

KLASIFIKASI BUAH MANGGA BERDASARKAN TINGKAT KEMATANGAN MENGGUNAKAN *LEAST-SQUARES SUPPORT VECTOR MACHINE*

Cahya Bagus Sanjaya¹⁾, dan Muhammad Imron Rosadi²⁾

^{1, 2)}Teknik Informatika, Universitas Yudharta Pasuruan
Pasuruan

e-mail: cahya.bagus@yudharta.ac.id¹⁾, imron.rosadi@yudharta.ac.id²⁾

Received : September, 2018

Accepted : Oktober, 2018

Published : Desember, 2018

ABSTRAK

Buah mangga adalah buah musiman yang sudah umum diketahui oleh masyarakat di Indonesia. Buah yang berasal dari India ini mempunyai banyak jenis yang banyak digemari oleh masyarakat Indonesia. Tanaman Mangga dapat tumbuh dengan baik di dataran rendah dan berhawa panas. Akan tetapi, ada juga yang dapat tumbuh di daerah yang memiliki ketinggian hingga 600 meter di atas permukaan laut. Batang pohon mangga tegak, bercabang agak kuat. Kulit tebal dan kasar dengan banyak celah-celah kecil dan sisik-sisik bekas tangkai daun.

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 ini 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.

Kata Kunci: *Mangga, GLCM, K-Means Clustering, LS-SVM*

I. PENDAHULUAN

Buah Mangga adalah buah musiman yang sudah umum diketahui oleh masyarakat di Indonesia. Buah yang berasal dari India ini mempunyai banyak jenis yang banyak digemari oleh masyarakat Indonesia. Tanaman Mangga dapat tumbuh dengan baik di dataran rendah dan berhawa panas. Akan tetapi, ada juga yang dapat tumbuh di daerah yang memiliki ketinggian hingga 600 meter di atas permukaan laut. Batang pohon mangga tegak, bercabang agak kuat. Kulit tebal dan kasar dengan banyak celah-celah kecil dan sisik-sisik bekas tangkai daun (Nurdayati dkk, 2016).

Salah satu penghasil buah mangga terbesar di Pulau Jawa terdapat di Kabupaten Probolinggo, Situbondo dan Pasuruan Jawa Timur. Menurut Kementerian Pertanian (2015) produksi buah mangga di Indonesia pada tahun 2012 sampai 2014 mengalami peningkatan, yaitu pada tahun 2012 produksi buah mangga sebesar 76.547 ton, pada tahun 2013 produksi mangga sebesar 98.958 ton, dan pada tahun 2014 produksi mangga sebesar 102.820 ton (Qodri, 2016).

Dalam kasus kematangan buah mangga ini terkadang ada buah mangga yang memiliki warna yang cukup matang tetapi masih terasa asam, begitu pun sebaliknya (Nurdayati dkk, 2015). Sehingga untuk para konsumen diperlukan sebuah alat bantu untuk mengetahui tingkat kematangan dari buah Mangga.

Mulato (2016) telah melakukan penelitian untuk mengetahui tingkat kematangan pada buah Jambu Biji Merah dengan menggunakan metode Fuzzy. Metode yang diusulkan oleh mulato adalah dengan mengubah tipe gambar jambu biji merah dari tipe red, green, dan blue (RGB) ke tipe grayscale yang kemudian diekstraksi menggunakan bantuan aplikasi MATLAB untuk memperoleh informasi gambar. Terdapat 11 informasi yang digunakan sebagai masukan model fuzzy,

hasil dari proses ini adalah tingkat kematangan dari buah jambu biji merah yang dibagi menjadi empat kategori yaitu merah, setengah matang, matang, dan busuk.

Qodri (2016) melakukan penelitian untuk mendapatkan mutu pada buah Mangga Manalagi dengan menggunakan pengolahan citra. Metode yang digunakan adalah dengan mengambil citra buah Mangga Manalagi yang kemudian dicari segmentasi serta ekstraksi dari citra. Dengan menggunakan metode ini akurasi yang didapatkan untuk pemutuan buah Mangga Manalagi adalah sebesar 100% untuk kelas mutu A, kemudian kelas mutu B, dan mutu D yang memiliki akurasi produksi yaitu sebesar 80%, dan yang terakhir yaitu kelas mutu C yaitu memiliki akurasi produksi sebesar 60%.

Nurdayati dkk (2016) mengklasifikasikan tingkat kematangan buah Mangga Harum Manis berdasarkan normalisasi warna dengan menggunakan metode Fuzzy Logic berbasis web. Fuzzy logic yang digunakan pada penelitian ini dengan menggunakan metode mamdani yang terdiri dari tiga variabel input berupa red, green dan blue, sedangkan variabel output berupa tingkat kematangan dari buah Mangga yang diklasifikasikan menjadi tiga bagian, mentah, sedang dan matang. Hasil dari penelitian ini adalah metode fuzzy mamdani dapat mengklasifikasikan kematangan buah Mangga Harum Manis dengan nilai akurasi sebesar 40%. Pada penelitian ini belum adanya ekstraksi fitur sehingga hanya dapat dilakukan untuk satu jenis buah mangga.

