

Optimasi Algoritma Naïve Bayes Untuk Klasifikasi Buah Apel Berdasarkan Fitur Warna RGB

M. Afriansyah¹, Joni Saputra¹, Yuan Sa'adati², Valian Yoga Pudya Ardhana^{2,*}

¹Fakultas Sains & Teknologi, Ilmu Komputer, Universitas Qamarul Huda Badaruddin, Lombok Tengah, Indonesia

²Fakultas Sains & Teknologi, Teknologi Informasi, Universitas Qamarul Huda Badaruddin, Lombok Tengah, Indonesia

Email: ¹mafriansyah7901@gmail.com, ²saputrajoni798@gmail.com, ³yuan@uniqhba.ac.id, ^{4,*}valianypa81@gmail.com

Email Penulis Korespondensi: valianypa81@gmail.com

Abstrak—Apel adalah salah satu jenis buah yang semakin populer di Indonesia. Buah ini tidak hanya digemari karena rasanya yang enak, tetapi juga kaya akan nutrisi dan serat yang bermanfaat untuk kesehatan tubuh. Seiring dengan semakin berkembangnya industri pertanian di Indonesia, produksi apel dalam negeri pun semakin meningkat. Penelitian ini bertujuan untuk mengklasifikasikan jenis buah apel berdasarkan warna RGB dengan menggunakan metode penelitian yang dilakukan meliputi pengumpulan data citra apel, ekstraksi fitur RGB, pembagian data dengan k-fold cross validation, model klasifikasi dengan Naive Bayes. Metode ini memanfaatkan fitur warna yang diambil dari gambar buah apel sebagai input untuk menentukan kelas atau jenis buah apel yang sesuai. Hasil pengujian bahwa akurasi pada tingkat manis memiliki nilai 100%, untuk tingkat sedang memiliki nilai 86.66% dan asam memiliki nilai 80%. Rata-rata akurasi metode Naïve Bayes sebesar 88.88%. Hasil klasifikasi menggunakan algoritma Naïve Bayes.

Kata Kunci: Apel; Klasifikasi; Naïve Bayes; RGB; Fitur Ekstraksi

Abstract—Apples are one type of fruit that is increasingly popular in Indonesia. This fruit is not only popular because it tastes good, but is also rich in nutrients and fiber which are beneficial for the health of the body. Along with the development of the agricultural industry in Indonesia, domestic apple production is also increasing. This study aims to classify types of apples based on RGB color using research methods that include apple image data collection, RGB feature extraction, data division with k-fold cross validation, classification model with Naive Bayes. This method utilizes color features taken from apple images as input to determine the appropriate class or type of apple. The test results show that the accuracy for the sweet level has a value of 100%, for the medium level it has a value of 86.66% and for sour it has a value of 80%. The average accuracy of the Naïve Bayes method is 88.88%. Classification results using the Naïve Bayes algorithm.

Keywords: Apples; Classification; Naïve Bayes; RGB; Extraction Feature

1. PENDAHULUAN

Klasifikasi buah apel berdasarkan fitur warna dan RGB adalah salah satu aplikasi dari metode klasifikasi yang dapat dilakukan menggunakan algoritma Naive Bayes. Buah apel adalah salah satu buah yang sangat populer dan dikonsumsi secara luas di seluruh dunia. Oleh karena itu, penelitian tentang klasifikasi buah apel menjadi sangat penting dalam bidang pertanian dan teknologi makanan. Machine Learning merupakan cabang ilmu dari Artificial Intelligence [1]-[3] yang telah membuat kemajuan signifikan pada bidang teknologi di berbagai industri [4][5]. Model Machine Learning terbagi menjadi tiga bagian, model yang diawasi atau supervised, tidak diawasi atau unsupervised, dan gabungan atau semi-supervised [6]-[8]. Salah satu jenis klasifikasi yang menggunakan algoritma kecerdasan buatan adalah algoritma Naïve Bayes. Naïve bayes adalah algoritma klasifikasi yang populer pada penerapan machine learning [9][10] Metode Naive Bayes adalah salah satu metode klasifikasi yang sederhana dan efektif dalam memprediksi kelas dari suatu data. Metode ini didasarkan pada teori probabilitas Bayes, dimana probabilitas suatu data diklasifikasikan ke dalam suatu kelas dihitung berdasarkan probabilitas kondisional dari setiap fitur yang terkait dengan kelas tersebut. Dalam klasifikasi buah apel, fitur warna dan RGB dapat digunakan untuk membedakan antara jenis buah apel yang berbeda. Warna buah apel dipengaruhi oleh kandungan pigmen pada kulit buah. RGB (Red-Green-Blue) adalah model warna yang digunakan untuk menghasilkan berbagai warna dengan menggabungkan warna merah, hijau, dan biru dalam berbagai proporsi. Dalam penelitian ini, akan dilakukan klasifikasi buah apel berdasarkan fitur warna dan RGB menggunakan metode Naive Bayes. Tujuannya adalah untuk mengidentifikasi tingkat kemanisan apel. Metode ini diharapkan dapat membantu dalam pengenalan jenis buah apel secara cepat dan efektif, sehingga dapat membantu dalam meningkatkan kualitas dan efisiensi produksi buah apel.

