

Klasifikasi Kematangan Manggis Berdasarkan Fitur Warna dan Tekstur Menggunakan Algoritma Naive Bayes

Raihan Abimanyu Suharman, Hartono *

Departemen Pendidikan Matematika, FMIPA Universitas Negeri Yogyakarta

* Corresponding Author. E-mail: hartono@uny.ac.id

ARTICLE INFO	ABSTRACT
<p>Article History: Received: 03-Oct. 2022 Revised: 14-Nov. 2022 Accepted: 14-Dec.2022</p> <p>Keywords: Klasifikasi, kematangan manggis, Naive Bayes, MRMR</p>	<p>Manggis merupakan salah satu komoditas buah asli Indonesia yang memiliki prospek pasar yang menjanjikan, terlebih dalam pasar ekspor. Namun masih ada permasalahan dalam hal penyortiran buah hasil panen. Buah manggis hasil panen disortir berdasarkan kematangan buahnya untuk tujuan pasar yaitu pasar ekspor dan pasar domestik. Faktor penentu kematangan buah manggis adalah warna dan tekstur dari kulit buahnya. Penelitian ini bertujuan untuk mengklasifikasi kematangan buah manggis berdasarkan warna dan tekstur menggunakan algoritma Naive Bayes. Fitur warna dan tekstur yang diekstraksi adalah kontras, korelasi, energi, homogenitas, entropi, standar deviasi, mean, varians, skewness, dan kurtosis. Fitur diekstraksi dari citra RGB, citra grayscale, citra HSV, dan citra CIELAB. Hasil ekstraksi fitur melewati tahap seleksi fitur menggunakan algoritma Minimum Redundancy Maximum Relevance (MRMR). Metode klasifikasi yang digunakan adalah metode Naive Bayes. Model klasifikasi Naive Bayes menggunakan parameter sebanyak 13 fitur dalam pembangunan modelnya yaitu mean R, mean G, standar deviasi G, mean Saturation, mean Hue, standar deviasi Hue, standar deviasi Value, mean a*, mean b*, standar deviasi a*, standar deviasi b*, varians a*, dan kontras. Hasil klasifikasi kematangan buah manggis menggunakan algoritma Naive Bayes memperoleh tingkat akurasi sebesar 95,7% dengan sensitivitas, spesifisitas, dan presisi untuk kelas matang sebesar 93,3%, 96,8%, dan 93,3%. Sensitivitas, spesifisitas, dan presisi untuk kelas mentah masing-masing sebesar 100%. Sensitivitas, dan presisi untuk kelas setengah matang sebesar 93,3%, 96,9%, dan 93,3%.</p>
	<p><i>Mangosteen is one of Indonesia's original fruit commodities that has promising market prospects, especially in the export market. However, there are still problems in sorting the harvested fruit. Harvested mangosteen fruit is sorted based on the maturity of the fruit for market purposes, namely the export market and the domestic market. The determining factor for the maturity of the mangosteen fruit is the color and texture of the fruit's skin. This study aims to classify mangosteen ripeness based on color and texture using the Naive Bayes algorithm. The extracted color and texture features are contrast, correlation, energy, homogeneity, entropy, standard deviation, mean, variance, skewness, and kurtosis. Features are extracted from RGB images, grayscale images, HSV images, and CIELAB images. The result of feature extraction passes through the feature selection stage using the Minimum Redundancy Maximum Relevance (MRMR) algorithm. The classification method used is the Naive Bayes method. The Naive Bayes classification model uses a parameter of 13 features in the construction of the model, namely mean R, mean G, standard deviation G, mean Saturation, mean Hue, standard deviation Hue, standard deviation Value, mean a*, mean b*, standard deviation a*, standard deviation b*, variance a*, and contrast. The results of the classification of mangosteen fruit maturity using the Naive Bayes algorithm obtained an accuracy rate of 95.7% with sensitivity, specificity, and precision for the ripe class are 93.3%, 96.8%, and 93.3%. The sensitivity, specificity, and precision for the raw class are 100% each. The sensitivity, specificity, and precision for the half-ripe class are 93.3%, 96.9%, and 93.3%.</i></p>

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



How to Cite:

