

Penerapan Metode *K-Nearest Neighbor* dan *Support Vector Machine* untuk Klasifikasi Kematangan Buah Mengkudu

Application of K-Nearest Neighbor Method and Support Vector Machine for Noni Fruit Ripeness Classification

SITI GAYATRI HEHANUSSA¹, SONY HARTONO WIJAYA^{*1}, TOTO HARYANTO¹

Abstrak

Buah mengkudu (*Morinda citrifolia*) merupakan salah satu komoditas ekspor buah-buahan di Indonesia yang selalu tersedia di setiap musim dan dikenal memiliki berbagai manfaat kesehatan. Buah mengkudu berasal dari wilayah Asia Tenggara, termasuk Indonesia, dan sering digunakan dalam pengobatan tradisional. Pada umumnya masyarakat menentukan kematangan buah mengkudu secara manual, yaitu dengan menggunakan penampakan visual. Hal ini menyebabkan adanya perbedaan persepsi dalam menentukan tingkat kematangan buah mengkudu. Oleh karena itu, penelitian ini bertujuan membangun model *machine learning* untuk klasifikasi tingkat kematangan buah mengkudu. Metode klasifikasi yang digunakan adalah *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) dengan menggunakan ekstraksi fitur warna *Hue Saturation Intensity* (HSI) dan ekstraksi fitur tekstur *Local Binary Pattern* (LBP). Pengklasifikasian yang dilakukan pada buah mengkudu dengan algoritma KNN menghasilkan model klasifikasi yang lebih baik daripada menggunakan algoritma SVM. Akurasi terbaik yang dihasilkan oleh KNN sebesar 88.62% pada $k=11$, sedangkan akurasi terbaik SVM dengan *kernel* polynomial sebesar 87.80%, menggunakan parameter $C=0.1$, $\text{Gamma}=1$, $\text{Degree}=5$, dan $\text{coef0}=1.0$. Hasil ini didapatkan dari data latih dan data uji dengan perbandingan 80:20.

Kata Kunci: *k-nearest neighbor*, mengkudu, *support vector machine*

Abstract

Noni fruit (*Morinda citrifolia*) is one of Indonesia's export commodities. It is available year-round and is well known for its numerous health benefits. Native to Southeast Asia, including Indonesia, noni fruit is widely used in traditional medicine. Typically, the ripeness of noni fruit is determined manually based on visual inspection, which can lead to subjective judgments and inconsistent results. Therefore, this study aims to develop a machine-learning model to classify the ripeness levels of noni fruit. The classification methods employed are *K-Nearest Neighbor* (KNN) and *Support Vector Machine* (SVM), utilizing *Hue-Saturation-Intensity* (HSI) color features and *Local Binary Pattern* (LBP) texture features. Experimental results show that the KNN algorithm outperforms the SVM algorithm in terms of classification accuracy. The highest accuracy achieved using KNN was 88.62% at $k = 11$, whereas the best accuracy obtained with SVM using a polynomial kernel was 87.80%, with parameters set to $C = 0.1$, $\text{Gamma} = 1$, $\text{Degree} = 5$, and $\text{coef0} = 1.0$. These results were achieved using an 80:20 split ratio for training and testing data.

Keywords: *k-nearest neighbor*, noni, *support vector machine*

¹ Program Studi Ilmu Komputer, Sekolah Sains Data, Matematika dan Informatika, IPB University, Bogor (16680)

* Penulis korespondensi: Surel: sony@apps.ipb.ac.id

PENDAHULUAN

Buah mengkudu (*Morinda citrifolia*) telah lama dikenal dalam pengobatan tradisional dan semakin mendapatkan perhatian dalam penelitian ilmiah modern. Khasiatnya meliputi sejumlah kondisi medis seperti menguatkan tulang, membersihkan darah, peluruh kencing, peluruh haid, melembutkan kulit, obat batuk, obat cacing, melancarkan peredaran darah, dan menghilangkan ketombe (Almukarramah *et al.* 2019). Selain itu, Manfaat lainnya termasuk meredakan sakit kepala, mengatasi pembengkakan limpa, serta membantu mengurangi keluhan pada organ hati dan saluran pencernaan lainnya (Mayanti dan Khusniati 2022).

Salah satu tantangan dalam produksi buah mengkudu adalah proses identifikasi tingkat kematangan buah mengkudu saat ini masih dilakukan secara manual. Terdapat tantangan variabilitas dalam penilaian kematangan oleh masyarakat atau pelaku dalam industri pertanian. Penilaian yang subjektif ini dapat berpotensi mengurangi kualitas dan nilai gizi buah. Kecenderungan terjadinya kesalahan dalam identifikasi kematangan juga meningkat akibat perbedaan persepsi dan pengetahuan yang dimiliki oleh individu yang terlibat dalam proses ini.

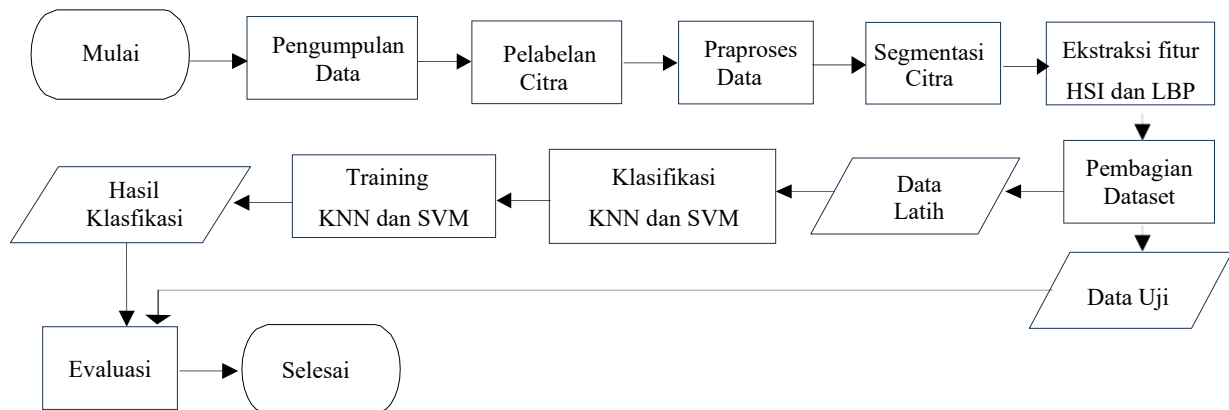
Dalam konteks analisis kematangan buah mengkudu, penggunaan teknologi pemanfaatan citra digital berbasis komputer memainkan peran yang penting dalam mengatasi tantangan yang dihadapi oleh pendekatan manual. Teknik pengolahan citra dalam penelitian ini menggunakan komponen warna *Hue Saturation Intensity* (HSI) pada citra dengan akurasi yang lebih tinggi (Auliasari *et al.* 2020) dan ekstraksi fitur tekstur *Local Binary Pattern* (LBP) untuk memahami karakteristik tata ruang lokal dari citra (Hidiya dan Lasulika 2019). Selain itu, penelitian ini menggunakan metode *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) untuk melakukan klasifikasi terhadap citra buah mengkudu.