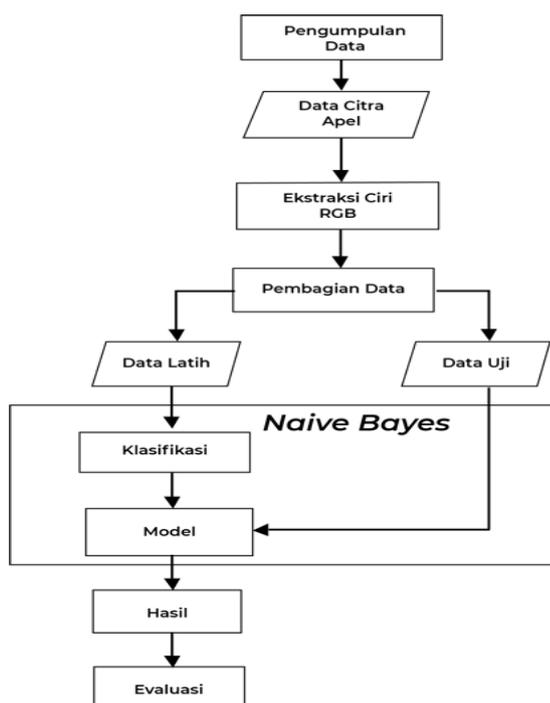
Terdapat penelitian sebelumnya yang telah melakukan sebuah penelitian diantaranya yaitu penelitian yang dilakukan oleh Roring, C. B. tentang Klasifikasi Tingkat Kematangan Buah Jambu Bol Berdasarkan Warna Kulit Menggunakan Metode Naïve Bayes, dengan klasifikasi menggunakan ekstraksi fitur RGB (Red Green Blue) dan metode Naïve Bayes, masalah kematangan buah jambu bol dapat diselesaikan. Dengan menggunakan data citra buah jambu bol sebanyak 30 citra, yang terdiri dari 10 citra buah masih mentah, 10 buah citra setengah mentah, dan 10 buah citra matang. Kemudian data diolah dan dijadikan dataset yang berupa data training dan data testing. Data training digunakan untuk proses pembuatan model dan acuan untuk data testing. Sedangkan data testing merupakan citra yang digunakan untuk menguji performa dari model yang telah dibuat. Dalam penelitian ini mendapatkan nilai akurasi mencapai 75%. [11]. Penelitian selanjutnya dilakukan oleh Syafi'i, A. M tentang Klasifikasi Kematangan Pada Buah Mangga Garifta Merah dengan Transformasi Ruang Warna HIS, klasifikasi kematangan buah mangga Garifta Merah. Selain mengklasifikasi kematangan buah mangga Garifta Merah secara manual, kami juga memakai software MATLAB R2015a yang terdapat program menggunakan metode transformasi ruang warna HSI dimana nilai RGB diubah kedalam nilai HSI. Dengan menggunakan 15 data uji diperoleh 12 data yang sesuai dengan datanya dan 3 data tidak sesuai dengan datanya.