Suharman, R. A. & Hartono, H. (2022). Klasifikasi kematangan manggis berdasarkan fitur warna dan tekstur menggunakan algoritma naive bayes. *Pythagoras: Jurnal Matematika dan Pendidikan Matematika*, 17(2), 377-388. <https://doi.org/10.21831/pythagoras.v17i2.53625>

 <https://doi.org/10.21831/pythagoras.v17i2.53625>

PENDAHULUAN

Sektor pertanian merupakan salah satu fokus utama pembangunan nasional dengan dampak signifikan pada ekonomi dan kesejahteraan masyarakat Indonesia, terutama dalam pengelolaan hasil pertanian, khususnya komoditas pangan. Pertanian memiliki peran penting dalam mendukung perekonomian nasional, termasuk ketahanan pangan, peningkatan daya saing, penyerapan tenaga kerja, dan mengurangi kemiskinan. Manggis (*Garcinia mangostana* L.) adalah salah satu buah tropis yang berpotensi besar dalam industri pertanian Indonesia. Dikenal sebagai "Queen of Fruit" atau ratu buah, manggis memiliki citarasa yang lezat dan sejumlah manfaat kesehatan yang signifikan, termasuk sebagai anti-inflamasi, antibakteri, obat infeksi dan luka (Widiastuti et al., 2013), antioksidan (Azima et al., 2017), dan anti-kanker (Li et al., 2013). Selain itu, produksi manggis di Indonesia telah mengalami peningkatan yang signifikan selama beberapa tahun terakhir, menunjukkan potensi pasar yang kuat, terutama di sektor ekspor. Data produksi buah manggis dari tahun 2017 hingga 2021, sebagaimana yang dipublikasikan oleh Badan Pusat Statistik (BPS), menunjukkan peningkatan yang signifikan, dengan jumlah produksi berturut-turut adalah 161,750 ton, 228,155 ton, 246,476 ton, 322,414 ton, dan 303,934 ton (Badan Pusat Statistik, 2021). Meskipun demikian, dalam konteks produksi buah di Indonesia, terdapat permasalahan yang signifikan terkait penyortiran buah setelah panen (Manik & Saragih, 2017), khususnya dalam hal analisis warna kulit buah secara visual oleh manusia, yang jelas memiliki keterbatasan. Metode penyortiran ini memerlukan tenaga kerja yang cukup besar, dan tingkat konsistensi dalam menilai tingkat kematangan buah tidak selalu terjamin karena manusia dapat mengalami kelelahan (Sugianto & Wibowo, 2015). Oleh karena itu, mengingat urgensi permasalahan ini, penelitian lebih lanjut diperlukan sebagai langkah menuju pengembangan sistem klasifikasi yang dapat memberikan akurasi tinggi dalam menentukan tingkat kematangan buah manggis. Hal ini diharapkan dapat mengoptimalkan proses penyortiran hasil panen buah manggis.

Permasalahan klasifikasi dapat diselesaikan dengan menggunakan *machine learning*. Dalam *machine learning*, terdapat beberapa metode yang dapat digunakan untuk mengklasifikasi, diantaranya yaitu *Neural Network* (NN), *Naive Bayes* (NB), Jaringan Syaraf Tiruan (JST), *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (KNN). Pada penelitian ini, metode yang digunakan untuk mengklasifikasi kematangan buah manggis adalah *Naive Bayes*. Pemilihan metode ini berdasarkan pada pertimbangan bahwa metode *Naive Bayes* mampu mengklasifikasi dengan cepat, mudah ditafsirkan, bekerja baik walaupun data latih berjumlah sedikit, serta mampu mengklasifikasi permasalahan multi-kelas (Ray, 2019). Kekurangan dari metode ini adalah parameter yang digunakan dalam model klasifikasi harus berdasar pada asumsi bahwa fitur saling bebas, sehingga perlu dilakukan pemilihan parameter yang sesuai untuk mendapatkan hasil klasifikasi yang maksimal. Hal tersebut dapat diatasi dengan melakukan seleksi fitur menggunakan algoritma *Minimum Redundancy Maximum Relevance* (MRMR) untuk mengetahui fitur yang berpengaruh dalam penentuan kelas atau label.

