

K-NN Klasifikasi Kematangan Buah Mangga Manalagi Menggunakan L*A*B dan Fitur Statistik

Arif Patriot Sri Pamungkas¹, Nur Nafi'iyah², Nur Qomariyah Nawafilah³

¹ Mahasiswa Program Teknik Informatika, Fakultas Teknik,
Universitas Islam Lamongan
Jln. Veteran Nomor 53A Lamongan
arifpatriot13@gmail.com

^{2,3} Program Studi Teknik Informatika, Fakultas Teknik,
Universitas Islam Lamongan
Jln. Veteran Nomor 53A Lamongan
mynaff26@gmail.com, nq.nawafil@yahoo.com

Abstrak

*Buah-buahan merupakan bahan pangan sumber vitamin. Buah cepat sekali rusak oleh pengaruh mekanik, kimia dan mikrobiologi sehingga mudah menjadi busuk. Klasifikasi dilakukan pada sekelompok buah mangga yang berbeda-beda jenis kematangannya. Ciri pembeda yang digunakan adalah fitur warna L*A*B. Tujuan penelitian ini yaitu memberikan hasil output klasifikasi kematangan buah mangga manalagi berdasarkan fitur warna menggunakan aplikasi Matlab. Pada penelitian ini akan diusulkan metode GLCM untuk ekstraksi fitur pada buah mangga. Dengan menggunakan K-Nearest Neighbor (KNN) untuk menentukan tingkat kematangan buah mangga. Dataset yang digunakan berjumlah 130 data, terdiri dari 65 data untuk mentah, 15 untuk setengah matang dan 50 untuk matang. Hasil Klasifikasi KNN dengan menggunakan metode GLCM dan L*A*B untuk ekstraksi fitur mendapatkan nilai akurasi sebesar 62.5% pada data uji.*

Kata kunci— Matlab, Mangga Manalagi, KNN, Lab, GLCM.

Abstract

*Fruits are a food source of vitamins. The fruit is quickly damaged by mechanical, chemical and microbiological influences, making it easy to rot. Classification is carried out on a group of mangoes which differ in type of maturity. The distinguishing feature was used is the L*A*B color feature. The purpose of this research gave the output of the maturity classification of Manalagi mangoes based on color features using the Matlab application. In this research the GLCM method will be proposed for feature extraction in mangoes. By using K-Nearest Neighbor (KNN) to determine the maturity level of the Mango fruit. The dataset used is 130 data, consisting of 65 data for raw, 15 for half-cooked and 50 for mature. The KNN Classification results using the GLCM and L*A*B methods for Feature Extraction get an accuracy value of 62.5% in the test data.*

Keywords— Matlab, Manalagi Mango, KNN, Lab, GLCM.

1. PENDAHULUAN

Langkah awal dalam penelitian Arissa Aprilia Nurcahyani, Ristu Saptono, (2015) dilakukan training terhadap 30 data untuk mendapatkan decision tree dengan model ID3 (Iterative Dichotomiser Tree). Data yang telah diakuisisi kemudian diklasifikasi ke dalam 3 kelas yaitu baik, kurang dan buruk dengan menggunakan aturan dari decision tree yang dihasilkan pada proses training. Hasil pengujian dengan metode k-fold cross validation dengan k=5 didapatkan akurasi sebesar 96.67%.

Berdasarkan hasil penelitian Yuda Permadi, Murinto, (2015) dengan menggunakan 20 sampel yang terdiri dari 10 citra mentimun matang dan 10 citra mentimun belum matang menunjukkan bahwa hasil untuk pengujian mentimun matang mencapai 70%, sedangkan untuk mentimun belum matang mencapai 80%. Secara keseluruhan tingkat keberhasilan aplikasi pengolahan citra untuk identifikasi kematangan mentimun berdasarkan tekstur kulit buah dengan metode ekstraksi ciri statistik, yaitu sebesar 75%.