Penelitian terkait klasifikasi kematangan buah mengkudu menggunakan SVM dan KNN telah di diterapkan oleh Kusuma *et al.* (2017), dengan hasil akurasi SVM terbaik sebesar 87.22% dibandingkan dengan KNN. Kekurangan pada penelitian tersebut adalah kurangnya set data yang digunakan (60 set data), sehingga menjadi kendala untuk merepresentasikan hasil klasifikasi yang baik. Penelitian lain dengan metode yang sama diterapkan oleh Irawan *et al.* (2023) pada klasifikasi kematangan buah mengkudu dengan model terbaik diperoleh berdasarkan metode KNN yang menghasilkan akurasi 95%. Namun, saran dalam penelitian tersebut adalah mengimplementasikan fitur ekstraksi LBP serta menghitung nilai *recall* dan presisi atau *F1-Score* sebagai evaluasi model. Dalam kasus yang sama, penelitian Haadi dan Hidayat (2021) juga menggunakan metode SVM sebagai klasifikasi kematangan buah mengkudu dengan warna RGB sehingga mencapai akurasi 92%.

Berdasarkan penelitian yang telah dilakukan, penelitian lanjutan dapat dilakukan dengan mengkaji kelas kematangan buah mengkudu matang dan tidak matang dengan melibatkan ekstraksi fitur LBP dan HSI yang belum pernah dilakukan sebelumnya. Penelitian ini bertujuan untuk menghasilkan model klasifikasi buah mengkudu untuk mengatasi tantangan identifikasi manual yang sering kali bersifat subjektif dan kurang konsisten. Pendekatan ini, diharapkan dapat diperoleh informasi yang lebih mendalam dan objektif terkait karakteristik buah mengkudu. Selain itu juga untuk melengkapi kekurangan pada penelitian terdahulu.

METODE

Penelitian ini dilakukan melalui beberapa tahapan-tahapan yang terstruktur dan terarah. Gambar 1 menunjukkan tahapan penelitian yang dilakukan pada penelitian ini.

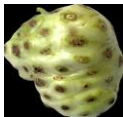


Gambar 1 Tahapan penelitian

Pengumpulan Data

Data primer citra buah mengkudu diperoleh melalui survei lapangan yang dilaksanakan di wilayah Desa Gemba, Kecamatan Kairatu, Kabupaten Seram Bagian Barat, Provinsi Maluku. Pengambilan data citra berlangsung efektif selama 5 bulan, mulai dari tanggal 20 September 2023 hingga 20 Januari 2024. Pengambilan citra dilakukan pagi hari antara pukul 07.00 sampai 10.00 WIT. Hasil pengumpulan data pada tahap ini yaitu telah dilakukan pengumpulan data dengan melakukan pemotretan buah mengkudu menggunakan kamera *smartphone* kapasitas 12MP dengan jarak pemotretan 40 cm. Pengambilan buah mengkudu menggunakan ujung galah, jaring, gunting panen, dan melibatkan petani mengkudu yaitu Bapak Wawan dalam penentuan label buah mengkudu sesuai kategori kelas tua dan muda. Kelas citra mengkudu dan jumlah tiap kelas dapat dilihat pada Tabel 1.

Tabel 1 Kelas citra mengkudu dan jumlah tiap kelas

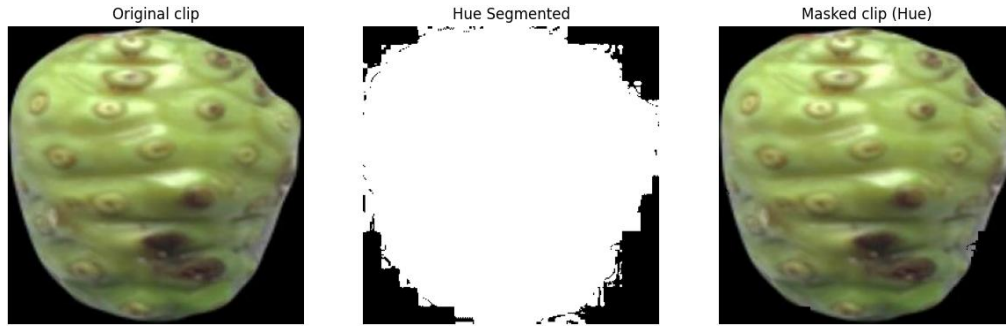
Citra	Kelas	Jumlah
	Matang	295
	Tidak Matang	319

Praproses Data

Praproses data dilakukan untuk mendapatkan set data yang berkualitas baik. Tahapan praproses dimulai dengan pemotongan citra (*cropping*) untuk menyamakan resolusi menggunakan perangkat lunak pengolah citra seperti aplikasi *image resizer* atau perpustakaan pemrosesan citra dalam bahasa pemrograman seperti OpenCV (Ektrada *et al.* 2023). Selanjutnya, dilakukan penghapusan latar belakang menggunakan aplikasi *remove.bg* agar tampilan citra menjadi lebih jelas. Proses *resizing* citra kemudian dilakukan untuk mengubah resolusi citra yang bervariasi menjadi ukuran 256×256 piksel. Setelah itu, citra mengalami segmentasi menggunakan *mask* yang terdiri dari nilai biner (0 dan 1) untuk menentukan area tertentu yang akan diproses, dengan tujuan menyempurnakan hasil segmentasi, serta melakukan normalisasi pada citra mengkudu. Normalisasi dilakukan untuk membantu mempermudah pemrosesan dan membuat model lebih akurat.