Pada penelitian ini diajukan metode GLCM dan K-Means Clustering untuk Ekstraksi Fitur dan metode LS-SVM untuk mendeteksi tingkat kematangan buah mangga. Penggunaan metode ini diharapkan mampu digunakan untuk mengklasifikasi lebih dari satu jenis buah mangga serta mendapatkan hasil akurasi yang optimal.

II. TINJAUAN PUSTAKA

2.1 Buah Mangga

Mangga merupakan tanaman buah tahunan yang berasal dari India yang pada saat ini telah banyak menyebar ke wilayah Asia Tenggara, khususnya di Indonesia dan Malaysia. Tanaman mangga tumbuh baik pada dataran rendah dengan ketinggian 0-500 meter yang memiliki tanah yang mengandung pasir dan lempung, tanaman mangga dapat tumbuh baik di daerah yang terbuka. (AAK, 1991).

Mangga manalagi merupakan suatu jenis buah mangga yang memiliki ciri-ciri yaitu memiliki ukuran sedang sampai besar dengan berat sekitar 350- 400 gram. Bentuk buah bulat, letak tangkai di tengah, pangkal buah runcing, sedikit berleher, dan kulit buah tebal. Jika dilihat dari warnanya, mangga manalagi dinyatakan matang jika pada pangkal buah telah menjadi kuning dan pucuk buah hijau. Kemudian jika dilihat dari daging buahnya, mangga manalagi dinyatakan matang jika daging buah tebal, lunak berwarna kuning, berserat halus dan memiliki aroma yang harum (Badan Standardisasi Nasional, 1992).

2.2 Skala Keabuan

Citra Mangga yang diterima adalah citra berwarna, sehingga terlebih dahulu perlu dilakukan proses skala keabuan untuk mendapatkan citra dengan arah keabuan. Jumlah warna pada citra keabuan adalah 256, karena citra keabuan jumlah bitnya adalah 8, sehingga jumlah warnanya adalah $2^8=256$, nilainya berada pada jangkauan 0-255. Untuk mendapatkan citra keabuan digunakan persamaan(1).

$$I(x, y) = \alpha \cdot R + \beta \cdot G + \gamma \cdot B \quad (1)$$

dimana $I(x,y)$ adalah level keabuan pada suatu koordinat yang diperoleh dengan mengatur komposisi warna R (merah), G (hijau), B (biru) yang ditunjukkan oleh nilai parameter α , β dan γ . Secara umum nilai α , β dan γ adalah 0.33. Nilai yang lain juga dapat diberikan untuk ketiga parameter tersebut asalkan total keseluruhan nilainya adalah 1 (Putra dan suarjana, 2010).

2.3 Ekstraksi Fitur GLCM

Pengukuran nilai tekstur yang digunakan didasarkan pada persamaan (Haralick et al, 1973 dan Conner et al. 1984). Menggunakan notasi berikut: G adalah jumlah tingkat abu-abu yang digunakan. m adalah nilai rata-rata dari $P_{\mu x}$, μy ,

σ_x and σ_y adalah means dan standard deviations P_x dan P_y . $P_x(i)$ adalah entri- i dalam matrik smarginal-probabilitas yang diperoleh dengan menjumlahkan barisan $P(i, j)$.

Berikut ini merupakan GLCM orde pertama:

Mean

$$\mu = \sum_n f_n \cdot p(f_n) \quad (2)$$

Skewness

$$a_3 = \sum_n \frac{E(p(f_n) - \mu)^3}{\sigma^3} \quad (3)$$

Variance

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (f_n - \mu)^2 \quad (4)$$

Kurtosis

$$a_4 = \sum_n \frac{E(p(f_n) - \mu)^4}{\sigma^4} \quad (5)$$

Entropy

$$H = - \sum_n P(f_n) \cdot \log_2 p(f_n) \quad (6)$$

Berikut ini merupakan GLCM orde kedua:

Entropi (Entropy) Menunjukkan ukuran ketidakteraturan bentuk. Persamaan :

$$= - \sum_{i=0}^{G-1} \sum_{j=0}^{G1} P(i, j) \times \log(P(i, j)) \quad (7)$$

Kontras (Contrast)

$$= \sum_{i=0}^{G-1} n^2 \left\{ \sum_{i=1}^{G1} \sum_{j=1}^{G1} P(i, j) \right\}, |i - j| \quad (8)$$

$= n$

Angular Second Moment (ASM)

ASM adalah ukuran homogenitas dari suatu gambar. Dan didefinisikan :

$$ASM = \sum_{i=0}^{G-1} \sum_{j=0}^{G1} \{p(i, j)\}^2 \quad (9)$$

Korelasi (Correlation)