Penelitian terkait penyelesaian permasalahan klasifikasi menggunakan metode *Naive Bayes* pernah dilakukan oleh Kusuma dan Putra (2018) untuk mengklasifikasi kematangan tomat. Pada penelitian tersebut digunakan fitur histogram sebagai parameter model klasifikasi dan memperoleh tingkat akurasi sebesar 76%. Khisanudin dan Murinto (2020) mengklasifikasi kematangan buah naga dengan menggunakan metode *Naive Bayes*. Pada penelitian tersebut digunakan fitur citra HSV sebagai parameter model klasifikasi dan memperoleh tingkat akurasi sebesar 86,6%. Liantoni dan Nugroho (2015) melakukan penelitian mengenai klasifikasi daun herbal menggunakan metode *Naive Bayes* dan *K-Nearest Neighbor* (KNN). Penelitian tersebut memperoleh tingkat akurasi menggunakan metode *Naive Bayes* sebesar 75%, sedangkan tingkat akurasi menggunakan metode *K-Nearest Neighbor* (KNN) sebesar 70,83%. Penelitian yang dilakukan oleh Vembandasamy et al. (2015) mengenai pendeteksian penyakit jantung dengan menggunakan metode *Naive Bayes* memperoleh tingkat akurasi sebesar 86,4%. Penelitian yang dilakukan oleh Sari et al. (2020) mengenai klasifikasi jenis daun pepaya berdasarkan ekstraksi fitur LBP menggunakan metode *Naive Bayes* memperoleh tingkat akurasi sebesar 96%.

Whidhiasih et al. (2012) melakukan penelitian mengenai pemutuan buah manggis ekspor dan lokal menggunakan *Fuzzy Neural Network* (FNN). Pada penelitian tersebut digunakan fitur dari citra RGB, citra *grayscale*, citra HSV, citra CIELAB, dan citra CEILUV sebagai parameter model klasifikasi dan memperoleh tingkat akurasi sebesar 90%. Susanto et al. (2020) mengklasifikasi mutu bawang merah menggunakan metode *Naive Bayes*. Pada penelitian tersebut digunakan fitur citra HSV sebagai parameter model klasifikasi, serta digunakan pengolahan citra dengan proses segmentasi, proses morfologi, dan proses pelabelan objek, dan memperoleh tingkat akurasi sebesar 91,67%. Sahagun (2020) melakukan penelitian mengenai determinasi tingkat kemanisan buah kersen berdasarkan

model warna HSV. Pada penelitian tersebut, pengolahan citra yang dilakukan meliputi konversi citra, segmentasi (*thresholding*), dan *masking*. Pada penelitian tersebut, hasil klasifikasi buah kersen dengan menggunakan metode *Naive Bayes* memperoleh tingkat akurasi sebesar 83%, sedangkan dengan menggunakan metode SVM memperoleh tingkat akurasi sebesar 87%.

Berdasarkan uraian permasalahan di atas, tujuan penelitian ini adalah untuk mengklasifikasi kematangan buah manggis berdasarkan fitur warna dan tekstur dari kulit buah manggis menggunakan algoritma *Naive Bayes*. Perbedaan penelitian ini dengan penelitian sebelumnya terletak pada pemilihan parameter model klasifikasi dan proses pengolahan citra. Fitur warna dan tekstur yang diekstraksi diantaranya yaitu kontras, korelasi, energi, homogenitas, entropi, standar deviasi, rerata, varians, *skewness*, dan kurtosis. Fitur diekstraksi dari citra RGB, citra *grayscale*, citra HSV, dan citra CIELAB. Pada penelitian ini, pengolahan citra meliputi proses *resizing*, proses segmentasi, proses komplemen, proses morfologi, proses pengurangan derau (*noise reduction*), dan proses *cropping*.