Segmentasi Citra

Teknik pengolahan gambar yang diperlukan untuk mendeteksi kematangan buah adalah segmentasi. Segmentasi citra merupakan proses mengubah representasi gambar menjadi sesuatu yang lebih mudah dianalisis (Rema 2019). Proses ini diperlukan untuk memisahkan buah dan daging buahnya, karena hanya daging buahnya saja yang perlu diproses untuk menentukan kematangan buah (Septiarini *et al.* 2022). Gambar 2 menunjukkan contoh hasil segmentasi citra.



Gambar 2 Segmentasi citra

Gambar 2 menggunakan *mask* dengan teknik *thresholding* pada *channel Hue* yang digunakan dalam pengolahan citra untuk memisahkan atau mengidentifikasi bagian-bagian tertentu dari sebuah gambar. *Mask* merupakan matriks biner yang menunjukkan nilai 1 untuk piksel yang berada dalam objek atau area yang diinginkan, dan nilai 0 untuk piksel di luar area tersebut (Medinah dan Sinurat 2020).

Ekstraksi Fitur

Setelah dilakukan praproses pada citra, tahapan selanjutnya adalah proses ekstraksi fitur. Proses ekstraksi fitur bertujuan untuk menghasilkan nilai-nilai yang merepresentasikan karakteristik khas objek yang sedang dianalisis. Ekstraksi fitur yang dilakukan pada penelitian ini adalah ekstraksi fitur tekstur dan ekstraksi fitur warna. Ekstraksi fitur warna menggunakan HSI yang terdiri dari *hue*, *saturation*, *intensity* dan fitur tekstur menggunakan LBP berupa nilai *mean*, *variance*, dan *entropy*.

- a. Teknik pengolahan citra yang menggunakan *Hue Saturation Intensity* (HSI), seperti yang digunakan dalam penelitian Bora *et al.* (2015) memungkinkan pemisahan antara komponen warna, saturasi, dan intensitas dalam citra. Dalam konteks analisis kematangan buah mengkudu, hal ini dapat membantu dalam mengidentifikasi perubahan warna yang terjadi seiring dengan tingkat kematangan. Langkah-langkah dalam membuat konversi RGB ke HSI dapat di lihat pada Persamaan 1, 2, 3 dan 4 (Yanto *et al.* 2021).

- Normalisasi RGB ke rentang 0-1

$$R', G', B' = \frac{R}{255}, \frac{G}{255}, \frac{B}{255} \quad (1)$$

yang mana, R adalah saluran warna merah, G adalah saluran warna hijau, dan B adalah saluran warna biru.

- Hitung *Hue*:

$$\theta = \cos^{-1} \left(\frac{(R' - G') + (R' - B')}{2\sqrt{(R' - G')^2 + (R' - B')(G' - B')}} \right) \quad (2)$$

$$H = \begin{cases} 0, & \text{jika } B' \leq G' \\ 360^\circ - \theta, & \text{jika } B' > G' \end{cases} \quad (3)$$

di mana, *Hue* diukur dalam derajat ($^\circ$) dalam rentang 0-360 $^\circ$.

- Hitung *Saturation* (S):

$$S = 1 - \frac{\min(R', G', B')}{I}, \text{ jika } I > 0 \quad (4)$$

- Hitung *Intensity* (I):

$$I = \frac{1}{3} (R' + G' + B') \quad (5)$$

- b. Teknik pengolahan citra *Local Binary Pattern* (LBP) merupakan sebuah citra informasi berupa susunan struktur permukaan suatu gambar (Alwi *et al.* 2019). Fitur tekstur cocok digunakan untuk memvisualisasikan pola dan sifat permukaan yang menyediakan informasi tentang variasi intensitas permukaan citra. *Local Binary Pattern* (LBP) merupakan salah satu metode yang digunakan untuk melakukan ekstraksi fitur tekstur pada sebuah citra yang pertama kali dikenalkan oleh Ojala *et al.* (1996). Pada ekstraksi fitur LBP ini akan menggunakan tiga jenis ciri fitur yaitu *Mean*, *Variance*, dan *Entropy*. Formulasi rumus dasar matematika dari operator (LBP) dituliskan pada Persamaan 6 (Lindahl 2007).

$$\text{LBP}_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) \cdot 2^p \quad (6)$$

yang mana : P = banyaknya piksel tetangga, R = nilai jarak/radius, g_c = nilai dari piksel x dan y . g_p = nilai piksel tetangga x_c, y_c = koordinat pusat. Adapun perhitungan dari tiga jenis fitur tersebut dapat dilihat dari Persamaan 7, 8, 9 (Ershad 2011).

- *Mean*

Merupakan nilai rata-rata pada sebuah citra. Nilai mean dapat ditentukan dengan membagi jumlah citra dengan banyaknya citra.

$$f1 = \sum_{i,j=0}^{N-1} P_{i,j} - \ln(P_{i,j}) P_{i,j} \quad (7)$$

- *Variance*

Menunjukkan variasi warna keabuan pada sebuah citra. Nilai *Variance* akan meningkat ketika nilai tingkat keabuannya berbeda dari *mean*nya.

$$f2 = \sum_{i,j=0}^{N-1} P_{i,j} (i - \mu)^2 \quad (8)$$

- *Entropy*

Menunjukkan kompleksitas suatu citra. Nilai *Entropy* akan tinggi ketika citra tidak seragam. Citra dengan tekstur yang kompleks akan cenderung memiliki nilai *Entropy* yang tinggi.

$$f3 = \sum_{i,j=0}^{N-1} -\ln(P_{i,j}) P_{i,j} \quad (9)$$

dengan: (i,j) = Baris i dan kolom j , $\sum_i$ = Jumlah nilai dari baris i , $\sum_j$ = Jumlah nilai dari kolom j , μ = Nilai rata-rata elemen pada matriks, $p_d(i, j)$ = Elemen dalam matriks pada baris i dan kolom j .