Korelasi menunjukkan ketergantungan linear derajat keabuan dari piksel-piksel yang saling bertetangga dalam suatu citra abu-abu. Persamaan :

$$= \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{\{ixj\}xP(i,j) - \{\mu_x\mu_y\}}{\sigma_x\sigma_y} \tag{10}$$

dimana :

μ_x = nilai rata-rata elemen kolom pada matriks P θ (i,j)

μ_y = nilai rata-rata elemen baris pada matriks P θ (i,j)

σ_x = nilai standar deviasi elemen kolom pada matriks P θ (i,j)

σ_y =nilai standar deviasi elemen baris pada matriks P θ (i,j)

Momentum Selisih invers (Inverse Difference Momentum)

$$IDM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1 + (i - j)^2} p (i,j) \tag{11}$$

IDM juga dipengaruhi oleh homogenitas gambar. Karena faktor bobot $(1 + (i-j)^2)^{-1}$ IDM akan mendapatkan kontribusi kecil dari daerah homogen ($i \neq j$). Hasilnya adalah nilai IDM rendah untuk gambar tidak homogen, dan nilai yang relatif lebih tinggi untuk gambar homogen.

Varians (Variance)

$$VARIANCE = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (k - \mu)^2 p (i,j) \tag{12}$$

2.4 K-Means Clustering

K-Means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data ke dalam cluster atau kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan dari clustering ini adalah untuk meminimalisasi objective function yang diset dalam proses clustering, yang pada umumnya berusaha meminimalisasi variasi di dalam suatu cluster dan memaksimalkan variasi antar cluster.

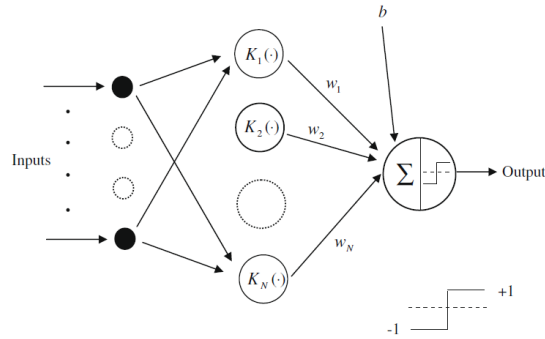
2.5 Support Vector Machines (SVM)

SVM yang diusulkan oleh Vapnik (1995) telah dipelajari secara ekstensif untuk klasifikasi, regresi dan estimasi kepadatan. Gambar 2.1. adalah arsitektur SVM. SVM memetakan pola input ke ruang fitur dimensi yang lebih tinggi melalui pemetaan nonlinier berdasar teori yang dipilih. Bidang pemisah linear ini kemudian dibangun dalam ruang fitur dimensi tinggi. Dengan demikian, SVM adalah linear classifier di ruang parameter, tapi itu menjadi nonlinear classifier sebagai akibat dari pemetaan nonlinear dari ruang pola input ke ruang fitur dimensi tinggi. Bila data pelatihan berdimensi m adalah x_i ($i = 1, \dots, M$) dan masing-masing kelas labelnya adalah y_i , di mana $y_i = 1$ dan $y_i = -1$ untuk kelas 1 dan 2. Jika data input terpisah secara linear di ruang fitur, maka fungsi keputusan dapat ditentukan:

$$D(x) = w \cdot g(x) + b \tag{13}$$

di mana $g(x)$ adalah fungsi pemetaan yang memetakan x ke dalam ruang dimensi 1, w adalah vektor dimensi dan 1, dan b adalah skalar. Untuk memisahkan data secara linier, fungsi keputusan memenuhi kondisi berikut:

$$y_i(w \cdot g(x_i) + b) > 1 \text{ untuk } i = 1, \dots, M \tag{14}$$



Gambar 2.1. Arsitektur SVM

Jika masalah terpisah secara linier dalam ruang fitur, maka fungsi keputusan yang memenuhi persamaan (16) jumlahnya tak terbatas. Di antara fungsi-fungsi tersebut, diperlukan hyperplane dengan margin terbesar antara dua kelas. Margin adalah jarak minimum yang memisahkan hyperplane terhadap data input dan ini dihasilkan dari \$|D(x)|/\|w\|\$. Sehingga didapatkan hyperplane pemisah dengan margin maksimal yang optimal memisahkan hyperplane.

Dengan asumsi bahwa margin adalah \$\rho\$, kondisi berikut harus memenuhi:

$$\frac{y_i D(x_i)}{\|w\|} \geq \rho \quad \text{untuk } i = 1, \dots, M \quad (15)$$

Hasil perkalian produk dari \$\rho\$ dan \$\|w\|\$ adalah tetap:

$$\rho \|w\| = 1 \quad (16)$$

Untuk mendapatkan hyperplane pemisah yang optimal dengan margin maksimal, \$w\$ dengan \$\|w\|\$ yang memenuhi persamaan (15) harus ditemukan. Persamaan (16) mengarahkan ke pemecahan masalah optimasi berikutnya. Dengan meminimalkan

$$\frac{1}{2} w^t w \quad (17)$$

dan mengikuti batasan:

$$y_i(w^t g(x_i) + b) > 1 \quad \text{untuk } i = 1, \dots, M \quad (18)$$

Bila data pelatihan tidak linier dipisahkan, digunakan slack variable \$\xi_i\$ ke persamaan (26):

$$y_i(w^t g(x_i) + b) > 1 - \xi_i \quad \xi_i > 0 \quad \text{untuk } i = 1, \dots, M \quad (19)$$