METODE

Deskripsi Data

Penelitian yang dilakukan merupakan penelitian kuantitatif pada data yang meliputi data primer dan data sekunder yang diperoleh dari *dataset Fruits-262* (Minut, 2021). Data penelitian berupa citra (gambar) buah manggis yang berjumlah 207 citra. Data dibagi menjadi dua, yaitu data latih (*trainingset*) sebanyak 114 citra dan data uji (*testingset*) sebanyak 93 citra. Data diklasifikasi menjadi tiga kategori, yaitu mentah, setengah matang, dan matang. Data latih yang digunakan terdiri dari 38 citra buah manggis dengan kematangan mentah, 37 citra buah manggis dengan kematangan setengah matang (tujuan ekspor), dan 39 citra buah manggis dengan kematangan matang (tujuan domestik). Data uji yang digunakan terdiri dari 33 citra buah manggis dengan kematangan mentah, 30 citra buah manggis dengan kematangan setengah matang (tujuan ekspor), dan 30 citra buah manggis dengan kematangan matang (tujuan domestik).

Tahapan Penelitian

- a. Studi Literatur
Pada tahap ini dilakukan pengumpulan bahan referensi yang berkaitan dengan metode *Naive Bayes* dari buku, jurnal, dan sumber lainnya.
- b. Pengelompokan Data
Pada tahap ini dilakukan pengumpulan data berupa citra (gambar) buah manggis. Data dibagi menjadi data latih (*trainingset*) dan data uji (*testingset*). Data diklasifikasi menjadi tiga kelas, yaitu matang, setengah matang, dan mentah. Pemberian label klasifikasi pada data dilakukan berdasarkan Standar Prosedur Operasi yang dipublikasi oleh Departemen Pertanian (2004) dijabarkan pada Gambar 1.
- c. Pra-Pemrosesan Citra
Citra melalui tahap pra-pemrosesan yang meliputi proses *resizing*, proses segmentasi, proses komplemen, proses morfologi, proses pengurangan derau (*noise reduction*), dan proses *cropping*.
- d. Ekstraksi Fitur
Pada tahap ini diambil nilai fitur dari citra RGB, citra *grayscale*, citra HSV, dan citra CIELAB. Fitur yang diekstraksi diantaranya yaitu kontras, korelasi, energi, homogenitas, entropi, mean, varians, standar deviasi, *skewness*, dan kurtosis.
- e. Seleksi Fitur
Data nilai fitur melalui tahapan seleksi fitur menggunakan algoritma Minimum Redundancy Maximum Relevance (MRMR) untuk mengetahui pengaruh fitur terhadap penentuan kelas. Pengaruh fitur diurutkan berdasarkan nilai F-test Correlation Difference (FCD). Formulasi FCD menurut Ding dan Peng (2005) dengan $F(i, h)$ menyatakan nilai statistik uji F antara variabel fitur i dan variabel kelas h , S menyatakan jumlah fitur, $c(i, j)$ menyatakan nilai korelasi pearson antara variabel fitur i dan variabel fitur j :

$$FCD = \max \left\{ I(i, h) - \frac{1}{|S|} \sum_{j \in S} |c(i, j)| \right\} \quad (1)$$

Tahap	Ciri
0 	<ul style="list-style-type: none"> Warna kulit buah kuning kehijauan Kulit buah masih banyak mengandung getah Buah belum siap dipetik
1 	<ul style="list-style-type: none"> Warna kulit buah hijau kekuningan Buah belum tua dan getah masih banyak, isi buah masih sulit dipisahkan dari daging Buah belum siap dipanen
2 	<ul style="list-style-type: none"> Warna kulit buah kuning kemerahan dengan bercak merah hampir merata Buah hampir tua dan getah mulai berkurang, isi buah masih sulit dipisahkan dari daging
3 	<ul style="list-style-type: none"> Warna kulit buah merah kecoklatan Kulit buah masih bergetah, isi buah sudah dapat dipisahkan dari daging kulit. Buah disarankan dapat dipetik untuk tujuan ekspor
4 	<ul style="list-style-type: none"> Warna kulit buah merah keunguan Kulit buah masih sedikit bergetah, isi buah sudah dapat dipisahkan dari daging kulit dan buah dapat dikonsumsi Dapat dipetik untuk tujuan ekspor
5 	<ul style="list-style-type: none"> Warna kulit buah ungu kemerahan Buah mulai masak dan siap dikonsumsi. Getah telah hilang dan isi buah mudah dilepaskan Buah lebih sesuai untuk pasar domestik
6 	<ul style="list-style-type: none"> Warna kulit buah ungu kehitaman Buah sudah masak Buah sesuai untuk kebutuhan pasar domestik dan siap saji

Gambar 1. Indeks Kematangan Manggis

f. Klasifikasi Naive Bayes

Pada tahap ini terbagi menjadi dua proses, yaitu proses pelatihan (*training*), dan proses pengujian (*testing*). Proses pelatihan dilakukan untuk membangun model klasifikasi dari data latih. Proses pengujian dilakukan untuk menguji coba model klasifikasi terhadap data uji. Menurut Prasetyo (2012), Rumus Naive Bayes untuk permasalahan klasifikasi adalah sebagai berikut.