Didapatkan hasil tingkat accuracy, precision, dan recal beturut-turut adalah 80%, 80% dan 87% [12]. Penelitian selanjutnya dilakukan oleh Yulianto Dwi tentang Klasifikasi Tahap Kematangan Pisang Ambon Berdasarkan Warna Menggunakan Naive Bayes, Dengan menggunakan 3 buah model yang terdiri dari variabel penduga yang berbeda didapatkan akurasi rata-rata terbesar yang sama yaitu dengan menggunakan model ke-2 yang mempunyai 9 nilai variabel (r, g, b, v, *a, *b, entropi, energi, dan homogenitas) dan model ke-3 yang mempunyai 7 nilai variabel (r, g, b, v, *a, entropi dan homogenitas) yaitu sebesar 90.48%. [13]. Penelitian selanjutnya dilakukan oleh Saputra yang berjudul Klasifikasi Pengenalan Buah Menggunakan Algoritma Naive Baiyes, Metodologi perangkat lunak ini melibatkan akuisisi gambar, pra-pemrosesan dan segmentasi, analisis dan klasifikasi varietas apel. Prototipe sistem klasifikasi Apple dibangun menggunakan lingkungan platform pengembangan MATLAB R2017. Hasil dalam penelitian menunjukkan bahwa nilai rata-rata estimasi akurasi, sensitivitas, presisi dan spesifisitas masing-masing adalah 81%, 73%, 100% dan 70%. MLP-Neural menunjukkan bahwa kinerja teknik Naive Bayes konsisten dengan analisis Komponen Utama, Fuzzy Logic dan MLP-Neural dengan 91%, 90%, 89 %, dan 83% masing-masing dalam hal akurasi [14]. Penelitian selanjutnya dilakukan oleh Mahran yang berjudul Penerapan Naive Bayes Gaussian Pada Klasifikasi Jenis Jamur Berdasarkan Ciri Statistik Orde Pertama, dilakukan proses klasifikasi jenis jamur dengan jumlah citra sebesar 240 citra menggunakan metode naive bayes gaussian dan metode feature extraction ciri statistic orde pertama (mean, variance, skewness, kurtosis, entropy) menggunakan metode cross validation dengan nilai k=60 didapatkan nilai akurasi sebesar 98.75% [15]. Penelitian selanjutnya dilakukan oleh Mustofa tentang Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes, Proses klasifikasi hoax dapat dilakukan melalui tahap preprocessing kemudian pembobotan kata dan dilakukan klasifikasi menggunakan naive bayes. Pengukuran dilakukan dengan metode 10-fold cross validation. Pengukuran dilakukan dengan metode 10-fold cross validation. Dari pengukuran tersebut diperoleh hasil, nilai fold 6 mempunyai keakuratan tertinggi, yaitu sebesar 85.28 % yang mana dokumen terklasifikasi yang relevan sebanyak 307 dan dokumen tidak relevan sebanyak 53 atau error rate sebesar 14.72% [16]. Penelitian selanjutnya dilakukan oleh Fauzi tentang Klasifikasi Jenis Tanaman Tembakau di Indonesia menggunakan Naive Bayes dengan Seleksi FiturInformation Gain, memudahkan menentukan tipe tanaman tembakau dibutuhkan sebuah sistem dengan metode klasifikasi. Salah satu metode klasifikasi yaitu dengan algoritme Naive Bayes. K-fold cross validation untuk menghilangkan bias pada datadengan k=10. Dari seluruh pengujian yang dilakukan akurasi rata-rata tertinggi untuk seluruh pengujian kelas seimbang yaitu 52,72% dengan menggunakan 17 fitur. Sedangkan untuk akurasi tertinggi dari seluruh pengujian kelas tidak seimbang adalah 64,06% ketika menggunakan 15 fitur [17]

2. METODOLOGI PENELITIAN

2.1 Metode Penelitian

Metode penelitian yang dilakukan meliputi pengumpulan data citra apel, ekstraksi fitur RGB, pembagian data dengan k-fold cross validation, model klasifikasi dengan Naive Bayes, dan evaluasi model dengan melihat hasil klasifikasi menggunakan confusion matrix.



Gambar 1. Tahapan Penelitian

2.2 Pengumpulan Data

Data yang digunakan adalah citra apel yang terdiri dari apel manis, sedang dan asam. Jumlah citra Apel yang diperoleh berjumlah 150 citra Apel. Apel adalah salah satu buah-buahan yang sangat populer dan dikonsumsi secara luas di seluruh dunia. Apel memiliki bentuk bulat atau lonjong dengan kulit yang halus atau kasar tergantung pada jenisnya. Buah apel memiliki banyak manfaat kesehatan, seperti meningkatkan sistem kekebalan tubuh, mencegah penyakit jantung, dan mengurangi risiko terkena kanker. Selain itu, apel juga mengandung banyak vitamin, serat, dan antioksidan.

2.3 Ekstraksi Ciri

Untuk mengetahui suatu citra, diperlukan adanya ekstraksi ciri. Ekstraksi ciri dapat diteliti dengan mengambil beberapa bagian citra yang bisa menunjukkan ciri khas dari citra tersebut, misalkan warna, pola citra, diameter, bentuk dan masih banyak lagi. Citra warna merupakan citra digital yang memiliki kombinasi warna Red, Green dan Blue. Teknik yang digunakan adalah dengan mengekstrak citra RGB (Red-Green-Blue) buah Apel menjadi beberapa nilai ciri, seperti jumlah R, jumlah G, jumlah B, mean dan standar deviasi dari citra buah Apel. Untuk satu buah, nilai ciri tersebut diperoleh dengan merata-ratakan atau menjumlahkan semua piksel yang ada, dan berdasar nilai inilah dilakukan pengenalan. Jadi dari teknik yang ada, proses pengenalan tidak berdasarkan semua piksel tetapi berdasarkan besaran yang merupakan rata-rata atau jumlah dari semua piksel [18]. RGB adalah warna dasar yang dijadikan patokan warna secara universal (primary colors) dengan citra jenis berwarna yang bisa diubah warnanya ke dalam kode-kode angka sehingga warna tersebut akan tampil universal [19].