Berdasarkan evaluasi uji coba yang telah dilakukan, dapat disimpulkan bahwa metode yang diajukan dalam mengidentifikasi tiga penyakit noda pada citra daun tebu dapat menunjukkan performa paling baik dengan akurasi tertinggi sebesar 93% dibandingkan dengan kombinasi fitur lainnya yang telah dilakukan pada percobaan. Informasi warna dan tekstur berpengaruh terhadap pengenalan jenis penyakit noda pada citra daun tebu, di mana terdapat kesalahan klasifikasi yang memiliki tekstur maupun warna yang mirip. Selain itu segmentasi yang baik diperlukan untuk mendapat informasi yang akurat dari lesi suatu penyakit noda. Segmentasi tersebut berpengaruh pada pengenalan citra, di mana terkadang terdapat bercak yang bukan merupakan lesi penyakit noda sehingga dapat menimbulkan kesalahan klasifikasi yang tidak sesuai dengan label ground truth (Evy Kamilah Ratnasari, Raden Venantius Hari Ginardi, Chastine Fatichah, 2017).

Penelitian Evy Kamilah Ratnasari, Raden Venantius Hari Ginardi, Chastine Fatichah, (2014) mengusulkan pengenalan penyakit noda tanaman tebu yang terdiri dari noda cincin, noda karat, dan noda kuning berdasarkan fitur tekstur yang merupakan kombinasi dari konsep Gray Level Co-Occurrence Matrix (GLCM) dan dimensi fraktal yang dinamakan Fractal Dimension Co-Occurrence Matrix (FDCM). Sedangkan fitur warna didapatkan dari perhitungan statistik color moments pada citra $L^*a^*b^*$. Kombinasi fitur tersebut menghasilkan 12 fitur warna dan 6 fitur tekstur yang kemudian digunakan sebagai masukan klasifikasi k-Nearest Neighbor (KNN). Pengenalan penyakit noda pada tanaman tebu menggunakan metode tersebut dapat menghasilkan akurasi tertinggi 90%.

Dalam kasus kematangan buah mangga, terkadang ada buah mangga yang memiliki warna yang cukup matang tetapi masih terasa asam, begitu pun sebaliknya. Sehingga untuk para konsumen diperlukan sebuah alat bantu untuk mengetahui tingkat kematangan dari buah Mangga. Metode sebelumnya telah berhasil mengklasifikasikan tingkat kematangan buah Mangga Harum Manis berdasarkan normalisasi warna dengan menggunakan metode Fuzzy Logic, namun kelemahan dari penelitian ini adalah metode yang digunakan hanya dapat digunakan untuk satu jenis buah mangga. Pada penelitian Cahya Bagus Sanjaya, Muhammad Imron Rosadi, (2018) akan diusulkan metode GLCM digabungkan dengan K-Means Clustering untuk ekstraksi fitur pada buah

mangga sehingga nantinya dapat digunakan pada berbagai jenis Mangga. Dengan menggunakan LS-SVM untuk menentukan tingkat kematangan buah Mangga. Dataset yang digunakan berjumlah 240 data, terdiri dari masing-masing 80 data untuk mangga jenis harum manis, mangga jenis manalagi, dan mangga jenis kent, masing-masing jenis terdiri dari 40 data mangga mentah dan 40 data data matang. Hasil Klasifikasi LS-SVM dengan menggunakan metode GLCM dan K-Means Clustering untuk Ekstraksi Fitur mendapatkan nilai akurasi sebesar 98,33% pada data uji (Cahaya Bagus Sanjaya, Muhammad Imron Rosadi, 2018).

Dalam menentukan kematangan mangga yang baik, manusia bersifat lebih subjektif dan tidak konsisten sehingga dari satu penilai dengan penilai lainnya hasilnya akan berbeda. Dengan adanya perkembangan teknologi informasi, memungkinkan klasifikasi kematangan buah pada mangga dengan bantuan komputer. Klasifikasi kematangan buah mangga dapat diketahui dengan teknik pencarian gambar menggunakan kemiripan karakteristik dengan menggunakan fitur statistik. Ciri merupakan suatu tanda yang khas pada buah mangga, yang membedakan antara satu gambar dengan gambar yang lain. Pada dasarnya suatu gambar memiliki ciri-ciri dasar yaitu: Warna, Bentuk, Tekstur. Oleh sebab itu, penelitian ini menggunakan ciri warna L^*A^*B untuk pengenalan warna dari buah mangga manalagi karena L^*A^*B merupakan ruang warna yang representative dalam menginterpretasikan permukaan citra buah atau sayuran.