Model Klasifikasi

Klasifikasi merupakan suatu proses pengelompokan sebuah data dengan data lain yang memiliki karakteristik yang sama (Widiastuti *et al.* 2017). Penelitian ini memanfaatkan metode SVM dan KNN untuk proses klasifikasi, dengan membandingkan akurasi hasil klasifikasi dari masing-masing algoritma. Data penelitian ini terdiri dari 80% sampel untuk data latih dan 20% sampel untuk data uji, seperti yang ditunjukkan pada Tabel 2.

Tabel 2 Pembagian data latih dan data uji

Kelas	Data Latih	Data Uji	Jumlah
Matang	230	65	295
Tidak matang	261	58	319
Total	491	123	614

Algoritma *k-Nearest Neighbor* (KNN)

Metode klasifikasi KNN menentukan kelas suatu data uji berdasarkan kedekatan jaraknya dengan data latih yang paling dekat (Hastuti *et al.* 2018). Dalam konteks analisis kematangan buah mengkudu, KNN dapat digunakan untuk mengidentifikasi kematangan buah berdasarkan karakteristik yang diambil dari citra buah tersebut. Dengan mengevaluasi data pembelajaran

yang memiliki kesamaan dengan objek yang akan diklasifikasikan, metode KNN dapat memberikan estimasi tingkat kematangan secara objektif. Jarak setiap objek dapat dihitung menggunakan jarak *Euclidean* yang dapat dilihat Persamaan 10 (Hadi dan Rachmawanto 2022).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (10)$$

dengan (d) = jarak kedekatan (x) = data *testing* (y) = data *training* (n) = jumlah atribut 1 sampai n.

Algoritma *Support Vector Machine* (SVM)

Metode klasifikasi (SVM) awalnya diusulkan oleh Cortes dan Vapnik (Nugroho *et al.* 2021). SVM merupakan salah satu teknik yang digunakan dalam klasifikasi citra dalam algoritma pembelajaran mesin yang membangun model untuk melakukan klasifikasi dengan memisahkan dua kelas secara optimal (Maneno *et al.* 2023). Algoritma ini didasarkan pada prinsip pembelajaran statistik, dengan fokus pada memaksimalkan margin, yaitu jarak antara *hyperplane* optimal sebagai garis pemisah dan titik-titik data terdekat dari masing-masing kelas (Wisudawati *et al.* 2020).

SVM juga dapat digunakan untuk pemisahan nonlinier dengan menggunakan fungsi *kernel* lainnya. Ada banyak fungsi pemetaan yang tersedia, termasuk fungsi basis linier, polinomial, dan radial (RBF). Pemilihan fungsi *kernel* Polinomial atau RBF biasanya tergantung pada set data pelatihan yang digunakan (Aziez dan Anuraga 2021). Secara matematika, konsep dasar SVM dapat dilihat pada Persamaan 11 dan 12.

$$\text{Fungsi objektif} = \min \frac{1}{2} |w|^2 \quad (11)$$

$$\text{Kendala} = y_t(x_t \cdot w + b) \geq 1 \quad (12)$$

dengan, x_t adalah titik untuk data ke-t, y_t adalah label kelas untuk x_t (+1 atau -1), $x_t \cdot w + b$ adalah fungsi keputusan. Parameter yang digunakan dapat dilihat pada Tabel 3.

Tabel 3 Keterangan parameter SVM

Parameter	Nilai
Kernel	Linier, RBF, Polynomial.
C	Mengontrol <i>trade-off</i> antara pencapaian margin maksimal dan minimalisir kesalahan. Nilai yang digunakan (C: 0.1, 1, 10, 100).
Gamma	Menentukan seberapa jauh pengaruh sebuah sampel pelatihan dalam menentukan <i>hyperplane</i> keputusan. Nilai yang digunakan gamma: (0.0001, 0.001, 0.01, 1).
Degree	Menciptakan fungsi polinomial dengan kompleksitas yang lebih tinggi, Ini memungkinkan model untuk membuat batas keputusan yang lebih kompleks. Nilai yang digunakan <i>degree</i> : (2,3,4,5).
Coef0	Dioptimalkan bersama dengan gamma, C, dan degree untuk mencapai hasil klasifikasi yang optimal, Nilai yang digunakan <i>coef0</i> : (0.0, 0.1, 0.5, 1.0).

Confusion Matrix

Confusion Matrix merupakan metrik pengujian yang digunakan dalam mengevaluasi model klasifikasi. Metrik ini merupakan hasil perhitungan dari data yang diklasifikasi sebagai benar atau salah yang dibandingkan dengan label sebenarnya pada data yang telah diklasifikasi. Nilai yang diperoleh dari *confusion matrix* tersebut dapat digunakan untuk menghitung *accuracy*, *precision*, *recall* dan *F1-score* sebagai hasil nilai evaluasi dari suatu model. Tabel 4 menunjukkan ilustrasi dari *confusion matrix* (Manurung *et al.* 2024).

Tabel 4 *Confusion matrix*

Prediksi	True	False
Positive	True Positive (TP)	False Positive (FP)
Negative	False negative (FN)	True Negative (TN)

- Akurasi (*Accuracy*): Akurasi merupakan metrik evaluasi yang menghitung persentase prediksi benar dari seluruh prediksi yang dihasilkan oleh model. Dalam klasifikasi, akurasi menunjukkan seberapa sering model berhasil mengklasifikasikan data dengan tepat, baik untuk kelas positif maupun negatif. Rumus matematika untuk menghitung akurasi adalah Persamaan 13.

$$Akurasi = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (13)$$

- b. Presisi (*Precision*): Presisi adalah metrik evaluasi yang mengukur keakuratan prediksi model dalam mengidentifikasi kelas positif, yaitu persentase prediksi positif yang benar dari seluruh prediksi positif yang dihasilkan. Presisi dinyatakan sebagai rasio antara jumlah prediksi benar positif dengan jumlah total prediksi positif. Rumus matematika untuk menghitung presisi dapat dilihat pada Persamaan 14.

$$Presisi = \frac{TP}{FP+FP} \quad (14)$$

- c. Recall (*Recall*): Sensitivitas (*Recall*) merupakan metrik evaluasi yang mengukur seberapa baik suatu model dari SVM dan KNN dalam mengidentifikasi kelas positif dengan benar. Rumus matematika untuk menghitung *recall* terdapat pada Persamaan 15.