2.4 Pembagian Data

Seluruh data hasil ekstraksi masing-masing ciri dibagi menjadi data latih dan data uji. Persentase data latih yang dicobakan pada penelitian ini ditentukan dengan menggunakan k-fold cross validation.

K-fold cross validation adalah metode yang digunakan untuk membagi dataset menjadi sejumlah k buah partisi secara acak. Data awal dibagi menjadi k subset secara acak yaitu D1, D2, D3, ... , Dk, dengan ukuran subset yang hampir sama dengan mempertahankan perbandingan antar kelas. Langkahnya adalah dengan melakukan iterasi sejumlah k kali iterasi untuk data latih dan data uji, dimana masing-masing iterasi menggunakan partisi ke k sebagai data latih dan sisa partisi lainnya sebagai data uji. Keuntungan k-fold cross validation adalah semua data digunakan baik untuk data uji maupun data latih. Hal ini dilakukan untuk mendapatkan nilai akurasi ataupun ukuran penilaian lainnya dari hasil eksperimen yang dilakukan [20].

Jumlah k yang digunakan adalah 3. Mengingat data yang digunakan untuk pelatihan sedikit, pemilihan 3-fold cross validation cukup mampu untuk membuat variasi data, sehingga semua data digunakan, baik untuk data uji maupun data latih. Seluruh data hasil ekstraksi ciri dibagi menjadi 3 subset, yaitu D1, D2, D3. Masing-masing subset memiliki ukuran yang sama. Pada proses pertama D2, D3 menjadi data pelatihan dan D1 menjadi data pengujian, pada proses kedua D1, D3, menjadi data pelatihan dan D2 menjadi data pengujian, dan seterusnya.

2.5 Pembagian Data Penerapan Metode Naive Bayes

Naive Bayes merupakan salah satu metode yang dapat digunakan untuk mengklasifikasikan data. Bayesian classification merupakan pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class [21]. Teorema ini dikemukakan oleh ilmuwan Inggris, Thomas Bayes, dengan memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya [22]. Teorema Bayes dikombinasikan dengan "Naive" yang berarti setiap atribut/variable bersifat bebas (independent) [23]. Sedangkan penjelasan frekuentis, model ini mengartikan bahwa representasi sebagai invers probabilitas melalui dua scenario kejadian [24].

Naive Bayes merupakan salah satu metode machine learning yang menggunakan perhitungan probabilitas. Algoritma ini memanfaatkan metode probabilitas dan statistic sederhana dengan asumsi bahwa antar satu kelas dengan kelas yang lain tidak saling tergantung (independen) [25]. Dasar dari Naive Bayes yang dipakai dalam pemrograman adalah persamaan (1) Bayes [26] :

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^q (P(X_i|Y))}{P(X)} \quad (1)$$

Keterangan:

$P(Y|X)$ = probabilitas data dengan vector X pada kelas Y

$P(Y)$ = probabilitas awal kelas Y (prior probability)

$\prod_{i=1}^q P(X_i|Y)$ = probabilitas independen kelas Y dari semua fitur dalam vektor X

$P(X)$ = probabilitas dari X

Data yang digunakan dapat bersifat kategorial maupun kontinyu. Untuk data kontinyu dapat diselesaikan dengan menggunakan langkah-langkah berikut.

1. Hitung probabilitas (Prior) tiap kelas yang ada.
2. Lalu hitung rata-rata (mean) tiap fitur dengan persamaan (2).

$$\mu = \frac{\sum^n}{k} \quad (2)$$

Keterangan:

k = Banyaknya data

n = nilai data

3. Kemudian hitung nilai standar deviasi dari fitur tersebut seperti pada persamaan (3) .

$$sd = \sqrt{\frac{n \sum_{i=1}^n (xi-\bar{x})^2}{(n-1)}} \tag{3}$$

Selanjutnya menghitung densitas probabilitasnya menggunakan persamaan (4).

$$\varphi_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4}$$

Setelah didapatkan nilai densitas probabilitas dan prior, hitung probabilitas masing-masing kelas dengan menggunakan persamaan (5).

$$P(X|Ci) \times P(Ci) \tag{5}$$

Nilai probabilitas terbesar adalah kelas yang sesuai. Sedangkan untuk data kategorial, hanya memerlukan semua kemungkinan yang terjadi.