Hyperplane pemisah yang optimal telah ditentukan sehingga maksimalisasi dari margin dan meminimalisasi dari kesalahan training didapatkan. Dengan meminimalkan

$$\frac{1}{2} w^t w + \frac{C}{2} \sum_{i=1}^n \xi_i^p \quad (20)$$

mengikuti batasan:

$$y_i(w^T g(x_i) + b) > 1 - \xi_i \quad \xi_i > 0 \text{ untuk } i = 1, \dots, M \quad (21)$$

di mana C adalah parameter yang menentukan tradeoff antara margin maksimum dan kesalahan klasifikasi minimum dan ρ adalah 1 atau 2. Jika $\rho = 1$, SVM disebut SVM dengan soft margin L1 (L1-SVM), dan jika $\rho = 2$, SVM dengan soft margin L2 (L2-SVM). Pada SVM konvensional, hyperplane pemisah yang optimal diperoleh dengan memecahkan masalah pemrograman kuadratik.

Fungsi kernel memungkinkan operasi yang akan dilakukan di ruang input bukan di ruang fitur dimensi tinggi. Beberapa contoh fungsi kernel adalah $K(u, v) = v^T u$ (SVM linier); $K(u, v) = (v^T u + 1)^n$ (SVM polinomial derajat n); $K(u, v) = \exp(-\|u - v\|^2 / 2\sigma^2)$ (SVM fungsi radial bases – SVM RBF); $K(u, v) = \tanh(Kv^T u + o)$ (neural SVM dua layer) di mana σ, κ, o adalah konstanta (Vapnik, 1995; Cortes, 1995). Namun, fungsi kernel yang tepat untuk suatu masalah tertentu tergantung pada data, dan sampai saat ini belum ada metode yang baik tentang cara memilih fungsi kernel.

2.6 Least Squares Support Vectors Machine (LS-SVM)

Least Squares Support Vectors Machine (LS-SVM) adalah salah satu modifikasi dari SVM (Suykens, 1999). Jika SVM dikarakteristik oleh permasalahan konveks quadratic programming dengan pembatas berupa pertidaksamaan, LS-SVM sebaliknya, diformulasikan dengan menggunakan pembatas yang hanya berupa persamaan. Sehingga solusi LS-SVM dihasilkan dengan menyelesaikan persamaan linier. Hal ini tentunya berbeda dengan SVM yang mana solusinya dihasilkan melalui penyelesaian quadratic programming. Saat ini, LS-SVM banyak dilakukan pada klasifikasi dan estimasi fungsi (Suykens, 1999).

LS-SVM di-training dengan meminimalkan

$$\frac{1}{2} w^T w + \frac{C}{2} \sum_{i=1}^n \xi_i^2 \quad (22)$$

dan mengikuti batasan persamaan:

$$y_i(w^T g(x_i) + b) > 1 - \xi_i \quad \xi_i > 0 \text{ untuk } i = 1, \dots, M \quad (23)$$

Pada LS-SVM, batasan persamaan digunakan sebagai pengganti pertidaksamaan yang digunakan pada SVM konvensional. Karena itu, solusi yang optimal dapat diperoleh dengan menyelesaikan sekumpulan persamaan linier bukan dengan penyelesaian quadratic programming. Untuk menurunkan dua masalah persamaan (25) dan (26), Lagrange multiplier digunakan, yaitu :

$$Q(w, b, \alpha, \xi) = \frac{1}{2} w^T w + \frac{C}{2} \sum_{i=1}^n \xi_i^2 - \sum_{i=1}^n \alpha_i \{ y_i (w^T g(x_i) + b) - 1 + \xi_i \} \quad (24)$$

dimana $\alpha = (\alpha_1, \dots, \alpha_M)^T$ adalah Lagrange multiplier yang bisa bernilai positif atau negatif pada rumus LS-SVM. Kondisi yang optimum diperoleh dengan mendiferensialkan persamaan di atas terhadap w, ξ_i, b , dan α_i dan persamaan dihasilkan sama dengan nol. (Suykens, 1999; Tsujinishi, 2003).