$$Recall = \frac{TP}{TP+FN} \quad (15)$$

- d. F1-Score (*F1-score*): *F1-Score* merupakan metrik evaluasi yang memadukan antara presisi dan *recall* yang mencerminkan keseimbangan antara keduanya. F1-Score merupakan metrik evaluasi yang mencerminkan keseimbangan antara Presisi dan *Recall*. Rumus matematika untuk menghitung F1-Score dilihat pada Persamaan 16.

$$F1\ score = \frac{2*(Presisi*Recall)}{(Presisi+Recall)} \quad (16)$$

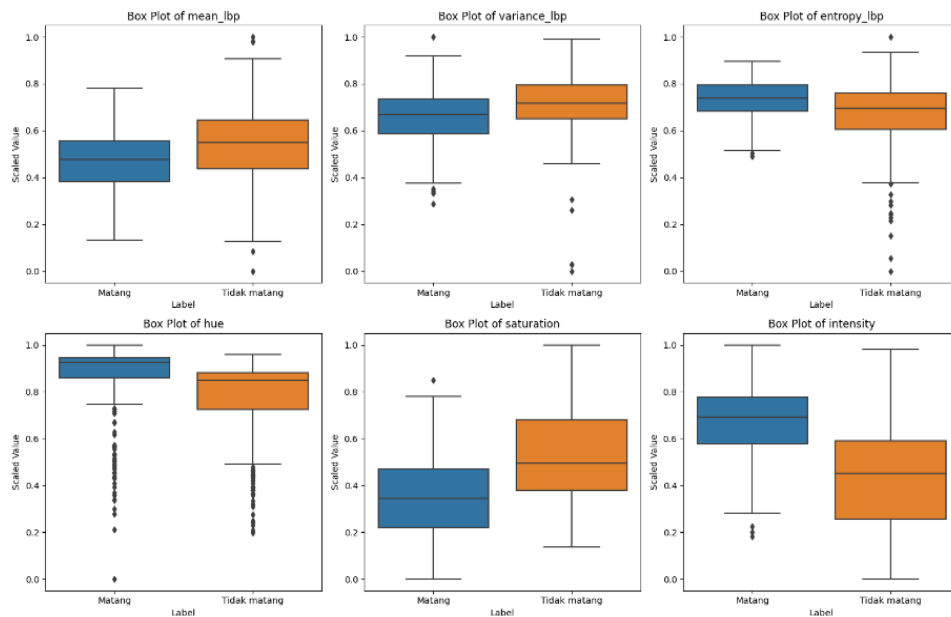
Analisis dan Evaluasi

Analisis dan evaluasi kinerja model dilakukan untuk memastikan bahwa model mampu memprediksi atau mengklasifikasikan data dengan akurat. Kinerja algoritma dievaluasi menggunakan *confusion matrix* (Prasetyawan dan Gatra 2022). Hasil analisis dalam penelitian ini menunjukkan bahwa model memiliki performa yang baik, terutama dalam hal presisi dan akurasi, meskipun *recall* rendah.

HASIL DAN PEMBAHASAN

Analisis Outlier pada Fitur HSI dan LBP Berdasarkan Boxplot

Hasil *boxplot* pada Gambar 3 menunjukkan bahwa HSI pada komponen warna '*Saturation*' memiliki outlier yang paling sedikit diantara fitur HSI lainnya dan LBP, sedangkan fitur yang paling banyak outlier nya ada pada komponen warna '*Hue*', dan fitur yang lain rata rata outlier nya sedikit.

Gambar 3 Tampilan *boxplot* tiap fitur HSI dan LBP

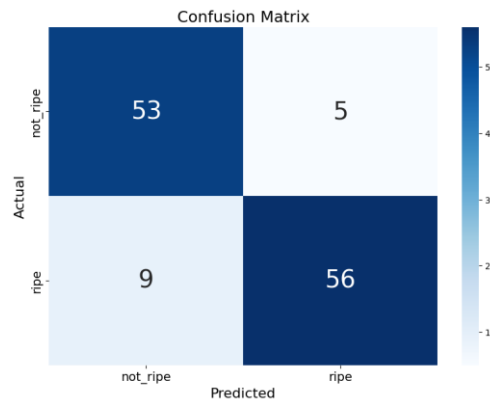
Hasil Pemodelan dengan KNN dan SVM

Proses hasil dari pengujian klasifikasi buah mengkudu menggunakan 2 skenario pengujian.

- a. Skenario 1: Melakukan klasifikasi kematangan buah mengkudu menggunakan algoritma KNN dengan jumlah k berbeda beda. Pengujian untuk nilai $k=1$ hingga $k=20$, waktu pelatihan dan prediksi juga dicatat untuk masing-masing nilai k . Proses memilih nilai k terbaik berdasarkan kinerja model dan efisiensi waktu dapat dilihat pada Tabel 5. Tabel 5 menunjukkan hasil klasifikasi menggunakan KNN dengan nilai $k1$ hingga $k20$ dengan hasil *Accuracy*, *Precision*, *Recall*, *F1-Score* menggunakan fitur warna HSI dan tekstur LBP. Hasil pengujian menunjukkan bahwa algoritma KNN pada kasus ini memiliki performa terbaik untuk semua metrik yang diujikan. Nilai $k=11$ dengan akurasi tertinggi sebesar 88.62%. Algoritma KNN memerlukan waktu *training* tercepat untuk kasus ini sebesar 2.3 detik dan waktu prediksi 9.7 detik. Gambar 4 menunjukkan analisis error model KNN pada *confusion matrix*.