Naïve Bayes adalah algoritma yang termasuk ke dalam supervised learning, maka akan dibutuhkan pengetahuan awal untuk dapat mengambil keputusan. Proses klasifikasi dengan Naïve Bayes dilakukan menggunakan data latih yang sebelumnya sudah dibagi menggunakan k- fold cross validation. Dalam melakukan pelatihan dan pengujian data, karakter akan diambil satu per satu dari fitur yang ada. Klasifikasi yang akan dilakukan adalah didasarkan ciri-ciri warna buah apel yang menunjukkan buah apel tersebut manis, sedang atau asam. Terdapat 2 proses dalam klasifikasi ini yaitu: tahap training dengan memakai data yang ada, membangun metode untuk mengestimasi parameter dari distribusi peluangnya dengan asumsi bahwa adanya independensi dari masing-masing kelas (data dengan karakteristik yang sama). Dalam tahapan ini dilakukan estimasi pada parameter θ dengan Maximum Likelihood (ML), dan tahap prediksi yaitu proses menggunakan model yang sudah dibangun tersebut untuk melakukan tes data untuk memperkirakan/mengukur akurasi dari aturan yang dibentuk dalam model dengan menghitung peluang posterior kemudian mengklasifikasi kedalam peluang posterior terbesar MAPH (Maximum A Posteriori Hypothesis).

2.6 Evaluasi

Tahapan ini merupakan tahapan untuk menganalisis dan mengevaluasi model yang diperoleh dari masing model yang digunakan. Proses perhitungan akurasi hasil klasifikasi menggunakan rumus Confusion matrix. Confusion matrix merupakan sebuah tabel yang terdiri atas banyaknya baris data uji yang diprediksi benar dan tidak benar oleh model klasifikasi. Tabel confusion matrix diperlukan untuk menentukan kinerja suatu model klasifikasi [27]. Ada empat istilah yang digunakan dalam confusion matrix yaitu:

1. True positive (TP): jumlah data positif yang benar diklasifikasi oleh classifier.
2. True negative (TN): jumlah data negatif yang benar diklasifikasi oleh classifier.
3. False positive (FP): jumlah data negatif yang salah diklasifikasi sebagai data positif.
4. False negative (FN): jumlah data positif yang salah diklasifikasi sebagai data negatif.

Contoh tabel confusion matrix prediksi kelas dapat dilihat pada Tabel 1.

Tabel 1. Confusion matrix Prediksi Kelas

Kelas Sebenarnya	Positive	Negative
Positive	A : True Positive	B : False Negative
Negative	C : False Positive	D : True Negative

Berdasarkan tabel confusion matrik di atas akurasi dapat dihitung dengan persamaan (6)

$$Akurasi = \frac{\sum A + \sum D}{\sum A + \sum B + \sum C + \sum D} \times 100\% \tag{6}$$

2.7 Klasifikasi

Klasifikasi adalah suatu proses yang bertujuan untuk menentukan suatu objek kedalam suatu kelas atau kategori yang sudah ditentukan sebelumnya. Klasifikasi juga didefinisikan sebagai proses dari pembangunan terhadap suatu model yang mengklasifikasikan suatu objek sesuai dengan atribut-atributnya [28]. Ada 2 proses dalam klasifikasi, yaitu Proses learning/training Melakukan pembangunan model menggunakan data training [29]

3. HASIL DAN PEMBAHASAN

Pengujian dilakukan dengan tujuan untuk mengetahui keberhasilan sistem dalam mengklasifikasi rasa dari buah apel, apakah manis, sedang dan asam. Skenario uji coba yang dilakukan dengan jumlah data gambar dari buah apel, yaitu sebanyak 150 gambar. Masing-masing jumlah gambar dari tingkat kemanisan buah apel seperti terlihat pada Tabel 2.

Tabel 2. Jumlah Data

No	Tingkat Apel	Jumlah
1	Manis	50
2	Sedang	50
3	Asam	50
	Jumlah Citra	150

Pengujian dilakukan dengan cara memisahkan data menjadi dua bagian yaitu 80% digunakan sebagai data latih dan 20% digunakan sebagai data uji. Beberapa sampel citra yang telah terkumpul seperti yang ditunjukkan pada Tabel 3.

