matrix (GLCM) dengan parameter contrast, correlation, energy dan homogeneity



Gambar 5.

Hasil Nilai Ekstraksi Ciri Bentuk dan Tekstur

Setelah model diimplementasikan di Matlab, langkah selanjutnya adalah evaluasi. Untuk evaluasi digunakan confusion matrix, dimana perhitungan yang dilakukan meliputi presisi, recall, dan akurasi. Hasil tes kemudian dimasukkan ke dalam confusion matrix. Confusion matrix terdiri dari true positive, false positive, true negative, dan false negative untuk menghitung presisi, recall dan akurasi.

SIMPULAN

Penelitian ini berhasil menunjukkan potensi penggunaan teknologi informasi dan algoritma K-Nearest Neighbor (K-NN) dalam mengklasifikasikan tingkat kematangan buah alpukat berdasarkan ciri warna. Beberapa kesimpulan utama dari penelitian ini adalah sebagai berikut:

1. Efektivitas Pengolahan Citra Digital
Penggunaan teknik pengolahan citra digital untuk identifikasi buah berdasarkan warna terbukti efektif. Transformasi citra dari ruang warna RGB ke Lab memungkinkan analisis warna yang lebih akurat dan mempermudah proses segmentasi citra.
2. Peran Algoritma K-Nearest Neighbor (K-NN)
Algoritma K-NN terbukti menjadi metode klasifikasi yang

berguna dalam menentukan tingkat kematangan alpukat. Dengan menggunakan jarak Euclidean untuk mengukur kedekatan fitur warna dan bentuk, K-NN dapat mengklasifikasikan citra alpukat dengan cukup akurat berdasarkan data pelatihan yang tersedia.

3. Evaluasi model menggunakan confusion matrix menunjukkan bahwa metode yang diterapkan memberikan hasil yang memuaskan dalam hal presisi, recall, dan akurasi. Hal ini menunjukkan bahwa algoritma K-NN dapat diandalkan dalam aplikasi praktis untuk menentukan tingkat kematangan buah alpukat berdasarkan ciri warna.
4. Penelitian ini memberikan kontribusi signifikan dalam pengembangan metode prediksi tingkat kematangan buah alpukat. Hasilnya membuka peluang untuk pengembangan lebih lanjut dalam penerapan teknologi data mining di bidang pertanian, khususnya untuk meningkatkan manajemen panen dan kualitas produk.

DAFTAR PUSTAKA

- J. Saputra, Y. Sa, V. Yoga Pudya Ardhana, and M. Afriansyah, "RESOLUSI: Rekayasa Teknik Informatika dan Informasi Klasifikasi Kematangan Buah Alpukat Mentega Menggunakan Metode K-Nearest Neighbor Berdasarkan Warna Kulit Buah," *Media Online*, vol. 3, no. 5, pp. 347–354, 2023, [Online]. Available: <https://djournals.com/resolusi>
- S. Siti Samsiyah, S.E., M.M. Nashrudin Latif, C. D. Aprilia, A. R. Auliansyah, and N. T. A. S. Gita Raditya, Maretha Zukar, "Perilaku Konsumen," *Mega Press Nusant.*, vol. 23, no. 3, pp. 14–20, 2023.
- L. Malangngi, M. Sangi, and J. Paendong, "Penentuan Kandungan Tanin dan Uji Aktivitas Antioksidan Ekstrak

- Biji Buah Alpukat (*Persea americana* Mill.),” *J. MIPA*, vol. 1, no. 1, p. 5, 2012, doi: 10.35799/jm.1.1.2012.423.
- F. Aprilliani, D. Atmiasih, and A. Ristono, “THE EVALUATION OF AVOCADO (*Persea americana* Mill.) MATURITY LEVEL USING IMAGE PROCESSING TECHNOLOGY,” *J. Penelit. Pascapanen Pertan.*, vol. 18, no. 1, p. 1, 2021, doi: 10.21082/jpasca.v18n1.2021.1-8.
- N. Windy Mardiyah, N. Rahaningsih, and I. Ali, “Penerapan Data Mining Menggunakan Algoritma K-Nearest Neighbor Pada Prediksi Pemberian Kredit Di Sektor Finansial,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 1491–1499, 2024, doi: 10.36040/jati.v8i2.9010.
- R. Rahmadini, Enjel Erika LorencisLubis, Aji Priansyah, Yolanda R.W.N, and Tuti Meutia, “Penerapan Data Mining Untuk Memprediksi Harga Bahan Pangan Di Indonesia Menggunakan Algoritma K-Nearest Neighbor,” *J. Mhs. Akunt. Samudra*, vol. 4, no. 4, pp. 223–235, 2023, doi: 10.33059/jmas.v4i4.7074.
- Sekar Setyaningtyas, B. Indarmawan Nugroho, and Z. Arif, “Tinjauan Pustaka Sistematis: Penerapan Data Mining Teknik Clustering Algoritma K-Means,” *J. Teknoif Tek. Inform. Inst. Teknol. Padang*, vol. 10, no. 2, pp. 52–61, 2022, doi: 10.21063/jtif.2022.v10.2.52-61.
- D. P. Utomo and M. Mesran, “Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung,” *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- Q. A. A’yuniyah and M. Reza, “Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Jurusan Siswa Di Sma Negeri 15 Pekanbaru,” *Indones. J. Inform. Res. Softw. Eng.*, vol. 3, no. 1, pp. 39–45, 2023, doi: 10.57152/ijirse.v3i1.484.
- F. Putra, H. F. Tahiyat, R. M. Ihsan, R. Rahmadden, and L. Efrizoni, “Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 1, pp. 273–281, 2024, doi: 10.57152/malcom.v4i1.1085.
- D. Nurnaningsih, D. Alamsyah, A. Herdiansah, A. Aristo, and J. Sinlae, “Identifikasi Citra Tanaman Obat Jenis Rimpang dengan Euclidean Distance Berdasarkan Ciri Bentuk dan Tekstur,” vol. 3, no. 3, pp. 171–178, 2021, doi: 10.47065/bits.v3i3.1019.
- [M. Natsir, “Penajaman Citra Penginderaan Jauh Komposit Warna Dengan Pemanfaatan Citra Sparsa Transformasi Curvelet,” *Pros. Pertem. Ilm. Tah.*, pp. 507–518, 2015.
- M. Yunus and N. K. A. Pratiwi, “Prediksi Status Gizi Balita Dengan Algoritma K-Nearest Neighbor (KNN) di Puskesmas Cakranegara,” *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 4, pp. 221–231, 2023, doi: 10.35746/jtim.v4i4.328.
- M. F. Grace and H. S. Scott, “An optional federal charter for insurance: Rationale and design,” *Futur. Insur. Regul. United States*, vol. 6, no. 2, pp. 55–96, 2009.
- D. Riana *et al.*, “Identifikasi Citra Pap Smear RepoMedUNM dengan Menggunakan K-Means Clustering dan GLCM,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 1–8, 2022, doi: 10.29207/resti.v6i1.3495.