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^q P(X_i|Y)}{P(X)} \tag{2}$$

dengan definisi:

$P(Y|X)$: probabilitas terjadinya kelas Y dengan syarat atribut X .

$P(Y)$: probabilitas terjadinya kelas Y .

$\prod_{i=1}^q P(X_i|Y)$: probabilitas terjadinya kelas Y dari setiap atribut X_i .

$P(X)$: probabilitas terjadinya atribut X .

Pada penelitian ini, data fitur berupa data kontinu, maka diasumsikan bahwa data fitur tersebut berdistribusi *Gaussian Naive Bayes* dengan rata-rata μ dan standar deviasi σ (Kamel et al., 2019):

$$g(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3}$$

maka,

$$P(X_i|Y) = g(X_i, \mu_Y, \sigma_Y) \tag{4}$$

g. Evaluasi Kinerja Model Klasifikasi

Model Klasifikasi dievaluasi dengan menghitung performansi berupa akurasi, sensitivitas, spesifisitas, dan presisi. Sensitivitas adalah kemampuan model klasifikasi untuk memilih data yang sesuai. Spesifisitas adalah kemampuan model klasifikasi untuk menolak data yang tidak sesuai. Presisi adalah proporsi banyaknya kasus yang ditemukan relevan. Akurasi adalah penghitungan performansi hasil klasifikasi secara keseluruhan. Rumus perhitungan akurasi dari sebuah model klasifikasi adalah sebagai berikut.

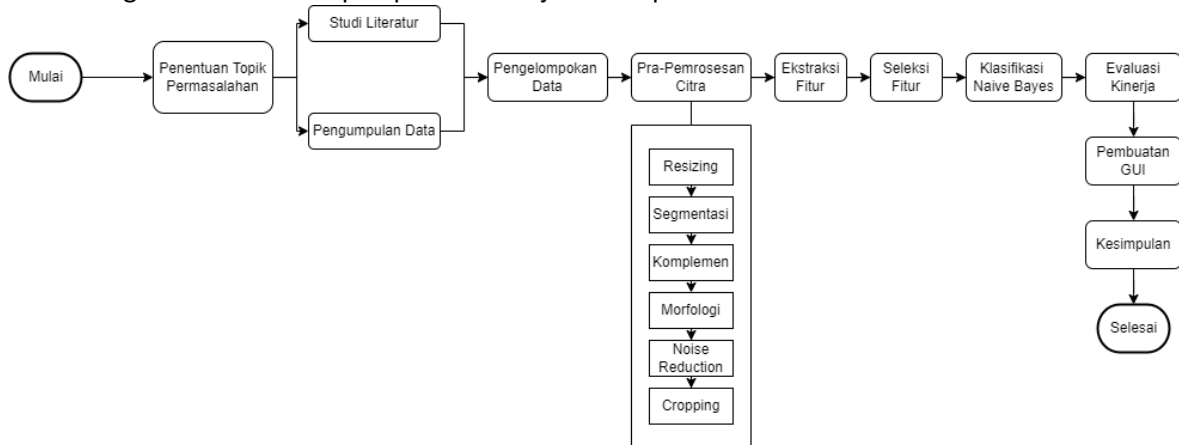
$$Akurasi = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \tag{5}$$

Atau dapat dituliskan sebagai:

$$Akurasi = \frac{\text{jumlah prediksi yang benar}}{\text{jumlah data keseluruhan}} \times 100\%$$

h. Pembuatan GUI (*Graphical User Interface*)

Program GUI dibuat berdasarkan model klasifikasi terbaik dengan menggunakan GUIDE Matlab. Diagram alir dari tahapan penelitian dijabarkan pada Gambar 2.