Tabel 3. Citra Buah Apel

Citra Buah Apel	Tingkat Kematangan Apel
	Manis (a)
	Sedang (b)
	Asam (c)

3.1 Hasil Ekstraksi

Praproses dilakukan dengan terlebih dahulu melakukan segmentasi. Segmentasi sangat diperlukan untuk menentukan piksel yang akan dianalisis. Pada tahap ekstraksi fitur dilakukan ekstraksi fitur yang berisi 5 nilai yaitu dari fitur warna R,G,B, mean dan fitur standar deviasi dari citra apel. Untuk satu buah apel, nilai ciri tersebut diperoleh dengan merata-ratakan atau menjumlahkan semua piksel yang ada, dan berdasar nilai inilah dilakukan pengenalan. Berikut nilai hasil ekstraksi fitur dapat dilihat pada Tabel 4.

Tabel 4. Ekstraksi Fitur Citra Apel RGB

Citra Apel	Warna	Tingkat Kematangan
Apel1	R = 191.42	Manis
	G = 182.03	Manis
	B = 160.13	Manis
Apel2	R = 187.79	Manis
	G = 161.53	Manis
	B = 125.34	Manis
Apel3	R = 205.08	Manis
	G = 188.52	Manis
	B = 155.08	Manis
Apel4	R = 197.77	Sedang
	G = 195.01	Sedang
	B = 160.35	Sedang
Apel5	R = 196.19	Sedang
	G = 188.86	Sedang
	B = 159.38	Sedang
Apel6	R = 182.24	Sedang
	G = 182.11	Sedang
	B = 145.12	Sedang
Apel7	R = 170.52	Asam
	G = 187.25	Asam
	B = 143.41	Asam

Citra Apel	Warna	Tingkat Kematangan
Apel8	R = 168.77	Asam
	G = 181.63	Asam
	B = 145.42	Asam
Apel9	R = 169.86	Asam
	G = 181.22	Asam
	B = 145.08	Asam

Berikut nilai hasil ekstraksi fitur Mean dan Standar Deviasi dapat dilihat pada Tabel 5.

Tabel 5. Ekstraksi Fitur Citra Apel Mean dan STD

Citra Apel	Mean	STD
Apel1	177.86	16.05
Apel2	158.22	31.35
Apel3	182.89	25.47
Apel4	184.37	20.85
Apel5	181.47	19.48
Apel6	169.82	21.39
Apel7	167.06	22.12
Apel8	165.27	18.35
Apel9	165.38	18.48

3.2 Hasil Klasifikasi

Proses training bertujuan untuk membangun model klasifikasi. Berdasarkan Tabel 6 diketahui bahwa akurasi terbesar yang diperoleh dari tiga kali percobaan terdapat pada fold 1 sebesar 85 %, akan tetapi untuk menghindari terjadinya overfitting, maka dihitung rata-rata dari model 3-fold cross validation. Kemudian hasil rata-rata dibandingkan dengan ketiga model tersebut. Fold 3 dipilih sebagai model Naive Bayes karena nilai fold 3 mendekati nilai rata-rata yang diperoleh.

Tabel 6. Akurasi Klasifikasi Tingkat Kemanisan Buah Apel

Fold	Akurasi
1	85%
2	70%
3	75%
Rata-rata	76.66%

Pengujian hasil klasifikasi dilakukan terhadap 40 citra buah apel yang diperoleh dari perbandingan data latih dan data uji sebesar 80% dan 20%. Data uji akan diklasifikasikan terhadap data training.

Tabel 7. Confusion Matrix Klasifikasi Buah Apel

		Class Prediction			Total
		Manis	Sedang	Asam	
Real Prediction	Manis	15	0	0	15
	Sedang	0	13	2	15
	Asam	0	2	8	10

Berdasarkan Tabel 7 diketahui metode Naive Bayes untuk tingkat kemanisan sedang dan asam masing mempunyai kesalahan klasifikasi, hal ini terjadi karena fitur warna buah apel yang manis berbeda warna dengan ciri warna buah apel yang memiliki tingkat kemanisan sedang dan asam

3.3 Evaluasi

Hasil klasifikasi dari tingkat kemanisan buah belimbing tersebut kemudian dihitung nilai akurasinya, dengan cara membagi jumlah dari data uji setiap kelas yang diklasifikasikan secara benar dengan total dari data uji. Berdasarkan hasil klasifikasi yang terlihat pada Tabel 5. diperoleh nilai akurasi untuk tiap tingkat kemanisan buah belimbing seperti pada Tabel 6.