Dalam penelitian menentukan ekstraksi ciri dilakukan untuk mendapatkan pola dari suatu citra yang akan dilatih maupun citra yang akan diuji. Metode ekstraksi ciri yang digunakan adalah ekstraksi ciri statistik. Metode ini menggunakan perhitungan statistik distribusi derajat keabuan (histogram) dengan mengukur tingkat kekontrasan, granularitas, dan kekasaran suatu daerah dari hubungan ketetanggaan antar piksel di dalam citra.. Dalam menentukan kategori suatu citra atau gambar dapat dilakukan dengan menggunakan teknik pengukuran jarak yaitu Euclidean Distance. Euclidean Distance merupakan suatu teknik yang digunakan untuk menghitung kesamaan antara dua vektor dengan cara menghitung akar dari kuadrat perbedaan dari dua vektor tersebut. Tujuan penelitian ini, yaitu: Membuat sistem klasifikasi kematangan buah mangga manalagi berdasarkan fitur GLCM, Lab.

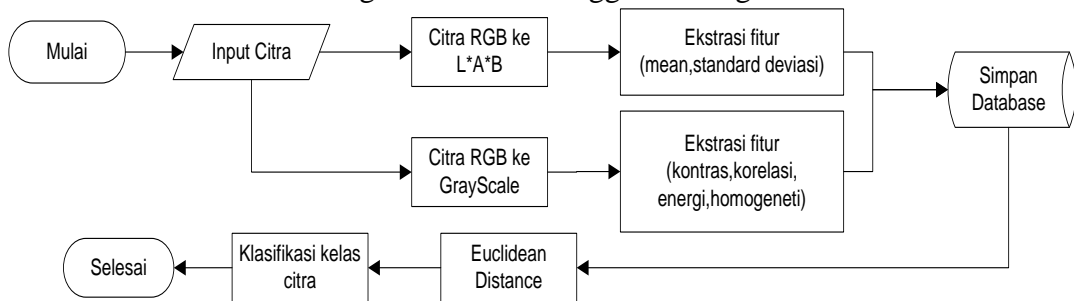
2. METODE PENELITIAN

Pada penelitian ini metodologi yang pertama dilakukan yaitu metode mulai dari tahapan pengumpulan data, perancangan sistem, perancangan proses dari awal hingga akhir. Adapun tahapan metode penelitian yang digunakan dalam penelitian yang berjudul “Klasifikasi Kematangan Buah Mangga Manalagi Menggunakan L^*A^*B dan Fitur Statistik” adalah seperti pada Gambar 1.

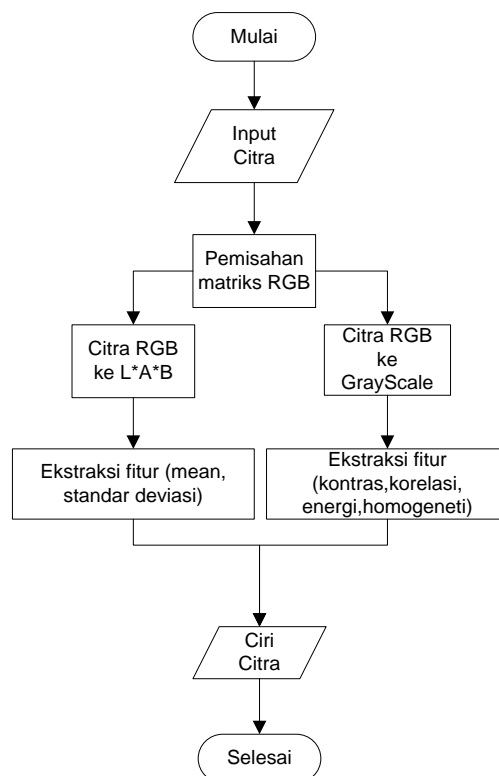
Tahapan metode penelitian:

1. Input gambar asli citra buah mangga manalagi dengan format JPG atau PNG.
2. Setelah itu gambar asli citra diubah menjadi citra warna L^*a^*b dan *grayscale* untuk diambil nilai piksel yang ada pada gambar citra.
3. Kemudian dilakukan pengambilan nilai ekstraksi ciri dari gambar yang sudah diubah ke L^*a^*b dan *grayscale* dengan menggunakan aplikasi Matlab.