Seperti pada SVM konvensional, fungsi kernel memungkinkan operasi yang akan dilakukan di ruang input bukan di ruang fitur dimensi tinggi. Beberapa penelitian menggunakan LS-SVM dan fungsi kernel RBF (LS-SVM RBF) secara empiris menghasilkan hasil yang optimal. (Suykens, 1999, Polat, 2007, Ubeyli, 2010a, 2010b). Untuk masalah klasifikasi dua-spiral yang kompleks dapat ditemukan dengan LS-SVM RBF dengan kinerja yang sangat baik dan komputasi rendah. (Suykens, 1999)

2.7 Evaluasi Kinerja Sistem

Evaluasi dilakukan dengan tujuan untuk mengevaluasi efektivitas metode dan sistem yang telah dibuat. Evaluasi dilakukan terhadap tingkat akurasi klasifikasi dan tingkat kesalahan klasifikasi. Ukuran atau parameter yang digunakan untuk evaluasi antara lain akurasi klasifikasi, sensitivitas, spesifisitas, nilai prediksi positif, nilai prediksi negatif, kurva ROC dan matriks konfusi (confusion matrix). Matriks konfusi berisi informasi tentang klasifikasi yang sebenarnya dan yang diperkirakan dari hasil sistem klasifikasi. Tabel 2.1 menunjukkan matriks konfusi untuk dua kelas klasifikasi. Akurasi klasifikasi, sensitivitas, spesifisitas, nilai prediksi positif dan nilai prediksi negatif dapat didefinisikan menggunakan elemen-elemen matriks konfusi sebagai berikut:

$$\text{Klasifikasi akurasi (\%)} = \frac{TP + TN}{TP + FP + FN + TN}, \quad (25)$$

$$\text{Sensitivitas (\%)} = \frac{TP}{TP + FN}, \quad (26)$$

$$\text{Spesifisitas (\%)} = \frac{TN}{FP + TN}, \quad (37)$$

$$\text{Nilai prediksi positif} = \frac{TP}{TP + FP} \times 100, \quad (28)$$

$$\text{Nilai prediksi negatif} = \frac{TN}{FN + TN} \times 100. \quad (29)$$

Kurva ROC adalah teknik yang dapat diandalkan karena didasarkan pada nilai-nilai true positive dan false positive sehinggamenunjukkan trade-off antara sensitivitas dan spesifisitas (Akay, 2009).

Selain itu juga analisis data hasil uji coba dilakukan terhadap waktu komputasi dari kombinasi fitur yang ada. Evaluasi dilakukan dengan melihat perubahan waktu komputasi yang dibutuhkan oleh tiap kombinasi fitur.

Tabel 2.1 Matriks Konfusi

Aktual	Prediksi	
	Positif	Negatif
Positif	True Positive (TP)	False Negative (FN)
Negatif	False Positive (FP)	True Negative (TN)

III. METODE PENELITIAN

3.1 Studi Literatur

Dalam tahap studi literatur, dikaji beberapa referensi yang berkaitan dengan praproses dan ekstraksi. Selain itu, dikaji pula metode-metode yang akan digunakan dan dikembangkan, antara lain : GLCM, K-Means Clustering dan LS-SVM.

3.2 Pengumpulan Data

Pada penelitian ini, dataset yang digunakan adalah foto citra buah Mangga local dan ditambah dengan citra foto yang didapatkan dari data publik. Data yang akan digunakan adalah sebanyak 240 data dengan rincian 80 dataset untuk manga jenis “Harum Manis”, 80 dataset untuk manga jenis “Manalagi” dan 80 data untuk manga jenis “Kent”, masing-masing jenis terdiri dari 40 Buah Mangga yang sudah matang dan 40 Buah Mangga yang masih mentah.

3.3 Analisis dan Desain

Tahap analisis dalam penelitian ini meliputi perumusan masalah, batasan-batasan masalah, dan penyelesaiannya. Permasalahan yang akan diselesaikan adalah bagaimana mengklasifikasikan mangga berdasarkan tingkat kematangan dengan menggunakan ekstraksi fitur GLCM dan K-Means Clustering dan Klasifikasi menggunakan LS-SVM.

Sistem klasifikasi kematangan mangga meliputi tahap pelatihan dan pengujian. Tahap pelatihan dimulai dengan menginputkan citra mangga, selanjutnya pada citra akan dilakukan proses pra-pengolahan. Citra diubah terlebih dahulu kedalam format grayscale green channel. Selanjutnya Ekstraksi ciri pada proses pelatihan dilakukan dengan menggunakan metode GLCM dan K-Means Clustering.

Ekstraksi ciri pada proses pengujian dilakukan dengan mengambil hasil ekstraksi ciri pada proses pelatihan diterapkan pada data pengujian. Hasil ekstraksi ciri pada data pengujian ini nantinya digunakan sebagai inputan pada proses klasifikasi pengujian.

Tahap terakhir dari metode penelitian adalah melakukan klasifikasi seleksi fitur dengan LS-SVM dengan RBF kernel (LS-SVM RBF). Data training untuk masing-masing seleksi fitur yang dihasilkan selanjutnya di-training dengan LS-SVM RBF. Proses training dilakukan dengan nilai parameter LS-SVM RBF (γ dan σ^2) yang ditentukan secara trial and error melalui uji coba penentuan parameter percobaan. γ adalah parameter regulerisasi, yang menentukan trade-off antara margin maksimum dan kesalahan klasifikasi minimum. Pada beberapa penelitian sebelumnya nilai γ disebut sebagai C penalty (Akay, 2009). Sedangkan σ^2 adalah bandwidth untuk fungsi kernel RBF. Penentuan nilai awal dan akhir parameter γ dan σ^2 secara trial and error dilakukan dengan merujuk pada nilai yang terdapat pada manual toolbox Matlab LS-SVMlab1.5 (Pelckmans 2002, 2003). Nilai parameter γ dan σ^2 yang dihasilkan tersebut digunakan untuk proses training untuk tiap kombinasi fitur.