Tabel 5 Hasil klasifikasi KNN

K	<i>Accuracy</i> (%)	<i>Precision</i> (%)	<i>Recall</i> (%)	<i>F1-Score</i> (%)	Waktu pelatihan (detik)	Waktu prediksi (detik)
1	85.37%	88.52%	83.08%	85.71%	3.9	12.7
2	82.11%	95.74%	69.23%	80.36%	3.0	11.4
3	85.37%	92.72%	78.46%	85.00%	2.8	11.4
4	84.55%	96.00%	73.85%	83.48%	2.9	11.5
5	86.99%	91.52%	83.08%	87.10%	3.0	11.5
6	86.18%	92.85%	80.00%	85.95%	2.8	11.5
7	87.80%	91.67%	84.61%	88.00%	2.8	11.5
8	86.18%	92.85%	80.00%	85.95%	4.2	12.5
9	86.18%	90.00%	83.08%	86.40%	2.5	10.3
10	86.99%	91.52%	83.08%	87.10%	2.5	10.8
11	88.62%	91.80%	86.15%	88.88%	2.7	11.2
12	87.80%	91.67%	84.61%	88.00%	2.5	10.7
13	87.80%	90.32%	86.15%	88.19%	2.4	10.3
14	86.18%	90.00%	83.08%	86.40%	2.4	09.7
15	86.99%	88.88%	86.15%	87.50%	2.3	10.2
16	86.99%	90.16%	84.61%	87.30%	2.3	09.8
17	87.80%	90.32%	86.15%	88.19%	2.3	10.5
18	86.99%	90.16%	84.61%	87.30%	2.2	09.8
19	87.80%	90.32%	86.15%	88.19%	2.7	11.5
20	87.80%	91.67%	84.61%	88.00%	2.7	11.3

Gambar 4 *Confusion matrix* KNN

Confusion matrix menunjukkan hasil pada model KNN, dengan buah mengkudu matang sebanyak 295 dan tidak matang sebanyak 319 telah diklasifikasikan, dengan tingkat akurasi yang bergantung pada jumlah prediksi yang benar dan salah dari model.

- b. Skenario 2: Melakukan klasifikasi kematangan buah mengkudu menggunakan algoritma SVM dengan *kernel* Linier, RBF dan Polynomial menggunakan parameter C , γ , degree , Coef0 . Pengukuran nilai akurasi klasifikasi dilakukan berdasarkan Persamaan 13.

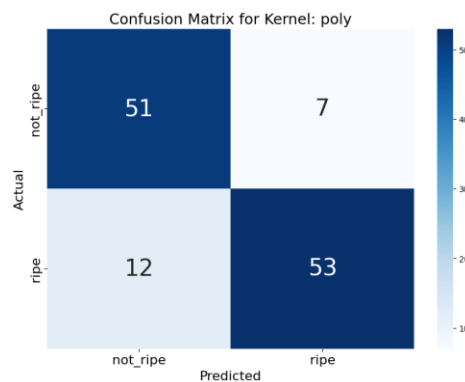
Tabel 6 Hasil klasifikasi SVM

No	Kernel	Akurasi (%)	Presisi (%)	Recall (%)	F1-score (%)	Waktu pelatihan (detik)	Waktu prediksi (detik)
1	Linier	82.93%	85.48%	87.69%	83.46%	6.5	1.8
2	RBF	86.18%	100%	87.69%	86.82%	6.5	2.3
3	Polynomial	87.80%	100%	89.23%	88.55%	2.0	2.0

Tabel 6 memperlihatkan hasil klasifikasi buah mengkudu berdasarkan fitur warna dan tekstur yang telah dilakukan menggunakan kernel linier, RBF, dan polynomial. Dari hasil tersebut, terlihat bahwa SVM dengan kernel Polynomial mencapai akurasi sebesar 87.80% dan *F1-score* tertinggi sebesar 88.55%, dengan waktu komputasi tercepat yaitu 2.0 detik. Parameter terbaik yang digunakan pada model ini adalah $C=0.1$, $\gamma=1$, $\text{degree}=5$, dan $\text{coef0}=1.0$. Hasil pengujian keseluruhan performa model menggunakan kedua algoritma ditampilkan pada Tabel 7 dengan *confusion matrix kernel* polynomial yang dapat dilihat pada Gambar 5.

Tabel 7 Nilai *accuracy*, *precision*, *recall*, dan *f1-score* pada masing – masing algoritma

Algoritma	Metrik Pengujian			
	<i>Accuracy(AC)</i>	<i>Precision(PR)</i>	<i>Recal(R)</i>	<i>F1-Score(FI)</i>
	$\frac{tp + tn}{tp + tn + fp + fn}$	$\frac{tp}{tp + tp}$	$\frac{tp}{tp + tn}$	$\frac{2xRCxPR}{RC + PR}$
KNN	88.62%	96%	86.15%	88.88%
SVM Linier	87.60%	85.48%	87.69%	83.46%
SVM RBF	86.18%	100%	87.69%	86.82%
SVM Polynomial	87.80%	100%	87.69%	88.55%

Gambar 5 *Confusion matrix* polynomial

Confusion matrix ini menunjukkan hasil klasifikasi menggunakan kernel polynomial dengan dua kelas: belum matang dan matang. Model berhasil mengklasifikasikan 51 sampel belum matang dengan benar, tetapi salah mengklasifikasikan 7 sampel belum matang sebagai ripe (matang), serta mengklasifikasikan 53 sampel matang dengan benar, namun salah mengklasifikasikan 12 sampel matang sebagai belum matang.

Hasil prediksi data uji menggunakan model yang telah dibangun menunjukkan hasil analisis seperti akurasi, presisi, recall dan *F1-score* pada algoritma KNN dan SVM dapat dijelaskan pada Tabel 8 dan Tabel 9.

Tabel 8 Hasil prediksi data uji menggunakan model KNN

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
Belum matang	0.85	0.91	0.88	58
Matang	0.92	0.86	0.89	65
Akurasi			0.89	123
Rata-Rata Makro	0.89	0.89	0.89	123
Rata-Rata Terbobot	0.89	0.89	0.89	123

Tabel 8 menunjukkan evaluasi kinerja model klasifikasi berdasarkan metrik *precision*, *recall*, dan *F1-score* untuk masing-masing kelas, yaitu "Belum Matang" dan "Matang". Nilai *F1-score* untuk kelas "Belum Matang" adalah 0.88, sedangkan untuk kelas "Matang" adalah 0.89. Hal ini menunjukkan bahwa model mampu mengenali kedua kelas dengan performa yang seimbang.

Rata-rata makro dan rata-rata terbobot dari semua metrik memiliki nilai yang sama, yaitu 0.89, mencerminkan konsistensi performa model terhadap distribusi data yang relatif seimbang. Akurasi keseluruhan model juga berada pada angka 0.89, yang menandakan bahwa 89% prediksi model sesuai dengan label sebenarnya.