4. Maka akan muncul nilai ekstraksi ciri rata-rata, standar deviasi, kontras, energi, korelasi dan homogeneti dari data observasi yang ada.
5. Simpan hasil nilai matrik fitur rata-rata, standar deviasi, kontras, energi, korelasi dan homogeneti. Untuk proses *data training* dilakukan sampai tahapan ini saja, selanjutnya hasil nilai matrik fitur rata-rata, standar deviasi, kontras, energi, korelasi dan homogeneti akan dipakai untuk mengklasifikasikan citra gambar pada *data testing*.
6. Pada tahap ini dilakukan pada data testing sejumlah 40 data citra gambar yakni pengklasifikasian citra gambar mangga manalagi matang, setengah matang dan mentah dengan rumus jarak *eclidean distance*. Yang mana perhitungannya juga mengacu pada nilai matrik fitur rata-rata, standar deviasi, kontras, energi, korelasi dan homogeneti.
7. Hasil dari proses klasifikasi metode *k-nearest neighbor* yakni data gambar yang dimasukkan adalah citra gambar buah mangga manalagi.



Gambar 1. Tahapan Metode Penelitian

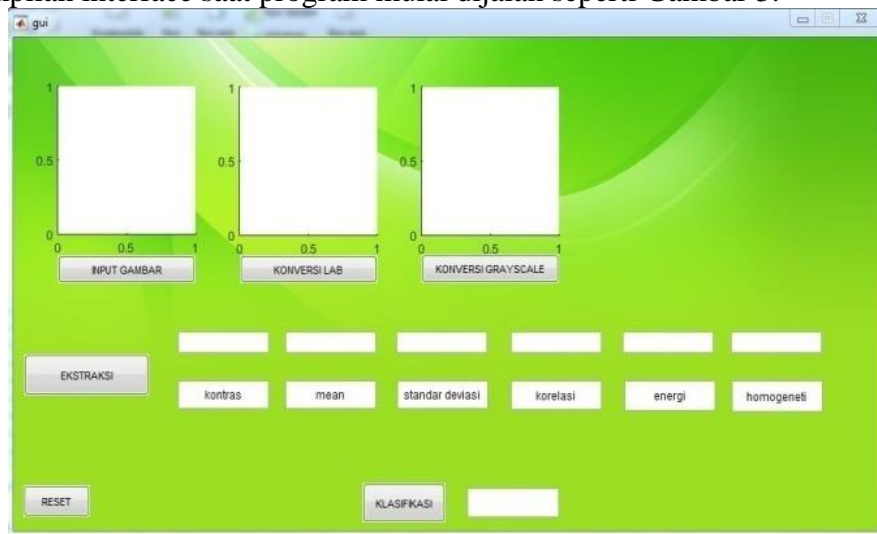


Gambar 2. Ekstraksi Ciri

Alur dari gambar 2 diawali dengan memasukan citra buah mangga manalagi sebanyak 90 sample citra. Citra tersebut berformat jpg. Selanjutnya, dilakukan proses perubahan fitur warna $l*a*b$ dan *grayscale*. Kemudian dilakukan proses ekstraksi fitur menggunakan rata-rata, standar deviasi, kontras, korelasi, energi dan homogeneti. Hasilnya didapatkan nilai ciri hasil rata-rata, standar deviasi, kontras, korelasi, energi dan homogeneti, pada tahapan testing, proses yang dilakukan adalah sama.

3. HASIL DAN PEMBAHASAN

Tampilan interface saat program mulai dijalankan seperti Gambar 3.