Tabel 8. Akurasi Hasil Klasifikasi Buah Apel

Kelas	Akurasi
Manis	100%
Sedang	86.66%
Asam	80%
Rata-rata	88.88%

Berdasarkan Tabel 8 diketahui bahwa akurasi pada tingkat manis memiliki nilai 100%, untuk tingkat sedang memiliki nilai 86.66% dan asam memiliki nilai 80%. Rata-rata akurasi metode Naïve Bayes sebesar 88.88%. Hasil klasifikasi menggunakan algoritma Naïve Bayes. Oleh karena itu, dibutuhkan ekstraksi ciri tambahan untuk mendukung model klasifikasi yang akan dibentuk.

4. KESIMPULAN

Ekstraksi fitur warna RGB dapat digunakan untuk ekstraksi ciri pada citra. Hasil ekstraksi ciri digunakan sebagai input bagi Naïve Bayes untuk mengenal pola citra dan mengklasifikasikan tingkat kemanisan buah apel. Hasil klasifikasi menunjukkan tingkat akurasi menggunakan ekstraksi ciri warna RGB sebesar 88.88%. Keberhasilan mengidentifikasi juga dipengaruhi oleh fitur-fitur yang digunakan sebagai penciri pada metode klasifikasi tersebut.

REFERENCES

- [1] A. Oikonomidis, C. Catal, and A. Kassahun, "Deep learning for crop yield prediction: a systematic literature review," *New Zeal. J. Crop Hortic. Sci.*, pp. 1–26, 2022.
- [2] V.Y.P. Ardhana, J. Saputra, and M. Afriansyah, "Klasifikasi Jenis Mangga Berdasarkan Tekstur Tulang Daun Menggunakan Metode Learning Vector Quantization (LVQ)," *Journal of Computer System and Informatics (JoSYC)*, vol. 4, no. 1, pp. 220-228, 2022.
- [3] M. Rifai *et al.*, "Dynamic time distribution system monitoring on traffic light using image processing and convolutional neural network method," *IOP Conf. Series: Mater. Sci. Eng.*, vol. 1175, no. 1, p. 012005, Aug. 2021.
- [4] Y. Xu, Y. Zhou, P. Sekula, and L. Ding, "Machine learning in construction: From shallow to deep learning," *Dev. Built Environ.*, vol. 6, p. 100045, 2021.
- [5] Tommy Wijaya Sagala, Eliyah Acantha Manapa, Valian Yoga Pudya Ardhana, Godfried, "Perbandingan Implementasi Manajemen Pengetahuan pada," *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 1, no. 4, pp. 327-335, 2020.
- [6] Nurul Salsabila Syam, et al, "Model Support Vector Machine untuk Prediksi pada Penggunaan Energi Listrik di Rumah Hemat Energi", *Jl*, vol. 1, no. 2, pp. 56-59, Sep. 2022.
- [7] M. Bibi *et al.*, "Class association and attribute relevancy based imputation algorithm to reduce twitter data for optimal sentiment analysis," *IEEE Access*, vol. 7, pp. 136535–136544, 2019
- [8] Valian Yoga Pudya Ardhana, Muh. Yusuf Syam, Eka Fitri Ramadani, Eliyah A M Sampetoding, Mohammad Syahril, Esther Sanda Manapa, Rahmat Mardzuki, "Prediksi Flight Delay Berbasis Algoritma Neural Network," *Journal of Informatics, Electrical and Electronics Engineering*, vol. 2, no. 1, pp. 26-30, 2022.
- [9] W. M. Shaban, A. H. Rabie, A. I. Saleh, and M. A. Abo-Elhoud, "Accurate detection of COVID-19 patients based on distance biased Naïve Bayes (DBNB) classification strategy," *Pattern Recognit.*, vol. 119, p. 108110, 2021
- [10] A. D. Wiratmoko *et al.*, "Design of Potholes Detection as Road's Feasibility Data Information Using Convolutional Neural Network(CNN)," in *2019 Int. Symp. Electron. Smart Devices (ISESD)*, Badung-Bali, Indonesia, Oct. 8–9, 2019. IEEE, 2019.
- [11] Roring, C. B., Mulyana, D. I., Lubis, Y. T., & Zamzami, A. R. (2022). Klasifikasi Tingkat Kematangan Buah Jambu Bol Berdasarkan Warna Kulit Menggunakan Metode Naïve Bayes. *Jurnal Pendidikan Tambusai*, 6(1), 2938-2948.