Tabel 9 Hasil prediksi data uji menggunakan model SVM *polynomial*

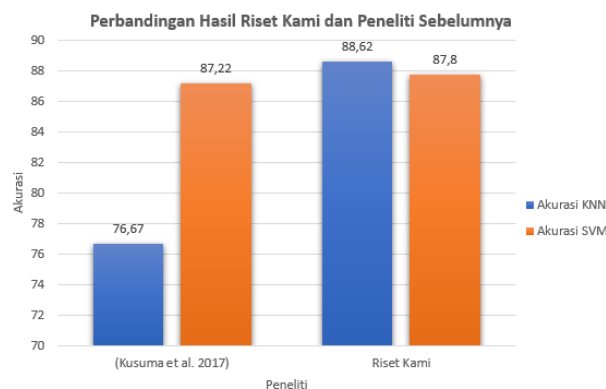
Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
Belum matang	0.81	0.88	0.84	58
Matang	0.88	0.82	0.85	65
Akurasi			0.85	123
Rata-Rata Makro	0.85	0.85	0.85	123
Rata-Rata Terbobot	0.85	0.85	0.85	123

Tabel 9 menyajikan hasil evaluasi performa model dalam mengklasifikasikan buah mengkudu ke dalam dua kelas, yaitu "Belum Matang" dan "Matang" menggunakan metrik *precision* (ketepatan), *recall* (kemampuan mendeteksi data sebenarnya), *F1-score* (harmoni antara *precision* dan *recall*), serta *support* (jumlah data pada tiap kelas).

Nilai *F1-score* untuk kelas "Belum Matang" adalah 0.84, sedangkan untuk kelas "Matang" sebesar 0.85. Ini menunjukkan bahwa model memiliki kemampuan yang cukup baik dan relatif seimbang dalam mengenali kedua kelas. Nilai akurasi keseluruhan sebesar 0.85 memperlihatkan bahwa 85% prediksi model sesuai dengan label sebenarnya dalam data uji.

Rata-rata makro dan rata-rata terbobot pada semua metrik juga menunjukkan nilai yang sama, yaitu 0.85. Hal ini menunjukkan bahwa model tidak hanya bekerja baik secara

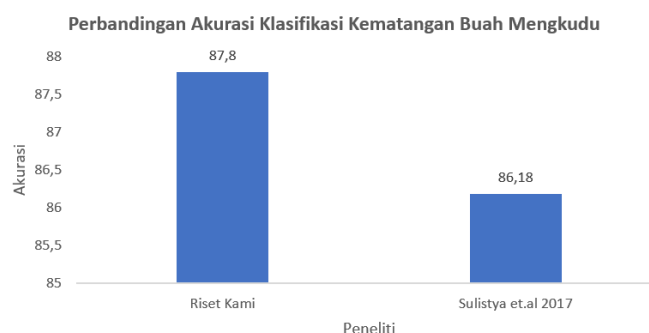
keseluruhan, tetapi juga memperhatikan keseimbangan performa antar kelas, meskipun terdapat sedikit variasi pada nilai *recall* masing-masing kelas.



Gambar 6 Hasil perbandingan riset saat ini dan penelitian sebelumnya (KNN dan SVM)

Gambar 6 menunjukkan perbandingan akurasi antara riset saat ini dan penelitian sebelumnya oleh Kusuma *et al.* (2017) menggunakan algoritma KNN dan SVM dengan ekstraksi fitur GLCM. Hasil riset saat ini dengan ekstraksi fitur LBP dan komponen warna HSI menunjukkan peningkatan akurasi pada KNN sebesar 88.62%, lebih tinggi dibandingkan penelitian sebelumnya (76.67%). Sementara akurasi SVM juga mengalami sedikit peningkatan dari 87.22% menjadi 87.8%.

Perbandingan lainnya kami lakukan terhadap hasil penelitian Sulistya *et al.* (2024), dengan menggunakan studi kasus yang sama namun menggunakan komponen warna yang berbeda (HSV dan HIS). Hasilnya dapat dilihat pada Gambar 7. Hasil menunjukkan bahwa riset saat ini (87.8%) memiliki akurasi yang lebih unggul dibandingkan riset sebelumnya (86.18%).



Gambar 7 Hasil perbandingan riset saat ini dan penelitian sebelumnya (SVM)

SIMPULAN

Penelitian ini menunjukkan bahwa klasifikasi kualitas buah mengkudu dapat dicapai melalui pengolahan citra digital dengan tingkat akurasi yang optimal saat menggunakan fitur warna HSI dan tekstur LBP. Algoritma KNN dengan pengujian *k11* menghasilkan akurasi tertinggi sebesar 88.62% unggul dibandingkan kinerja algoritma SVM dengan *kernel* polynomial menghasilkan akurasi sebesar 87.80% pada percobaan ini. Penggunaan metode ekstraksi fitur lain dapat diuji dalam penelitian selanjutnya dengan mengembangkan aplikasi *mobile* yang mampu melakukan analisis kematangan buah mengkudu secara efektif. Selain itu, dapat ditambahkan lebih banyak set data pada kelas setengah matang buah mengkudu guna memperluas penerapan teknologi pengolahan citra dalam klasifikasi.

DAFTAR PUSTAKA

- Almukarramah A, Ibrahim I, Sufriadi S. 2019. Tanaman Berkhasiat Obat dari Sub Kelas Sympetaleae yang digunakan Masyarakat. *Serambi Saintia J Sains dan Apl.* 7(1):18–25. doi:10.32672/jss.v7i1.988.
- Alwi L, Hermawan AT, Kristian Y. 2019. Identifikasi Biji-Bijian Berdasarkan Ekstraksi Fitur