Gambar 3. Tampilan Program

Saat kita klik tombol Input Gambar maka kita bisa memilih gambar yang akan diproses oleh sistem. Citra gambar yang akan diproses harus berada dalam folder Matlab. Fungsi tombol Konversi LAB yakni merubah gambar asli menjadi gambar LAB untuk diambil nilai rata-rata dan standar deviasi kemudian menampilkannya. Fungsi tombol Konversi Grayscale yakni merubah gambar asli menjadi gambar *grayscale* untuk diambil nilai kontras, korelasi, energi dan homogeneti kemudian menampilkannya. Tombol Ekstraksi Ciri digunakan untuk menghitung nilai kontras, korelasi, energi, homogeneti, rata-rata dan standar deviasi dari gambar asli yang sudah di ubah ke LAB dan *grayscale* kemudian menampilkan hasilnya. Tombol Klasifikasi digunakan untuk mengelompokkan *data testing* dengan *data training* termasuk kelompok buah matang, setengah matang dan mentah menggunakan rumus jarak *euclidean distance*. Tombol RESET digunakan untuk me-*reset* atau menghapus data yang sudah diproses mulai dari proses pengambilan gambar asli sampai dengan proses *clustering* sehingga kembali seperti awal.

Fitur pertama yang dihitung adalah rerata intensitas atau biasa disebut *mean*, contoh perhitungannya sebagai berikut:

$$m = \sum_{i=0}^{L-1} i \cdot p(i)$$

$$m = \frac{0 \times 0}{400 \times 400} + \frac{1 \times 0}{400 \times 400} + \frac{2 \times 0}{400 \times 400} + \frac{3 \times 0}{400 \times 400} + \frac{4 \times 0}{400 \times 400} + \frac{5 \times 0}{400 \times 400} +$$

$$\frac{6 \times 0}{400 \times 400} + \frac{7 \times 0}{400 \times 400} + \frac{8 \times 0}{400 \times 400} + \frac{9 \times 0}{400 \times 400} + \frac{10 \times 0}{400 \times 400} + \dots +$$

$$\frac{255 \times 0}{400 \times 400} +$$

$m=41,8881$. Dari perhitungan di atas didapatkan nilai m di mana m merupakan nilai rerata intensitas. Hasil tersebut didapatkan dari hasil perkalian i dengan $p(i)$. Di mana i merupakan nilai aras keabuan suatu citra dan $p(i)$ merupakan probabilitas kemunculan nilai i .

$$\Sigma = \sqrt{\sum_{i=0}^{L-1} (i - m)^2 \cdot p(i)}$$

$$\Sigma = \sqrt{(0 - 41,8881)^2 \times \frac{0}{400 \times 400}} + \sqrt{(1 - 41,8881)^2 \times \frac{0}{400 \times 400}} +$$

$$\sqrt{(2 - 41,8881)^2 \times \frac{0}{400 \times 400}} + \sqrt{(3 - 41,8881)^2 \times \frac{0}{400 \times 400}} +$$

$$\sqrt{(4 - 41,8881)^2 \times \frac{0}{400 \times 400}} + \sqrt{(5 - 41,8881)^2 \times \frac{0}{400 \times 400}} +$$

$$\sqrt{(6 - 41,8881)^2 \times \frac{0}{400 \times 400}} + \sqrt{(7 - 41,8881)^2 \times \frac{0}{400 \times 400}} +$$

$$\sqrt{(8 - 41,8881)^2 \times \frac{0}{400 \times 400}} + \sqrt{(9 - 41,8881)^2 \times \frac{0}{400 \times 400}} +$$

$$\dots + \sqrt{(255 - 41,8881)^2 \times \frac{0}{400 \times 400}} +$$

$std=27,1680$ Dari perhitungan di atas diperoleh hasil *deviasi*. Fitur *deviasi* ini menunjukkan kekontrasan suatu citra. Hasil *deviasi* didapatkan dari pengurangan i dengan m dikuadratkan dan dikalikan $p(i)$ setelah itu hasil baru diakar. Di mana i merupakan nilai aras keabuan suatu citra $p(i)$ merupakan probabilitas kemunculan i dan m merupakan rerata intensitas.