Hasil proses dari masing-masing seleksi fitur training pada classifier LS-SVM RBF digunakan untuk menguji seleksi fitur data testing dengan LS-SVM RBF. Hasil klasifikasi berupa class label tersebut dibandingkan dengan class label sebenarnya. Penelitian ini disebut sebagai supervised learning karena class label telah diketahui sebelumnya. Pengujian tersebut menggunakan nilai parameter γ dan σ^2 yang sama dengan saat training. Hasil pengujian tersebut digunakan untuk proses evaluasi dari sistem yang telah dikembangkan. Proses training dilakukan menggunakan fungsi `trainlssvm` dan proses testing menggunakan fungsi `latentlssvm` yang telah disediakan oleh toolbox Matlab LS-SVMlab1.5 (Pelckmans 2002, 2003). Tahapan seluruh proses klasifikasi seleksi fitur dengan LS-SVM RBF seperti yang telah dijelaskan tersebut dapat dilihat pada Gambar 3.1.



Gambar 3.1. Tahap Klasifikasi Seleksi Fitur

Pada proses klasifikasi pengujian menggunakan hasil ekstraksi ciri data pengujian dan hasil proses klasifikasi pelatihan. Hasil dari proses ini berupa nilai indeks dari fungsi keputusan yang terbesar yang menyatakan kelas dari data pengujian. Jika kelas yang dihasilkan dari proses klasifikasi pengujian sama dengan kelas data pengujian, maka pengenalan dinyatakan benar.

Pada Gambar 3.2 merupakan tahapan proses sistem pengklasifikasian kematangan mangga. Pada proses pelatihan terdapat metode yang digunakan untuk mengekstraksi ciri, ciri-ciri yang terpilih pada saat proses pelatihan digunakan dalam proses klasifikasi waktu uji coba. Masing-masing dataset citra mangga yang digunakan dibagi menjadi dua, sebagian digunakan untuk proses pelatihan (training) dan sisanya digunakan untuk proses pengujian (testing).



Gambar 3. 2 Gambaran Umum Sistem

3.4 Implementasi Perangkat Lunak

Pada tahap ini, desain yang telah dirancang sebelumnya diimplementasikan menggunakan Matlab 2013a.

3.5 Uji Coba dan Analisis

Melakukan pengujian dan analisis terhadap hasil dan performa metode yang diimplementasikan.

IV. HASIL DAN PEMBAHASAN

4.1 Lingkungan Implementasi

Untuk mengerjakan semua metode yang diusulkan pada penelitian ini maka dibutuhkan lingkungan implementasi yang mencukupi, agar semua tahap kegiatan dapat berjalan dengan baik. Tabel 4.1. memperlihatkan lingkungan implementasi yang digunakan pada Penelitian ini.

Tabel 4.1. Lingkungan Implementasi

Jenis Perangkat	Spesifikasi
Laptop	Asus FX504
Prosesor	Inter Core i5-8300H
Memory	8GB
Hard Disk	1TB
Sistem Operasi	Windows 10 Home
Perangkat Pengembang	Matlab 2013a dan toolbox LS-SVM 1.8

4.2 Data Set Mangga

Pada Penelitian ini akan digunakan 240 data set Citra Mangga. 160 Dataset Citra berupa Mangga jenis “Harum Manis” dan jenis “Manalagi” diambil dengan menggunakan Camera Handphone, sedangkan 80 Dataset Berupa Mangga jenis “Kent” diambil dari database public pada link http://www.cofilab.com/wp-content/downloads/DB_Mango.rar.

Seluruh dataset pada masing masing jenis mangga terdiri dari 40 Mangga mentah dan 40 Mangga yang sudah Matang. Dari 240 Mangga akan dibagi menjadi 120 data latih dan 120 data uji, untuk data latih setiap jenis mangga terdiri dari 20 Mangga Mentah dan 20 Mangga yang sudah Matang. Begitu juga, untuk data uji setiap jenis mangga terdiri dari 20 buah Mangga Mentah dan 20 buah Mangga yang sudah Matang.

4.3 Uji Coba Normalisasi Citra

Normalisasi Citra adalah perubahan seluruh ukuran citra. Seluruh data, baik data yang diambil secara manual maupun data yang diambil dari internet diubah ukurannya menjadi 480 X 640 Pixel. Normalisasi Citra ini dilakukan untuk mempercepat proses komputasi.