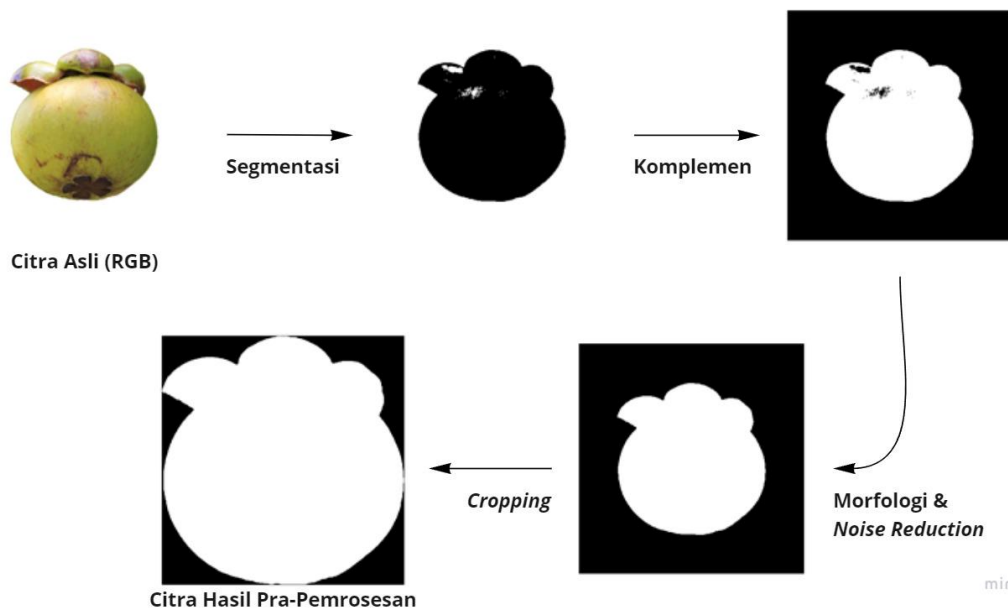
Gambar 2. Diagram alir tahapan penelitian

HASIL PENELITIAN

Hasil penelitian diperoleh melalui aplikasi Matlab yang dapat melakukan pengolahan terhadap citra ataupun data dengan baik.

Hasil Pra-Pemrosesan Citra

Pra-pemrosesan dilakukan agar saat melakukan ekstraksi fitur, nilai yang diperoleh adalah nilai fitur dari objek saja. Untuk mengurangi waktu komputasi, maka dilakukan proses *resizing* citra. Contoh pra-pemrosesan dari salah satu citra buah manggis digambarkan pada Gambar 3.