Pembahasan hasil penelitian yang telah dilakukan dalam pengelompokan citra gambar buah mangga manalagi matang, setengah matang dan mentah dengan menggunakan K-Nearest Neighbor ini. Dari 130 data citra gambar yang terbagi menjadi 90 *data training* dan 40 *data testing*, kemudian 90 *data training* terdiri dari 40 citra gambar mentah, 10 citra gambar setengah matang dan 40 citra gambar matang sedangkan untuk 40 citra *data testing* terdiri dari 25 citra gambar buah mangga mentah, 5 citra gambar setengah matang dan 10 citra gambar matang. Peneliti menemukan beberapa kendala terkait dengan pembacaan nilai ciri gambar sehingga nilai kontras, korelasi, energi, homogeniti, rata-rata dan standar deviasi dari masing-masing citra gambar mentah, setengah matang dan matang memiliki nilai kontras, korelasi, energi, homogeniti, rata-rata dan standar deviasi yang hampir mirip sehingga sistem kesulitan dalam mengklasifikasi. Seperti pada saat melakukan proses klasifikasi atau pengelompokan pada citra gambar data training ada beberapa citra gambar yang seharusnya terklasifikasi citra buah mentah tapi menjadi buah matang. Begitu pula saat melakukan proses klasifikasi citra *data testing* ada citra gambar yang terklasifikasi menjadi buah mentah padahal seharusnya citra gambar adalah buah setengah matang. Tentunya hal ini membuat peneliti melakukan beberapa kali uji coba dan berusaha

memperbaiki sistem yang ada, dikarenakan kemungkinan ada *human error* yang terjadi. Setelah dilakukan uji coba dan perbaikan sistem peneliti menemukan kesimpulan yang salah satunya adalah diakibatkan karena citra gambar *training* maupun *testing* yang kurang baik, sehingga sistem kurang bisa membaca nilai piksel dengan benar.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan oleh penulis, maka dapat diambil beberapa kesimpulan diantaranya adalah sebagai berikut:

1. Penerapan klasifikasi menggunakan L*A*B dan Fitur Statistik dapat diterapkan dengan baik dalam sistem klasifikasi kematangan buah mangga manalagi berdasarkan warna.
2. Ekstraksi ciri pada sistem klasifikasi kematangan buah mangga manalagi memperoleh nilai yang diambil yaitu nilai kontras, rata-rata, standar deviasi, korelasi, energi dan homogeniti.
3. Akurasi yang didapatkan dari pengujian data testing sebesar 62.5%.

5. SARAN

Agar penelitian ini memberikan hasil yang lebih baik, maka perlu adanya proses klasifikasi dengan metode yang berbeda untuk pengambilan nilai dari objek. Memperbanyak ekstraksi ciri yang diambil supaya akan memberikan hasil klasifikasi yang lebih baik. Memperbanyak data training sebagai acuan jarak terdekat

UCAPAN TERIMA KASIH

Terima kasih banyak kepada orang tua yang memberikan dukungan serta almamater prodi Teknik Informatika Universitas Islam Lamongan.

DAFTAR PUSTAKA

- Arissa Aprilia Nurcahyani, Ristu Saptono. (2015). Identifikasi Kualitas Beras dengan Citra Digital. *Scientific Journal of Informatics*, 63-72.
- Cahya Bagus Sanjaya, Muhammad Imron Rosadi. (2018). Klasifikasi Buah Mangga Berdasarkan Tingkat Kematangan Menggunakan Least-Squares Support vector Machine. *Explore IT*, 1-13.
- Evy Kamilah Ratnasari, Raden Venantius Hari Ginardi, Chastine Faticah. (2014). Pengenalan Penyakit Noda pada Citra Daun Tebu Berdasarkan Ciri Tekstur Fractal Dimension Co-occurrence Matrix dan Lab Moments. *Juti*, 27-36.
- Evy Kamilah Ratnasari, Raden Venantius Hari Ginardi, Chastine Faticah. (2017). Klasifikasi Penyakit Noda pada Citra Daun Tebu Berdasarkan Ciri Tekstur dan Warna Menggunakan Segmentation-Based Gray Level Cooccurrence Matrix dan Lab Color Moments. *Register*, 1-7.

Yuda Permadi, Murinto. (2015). Aplikasi Pengolahan Citra untuk Identifikasi Kematangan Mentimun Berdasarkan Tekstur Kulit Buah Menggunakan Metode Ekstraksi Ciri Statistik. *Jurnal Informatika*, 1028-1038.