- Warna, Bentuk dan Tekstur Menggunakan Random Forest. *J Intell Syst Comput.* 1(2):92–98. doi:10.52985/insyst.v1i2.93.
- Auliasari R, Novamizanti L, Ibrahim N. 2020. Identifikasi Kematangan Daun Teh Berbasis Fitur Warna Hue Saturation Intensity (HSI) dan Hue Saturation Value (HSV). *JUITA J Inform.* 8:217. doi:10.30595/juita.v8i2.7387.
- Al Aziez H, Anuraga G. 2021. Klasifikasi Daerah Tertinggal di Indonesia Menggunakan Algoritma SVM dan k-NN. *J Ilmu Dasar.* 22(1)(1):31–38.
- Bora DJ, Gupta AK, Khan AF. 2015. Comparing the Performance of L*A*B* and HSV Color Spaces with Respect to Color Image Segmentation. *Int J Emerg Technol Adv Eng.* 5(2).
- Ektrada E, Hakim L, Kristanto SP. 2023. Sistem Tracking dan Counting Kendaraan Berbasis YOLO untuk Pemetaan Slot Parkir Kendaraan. *Softw Dev Digit Bus Intell Comput Eng.* 1(02):55–60. doi:10.57203/session.v1i02.2023.55-60.
- Ershad SF. 2011. Texture classification approach based on combination of edge & Co-occurrence and Local binary pattern. *Proc 2011 Int Conf Image Process Comput Vision, Pattern Recognition, IPCV 2011.* 2:626–629.
- Hadi HP, Rachmawanto EH. 2022. Analisa Fitur Ekstraksi Ciri Dan Warna Dalam Proses Klasifikasi Kematangan Buah Rambutan Berbasis K-Nearest Neighbor. *Skanika.* 5(2):177–189. doi:10.36080/skanika.v5i2.2944.
- Hastuti MT, Widodo AW, Dewi C. 2018. Identifikasi Kondisi Kesehatan Ayam Petelur Berdasarkan Ciri Warna HSV Dan Gray Level Cooccurrence Matrix (GLCM) Pada Citra Jengger Dengan Klasifikasi K-Nearest Neighbour. Volume ke-2.
- Hidiya SR, Lasulika ME. 2019. Fitur Ekstraksi LBP Untuk Mengidentifikasi Kematangan Tomat Sayur Menggunakan Metode K-Nearest Neighbor. *J Nas cosPhi.* 3(1):2597–9329.
- Irawan C, Rachmawanto EH, Atika Sari C, Umah Nur R. 2023. Klasifikasi Citra Mengkudu Berdasarkan Perhitungan Jarak Pikel pada Algoritma K-Nearest Neighbour. *Infotekmesin.* 14(2):200–207. doi:10.35970/infotekmesin.v14i2.1827.
- Kusuma SF, Pawening RE, Dijaya R. 2017. Otomatisasi klasifikasi kematangan buah mengkudu berdasarkan warna dan tekstur. *Regist J Ilm Teknol Sist Inf.* 3(1):17–23. doi:10.26594/register.v3i1.576.
- Lindahl T. 2007. Study of Local Binary Patterns. *Sci Technol.*, siap terbit.
- Maneno R, Baso B, Manek PG, Fallo K. 2023. Deteksi Tingkat Kematangan Buah Pinang Menggunakan Metode Support Vector Machine Berdasarkan Warna Dan Tekstur. *J Inf Technol.* 3(2):60–66. doi:10.32938/jitu.v3i2.5323.
- Manurung J, Setiawan A, Cahyo Untoro M. 2024. Identification of Fatigue from Facial Expressions Using Transfer Learning. *Media Comput Sci.* 1(1):1–16. doi:10.69616/mcs.
- Mayanti ANR, Khusniati M. 2022. Konsep Sains dalam Kebiasaan Mengonsumsi Wedang Mengkudu untuk Menurunkan Tekanan Darah pada Masyarakat Desa Pladen dan Terban. *Proceeding Semin Nas.*, siap terbit.
- Medinah DRE, Sinurat S. 2020. Analisa dan Perbandingan Algoritma Otsu Thresholding dengan Algoritma Region Growing Pada Segmentasi Citra Digital. *J Comput Syst Informatics.* 2(1):9–16.
- Nugroho BA, Pradana AKA, Nurfarida E. 2021. Prediksi Waktu Kedatangan Pelanggan Servis Kendaraan Bermotor Berdasarkan Data Historis menggunakan Support Vector Machine. *J Edukasi dan Penelit Inform.* 7(1):25. doi:10.26418/jp.v7i1.42964.
- Ojala T, Pietikäinen M, Harwood D. 1996. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.* 29(1):51–59. doi:https://doi.org/10.1016/0031-3203(95)00067-4.
- Prasetyawan D, Gatra R. 2022. Algoritma K-Nearest Neighbor untuk Memprediksi Prestasi Mahasiswa Berdasarkan Latar Belakang Pendidikan dan Ekonomi. *JISKA (Jurnal Inform Sunan Kalijaga).* 7(1):56–67. doi:10.14421/jiska.2022.7.1.56-67.
- Rema YOL. 2019. Deteksi Plat Nomor Kendaraan Bermotor dengan Segmentasi Gambar. *J Saintek Lahan Kering.* 2(1):20–23. doi:10.32938/slk.v2i1.794.

- Septiarini A, Hamdani H, Khoemah Y, Puspitasari N, Wati M. 2022. Segmentasi Tomat Menggunakan Metode K-Means Clustering dan Pengolahan Citra Digital. *Semin Nas CORIS 2022.*, siap terbit.
- Sulistya YI, Istighosah M, Septiara M, Dwi A. 2024. Classification of Noni Fruit Ripeness Using Support Vector Machine (SVM) Method. 5(3):206–215.
- Widiastuti NI, Rainarli E, Dewi KE. 2017. Peringkasan dan Support Vector Machine pada Klasifikasi Dokumen. *J Infotel.* 9(4):416. doi:10.20895/infotel.v9i4.312.
- Wisudawati DsT, Utami T iani W, Arum PR. 2020. Analisis Sentimen Terhadap Dampak Covid-19 Pada Performa Tokopedia Menggunakan Support Vector Machine. *Seminar Nasional Variansi.* 2:87–96. <https://ojs.unm.ac.id/variantsistatistika/article/view/19508>.
- Yanto B, Jufri J, Lubis A, Hayadi BH, Armita, NST E. 2021. Klarifikasi Kematangan Buah Nanas Dengan Ruang Warna Hue Saturation Intensity (Hsi). *INOVTEK Polbeng - Seri Inform.* 6(1):135. doi:10.35314/isi.v6i1.1882.