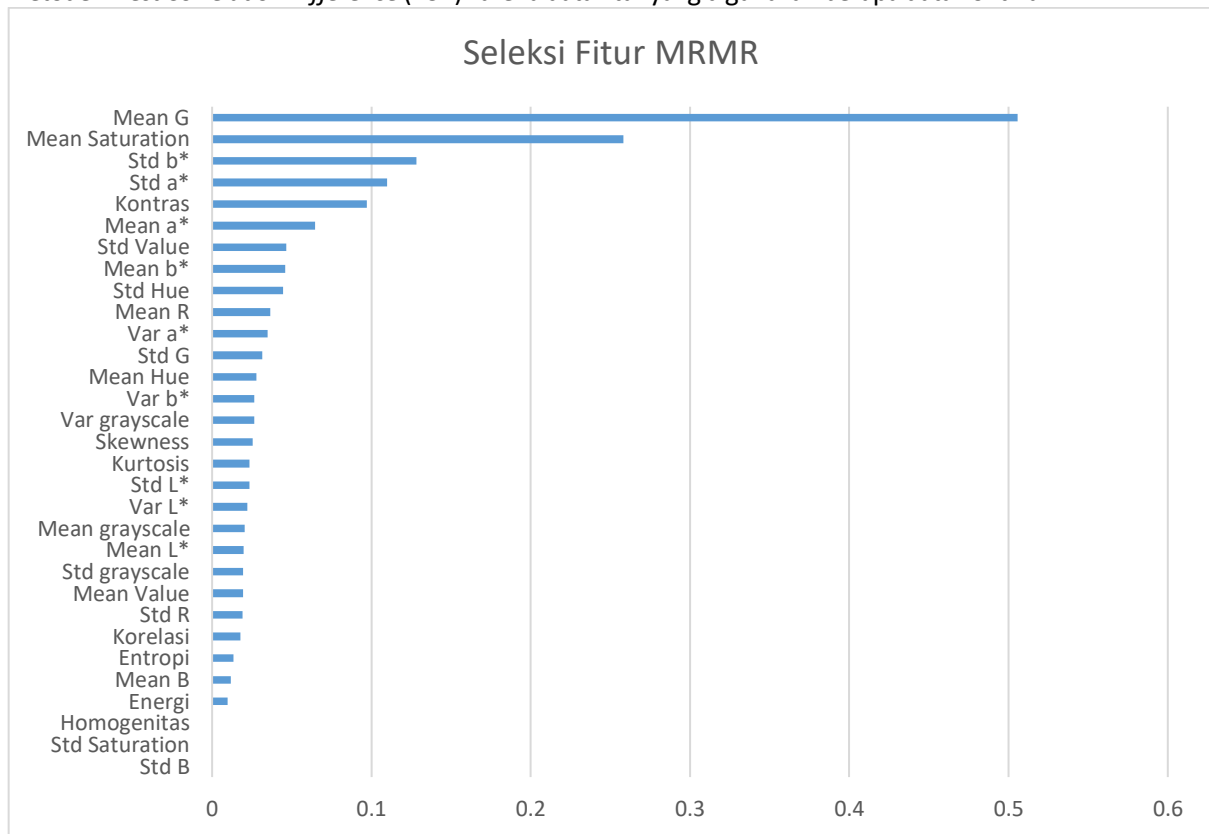
Gambar 3. Contoh pra-pemrosesan citra

Proses segmentasi dilakukan untuk memisahkan objek dan latar belakang dalam dua nilai intensitas yaitu 0 (hitam) dan 1 (putih). Pada ekstraksi fitur, nilai fitur yang diinginkan adalah nilai objek saja, sehingga prioritas warna latar belakang adalah warna hitam dengan nilai intensitas 0. Proses komplemen dilakukan untuk membalik warna hitam dan warna putih pada citra hasil segmentasi. Morfologi merupakan proses perubahan bentuk dalam sebuah citra, dengan proses morfologi yang dilakukan adalah *filling holes* untuk mengisi lubang pada objek. Proses pengurangan derau (Noise reduction) dilakukan untuk mengurangi titik ataupun sekumpulan titik yang terletak di luar objek. Proses *cropping* dilakukan untuk mengubah ukuran citra sesuai dengan hasil perhitungan batasan (*boundary*) objek. Batasan (*boundary*) objek dihitung dengan cara menentukan koordinat pusat objek (*centroid*),

lalu menghitung lebar dan tinggi objek berdasarkan pusat objek tersebut sehingga diperoleh batasan (*boundary*) objek.

Hasil Seleksi Fitur

Seleksi fitur dilakukan menggunakan algoritma *Minimum Redundancy Maximum Relevance* (MRMR) ([Gambar 4](#)) untuk mengetahui pengaruh fitur terhadap penentuan kelas. Metode MRMR yang digunakan adalah metode *F-Test Correlation Difference* (FCD) karena data fitur yang digunakan berupa data kontinu.



Gambar 4. Histogram seleksi fitur MRMR

Berdasarkan [Gambar 4](#), diketahui bahwa berdasarkan algoritma MRMR, 15 fitur dengan pengaruh terbesar adalah mean G, mean *Saturation*, std b*, std a*, kontras, mean a*, std *Value*, mean b*, std *Hue*, mean R, var a*, std G, mean *Hue*, var b*, dan var *grayscale*.

Klasifikasi Naive Bayes

Klasifikasi melalui proses pelatihan (*training*) dan proses pengujian (*testing*).