4.4 Uji Coba Segmentasi Citra

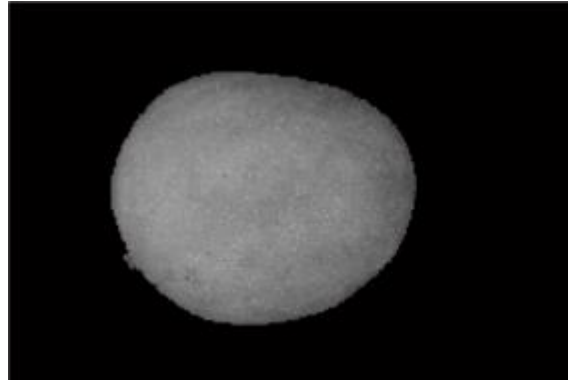
Setelah Citra melalui proses normalisasi kemudian citra mangga akan disegmentasi dengan menggunakan K-Means Clustering. Proses ini perlu dilakukan untuk mendapatkan ciri bentuk dari citra Mangga yang akan diklasifikasi. Dari proses hanya akan diambil nilai metric dan eccentricity untuk dimasukkan sebagai ciri bentuk dari citra mangga. Contoh Citra Hasil dari proses segmentasi dapat dilihat pada Gambar 4.1.



Gambar 4.1. Citra Hasil Segmentasi

4.5 Uji Coba Ekstraksi Fitur menggunakan GLCM

Proses Ekstraksi Fitur dilakukan dengan terlebih dahulu mengubah citra dari RGB menjadi citra abu dengan memanfaatkan fungsi `rgb2gray` pada MATLAB. Seperti yang terlihat pada Gambar 4.2.



Gambar 4.1. Citra Grayscale

Kemudia, dari citra grayscale buah Mangga dilakukan proses ekstraksi fitur dengan menggunakan GLCM. Setiap data citra akan diambil 4 fitur utama dari proses ekstraksi, yaitu: contrast, correlation, energy, homogeneity. Seluruh fitur tersebut yang selanjutnya digunakan untuk proses klasifikasi. Pada Tabel 4.2 dapat dilihat contoh seluruh fitur yang dihasilkan setelah melalui proses segmentasi dengan menggunakan metode K-Means Clustering dan proses ekstraksi fitur dengan menggunakan metode GLCM.

Tabel 4.2 Fitur Citra Mangga

No	Fitur	Nilai
1	Metric	0,4942
2	Eccentricity	0,6646
3	Contrast	0,0523
4	Correlation	0,9906
5	Energy	0,5207
6	Homogeneity	0,9789

4.6 Uji Coba Klasifikasi Menggunakan LS-SVM

Uji coba Klasifikasi dengan Menggunakan LS-SVM dengan RBF kernel (LS-SVM RBF) dimulai dengan proses pencarian nilai gamma dan sig2 (γ dan σ^2) yang optimal dengan memanfaatkan fungsi tunnelssvm pada modul ls-svm matlab. Setelah melakukan beberapa kali percobaan ditemukan nilai optimal untuk gamma adalah 12.6361 dan nilai optimal untuk sig2 adalah 1.3253.

Langkah selanjutnya adalah melakukan proses training pada data training dengan menggunakan fungsi trainlssvm pada modul ls-svm matlab, pada fungsi ini ada enam parameter yang digunakan, yaitu: (1) Daftar Fitur dari data training, (2) Label Kelas untuk data training, (3) tipe ls-svm menggunakan klasifikasi, (4) nilai gamma yang diperoleh dari proses sebelumnya, (5) nilai sig2 yang diperoleh dari proses sebelumnya, dan (6) RBF Kernel, untuk kernel klasifikasi.

Nilai alpha dan beta yang didapatkan pada proses training kemudian digunakan pada proses uji. Proses uji memanfaatkan fungsi simlssvm pada modul ls-svm matlab. Fungsi ini membutuhkan 3 parameter yaitu : (1) Model LS-SVM yang digunakan, (2) nilai alpha dan beta dari proses training, dan (3) Fitur dari data uji. Hasil dari proses ini adalah nilai uji klasifikasi berupa mangga mentah atau mangga matang dengan diwakili nilai 0 dan 1.

4.7 Tingkat Akurasi Klasifikasi.

Proses uji coba dengan menggunakan LS-SVM dilakukan pada data training dan pada data uji. Proses uji coba pada data training dilakukan untuk mendapatkan nilai gamma dan sigma yang paling optimal. Hasil dari proses uji coba pada data training dilakukan berkali-kali sehingga mendapatkan nilai akurasi 100% dengan nilai optimal untuk gamma

adalah 12.6361 dan nilai optimal untuk sigma adalah 1.3253. Proses uji coba pada data uji dengan menggunakan nilai gamma dan sigma optimal mendapatkan nilai akurasi sebesar 98,33 %, nilai presisi sebesar 98,33, dan nilai recall sebesar 98,33. Hasil Konfusi Matriks untuk uji coba pada data uji bisa dilihat pada Tabel 4.2. dengan 59 data True Positive, 1 data False Negative, 1 data False Positive, dan 59 data True Negative.