- [12] Syafi'i, A. M., Ahadi, M. F., Rasyid, M. I., Adhinata, F. D., & Junaidi, A. (2021). Mendeteksi Kematangan Pada Buah Mangga Garifta Merah Dengan Transformasi Ruang Warna HSI. *Journal of Applied Informatics and Computing*, 5(2), 117-121.
- [13] Yulianto, D., Whidhiah, R. N., & Maimunah, M. (2017). Klasifikasi Tahap Kematangan Pisang Ambon Berdasarkan Warna Menggunakan Naïve Bayes. *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic*, 5(2), 60-67.
- [14] Saputra, A. (2019). KLASIFIKASI PENGENALAN BUAH MENGGUNAKAN ALGORITMA NAIVE BAIYES. *Jurnal RESISTOR (Rekayasa Sistem Komputer)*, 2(2), 83-88.
- [15] Mahran, A. A., Hapsari, R. K., & Nugroho, H. (2020). Penerapan Naive Bayes Gaussian Pada Klasifikasi Jenis Jamur Berdasarkan Ciri Statistik Orde Pertama. *Network Engineering Research Operation*, 5(2), 91-99.
- [16] Mustofa, H., & Mahfudh, A. A. (2019). Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes. *Walisongo Journal of Information Technology*, 1(1), 1-12.
- [17] Fauzi, F. A., Furqon, M. T., & Yudistira, N. (2021). Klasifikasi jenis tanaman tembakau di Indonesia menggunakan Naïve Bayes dengan seleksi fitur information Gain. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2548, 964X.
- [18] Saputra K dan Manik F.Y, 2016, Klasifikasi Belimbing Menggunakan K-Nearest Neighbors (KNN) Berdasarkan Citra Red-Green-Blue (RGB), *Prosiding SEMMAU 2016*, Kupang, 17 September
- [19] Riska, S. Y. Klasifikasi Level Kematangan Tomat Berdasarkan Perbedaan Perbaikan Citra Menggunakan Rata-Rata RGB Dan Index Pixel. (2015b). *Jurnal Ilmiah Teknologi Dan Informasia ASIA (JITIKA)* Retrieved from <http://lp3m.asia.ac.id/wpcontent/uploads/2015/11/7.-Bu-Riska.pdf>.
- [20] Han J, Kamber M, Pei J, 2012, *Data Mining: Concepts and Techniques*. 3th ed, New York (US): Morgan Kaufman Elsevier Academic Pr.
- [21] R. I. Borman Dan M. Wati, "Penerapan Data Maining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naïve Bayes," *Jurnal Ilmiah Fakultas Ilmu Komputer*, Vol. 9, No. 1, Pp. 25-34, 2020.
- [22] B. S. Gandhi, D. A. Megawaty Dan D. Alita, "Aplikasi Monitoring Dan Penentuan Peringkat Kelas Menggunakan Naïve Bayes Classifier," *Jurnal Informatika Dan Rekayasa Perangkat Lunak (Jatika)*, Vol. 2, No. 1, Pp. 54-63, 2020
- [23] S. E. Y. Putri Dan A. Surahman, "Penerapan Model Naive Bayes Untuk Memprediksi Potensi Pendaftaran Siswa Di Smk Taman Siswa Teluk Betung Berbasis Web," *Jurnal Informatika Dan Rekayasa Perangkat Lunak (Jatika)*, Vol. 1, No. 1, Pp. 93-99, 2020.
- [24] Kusumadewi, S., 2003, *Artificial Intelligence Teknik dan Aplikasi*, Graha Ilmu, Yogyakarta.
- [25] Nugroho A dan Subahar, 2013, *Klasifikasi Naive Bayes Untuk Prediksi Kelahiran Pada Data Ibu Hamil*. *Berkala Mipa*, Vol 23, Ed 3.
- [26] Prasetyo E, 2012, *Data Mining : Konsep dan Aplikasi Menggunakan Matlab*, Penerbit Andi, Yogyakarta

- [27] Tan PN, Steinbach M, Kumar V, 2005, Introduction to data mining, New York (US): Addison Wesley
- [28] C. Paramita, E. Hari Rachmawanto, C. Atika Sari, and D. R. Ignatius Moses Setiadi, “Klasifikasi Jeruk Nipis Terhadap Tingkat Kematangan Buah Berdasarkan Fitur Warna Menggunakan K-Nearest Neighbor,” *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 4, no. 1, pp. 1–6, 2019, doi: 10.30591/jpit.v4i1.1267.
- [29] S. R. Raysyah, Veri Arinal, and Dadang Iskandar Mulyana, “Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode Knn Dan Pca,” *JSiI (Jurnal Sist. Informasi)*, vol. 8, no. 2, pp. 88–95, 2021, doi: 10.30656/jsii.v8i2.3638.