Tabel 4.2 Konfusi Matriks Hasil Klasifikasi

Aktual	Prediksi	
	Positif	Negatif
Positif	59 (TP)	1 (FN)
Negatif	1 (FP)	59 (TN)

4.8 Waktu Komputasi

Waktu Komputasi yang dibutuhkan untuk proses ekstraksi fitur untuk setiap dataset adalah rata-rata sebesar 0,85 detik, sedangkan waktu komputasi untuk uji coba pada data uji adalah rata-rata sebesar 0,006 detik. Dapat disimpulkan bahwa waktu komputasi untuk setiap dataset adalah rata-rata sebesar 0,856 detik.

V. KESIMPULAN DAN SARAN

Penggunaan Metode GLCM dan K-Means Clustering untuk Ekstraksi Fitur digabungkan dengan metode LS-SVM untuk klasifikasi dapat digunakan untuk mendapatkan tingkat kematangan lebih dari satu jenis buah Mangga dengan nilai akurasi yang tinggi.

Dari Hasil Pengujian Pada 240 Data set uji Mangga jenis Harum Manis, Manalagi, dan Kent dengan label klasifikasi Matang dan Mentah didapatkan hasil akurasi, presisi dan recall dengan nilai yang sama yaitu sebesar 98,33%.

Jumlah Fitur yang digunakan untuk proses klasifikasi sangat mempengaruhi tingkat akurasi dan waktu komputasi, semakin banyak fitur yang digunakan maka tingkat akurasi semakin tinggi tetapi membuat waktu komputasi yang semakin lama, sedangkan jika menggunakan fitur yang sedikit maka waktu komputasi cenderung lebih cepat tapi akan mendapatkan nilai akurasi yang relative lebih rendah.

Diperlukan dataset Mangga Jenis Lain untuk menkonfirmasi bahwa metode ini bisa diterapkan untuk seluruh jenis buah Mangga. Diperlukan metode seleksi fitur agar mendapatkan nilai komputasi dan nilai akurasi yang optimal. Diperlukan perbandingan metode ekstraksi Fitur selain dengan Menggunakan GLCM dan K-Means Clustering. Diperlukan perbandingan metode Klasifikasi selain dengan menggunakan metode LS-SVM.

VI. DAFTAR PUSTAKA

- [1] AAK. 1991. Mangga. Yogyakarta: Kanisius. [Serial on line]. <https://books.google.co.id/books?id=j02Pj6omzlsC&pg=PA19&dq=mangga&hl=id&sa=X&ved=0ahUKEwiMqOCrcDPahVDuo8KHcMnBaAQ6AEIHzAB#v=onepage&q=mangga&f=false>. [04 Juni 2017].
- [2] Cortes, C., & Vapnik, V. 1995. Support vector networks. *Machine Learning*, 20 (3), 273–297.
- [3] Hermawati, Fajar Astuti. 2013. *Pengolahan Citra Digital*. Yogyakarta: Andi Offset.
- [4] QN Setiawan, dkk. 2015. *Pengelompokan Kayu Kelapa Menggunakan Algoritma K-Means Berdasarkan Tekstur Citra Kayu Kelapa Dua Dimensi*. Universitas Dian Nuswantoro, Semarang.
- [5] Nurdayati, Niki, dkk. 2016. *Klasifikasi Kematangan Buah Mangga Harum Manis Berdasarkan Normalisasi Warna Menggunakan Metode Fuzzy Logic Berbasis Web*. FMIPA Universitas Pakuan Bogor.
- [6] Mulato Febry, 2014. *Klasifikasi kematangan buah jambu biji merah (psidium guajava) dengan menggunakan model fuzzy*. Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Yogyakarta, Yogyakarta.

- [7] Zheng, Hong, dkk. 2012. A least-squares support vector machine (LS-SVM) based on fractal analysis and CIELab parameters for the detection of browning degree on mango (*Mangifera indica* L.). *Computer and Electronic in Agriculture*. Vol. 83, 47-51.
- [8] Kolar, R., Harabis, V. 2009. Automatic Rigid Registration and Analysis of Colour Fundus Image in Patient with Diabetic Retinopathy, *IFMBE proceedings*, vol 25/11, 251-24.
- [9] Cortes, C., & Vapnik, V. 1995. Support vector networks. *Machine Learning*, 20 (3), 273–297.
- [10] Putra, D., Suarjana, I.G., 2010. Segmentasi citra retina digital retinopati diabetes untuk membantu pendeteksian mikroaneurisma. *Jurnal teknologi teknik elektro* vol 9 no 1.
- [11] Hermawati, Fajar Astuti. 2013. *Pengolahan Citra Digital*. Yogyakarta: Andi Offset.
- [12] Wahyudi, dkk. 2012. Sistem Deteksi Retinopati Diabetik Menggunakan Support Vector Machine. *Jurnal Sistem Informasi Bisnis* 03.