Pelatihan (Training)

Proses pelatihan (*training*) bertujuan untuk membangun model klasifikasi *Naive Bayes* dari data latih. Setiap model klasifikasi memiliki pemilihan parameter yang berbeda. Masing-masing model klasifikasi tersebut diuji performasinya untuk mengetahui model *Naive Bayes* terbaik dalam mengklasifikasi kematangan buah manggis. Performasi suatu model dapat diukur berdasarkan tingkat akurasi. Rumus untuk menghitung akurasi dari sebuah model klasifikasi adalah sebagai berikut.

$$Akurasi = \frac{TP + TN}{TP + FP + TN + FN} \times 100$$

Atau dapat dituliskan sebagai:

$$Akurasi = \frac{\text{jumlah prediksi yang benar}}{\text{jumlah data keseluruhan}} \times 100\%$$

Model klasifikasi *Naive Bayes* yang dibangun tanpa mempertimbangkan hasil seleksi fitur MRMR, dan 2 model klasifikasi *Naive Bayes* terbaik yang dibangun berdasarkan hasil seleksi fitur MRMR dijabarkan pada [Tabel 1](#).

Tabel 1. Model klasifikasi naive bayes

Model	Pilihan Fitur						Seleksi Fitur	Total Fitur	Akurasi Pelatihan	
	RGB		CIELAB		HSV					Grayscale
NB1	Mean mean mean R, std std B	R, G, B, dan std G	Mean mean std L*, std b*, var a*, dan var b*	L*, a*, std L*, b*, var a*, dan var b*	Mean mean std H, S, dan std V	H, S, dan std V	Kontras, korelasi, energi, homogenitas, entropi, standar deviasi, mean, varians, skewness, dan kurtosis	-	31	94,7%
NB2	Mean mean std G, dan std G	R, dan std G	Mean mean std a*, b*, var a*, dan var b*	a*, std a*, std b*, var a*, dan var b*	Mean mean std H, S, dan std V	H, S, dan std V	Kontras, grayscale, skewness, dan kurtosis	MRMR	18	95,6%
NB3	Mean mean std G, dan std G	R, dan std G	Mean mean std a*, b*, std a*, std b*, dan var a*	a*, std b*, std a*, std b*, dan var a*	Mean mean std H, S, dan std V	H, S, dan std V	Kontras	MRMR	13	95,6%

Berdasarkan Tabel 1, diketahui bahwa model NB2 dan NB3 memperoleh akurasi pelatihan yang sama yaitu sebesar 95,6%, sedangkan model NB1 memperoleh akurasi pelatihan sebesar 94,7%. Model NB3 mampu memperoleh akurasi pelatihan yang optimal dengan pilihan fitur yang lebih sedikit.

Pengujian (Testing)

Proses pengujian (*testing*) bertujuan untuk menguji coba model klasifikasi *Naive Bayes* yang telah dibangun terhadap data uji. Rincian hasil klasifikasi pengujian dari masing-masing model klasifikasi dapat dilihat pada Tabel 2.

Tabel 2. Rincian hasil klasifikasi pengujian

Model	Aktual	Prediksi			Akurasi Pengujian	Jumlah Citra
		Matang	Mentah	Setengah matang		
NB1	Matang	25	0	5	91,4%	30
	Mentah	0	32	1		33
	Setengah Matang	1	1	28		30
	Total					93
NB2	Matang	27	0	3	94,6%	30
	Mentah	0	33	0		33
	Setengah Matang	1	1	28		30
	Total					93
NB3	Matang	28	0	2	95,7%	30
	Mentah	0	33	0		33
	Setengah Matang	2	0	28		30
	Total					93

Dari Tabel 2, dapat dilihat bahwa model NB1 memperoleh akurasi pengujian sebesar 91,4%, model NB2 memperoleh akurasi pengujian sebesar 94,6%, sedangkan model NB3 memperoleh akurasi pengujian sebesar

95,7%. Model NB3 memiliki performasi terbaik, dengan akurasi pelatihan sebesar 95,6% dan akurasi pengujian sebesar 95,7%.

Evaluasi Kinerja Model Klasifikasi

Pada tahap ini, dilakukan evaluasi kinerja terhadap model klasifikasi yang memiliki akurasi klasifikasi terbaik. Model klasifikasi dievaluasi performasinya dengan menghitung akurasi, sensitivitas, spesifisitas, dan presisi dari model tersebut. Perhitungan akurasi, sensitivitas, spesifisitas, dan presisi dapat dilakukan dengan menggunakan matriks konfusi dari model klasifikasi. Matriks konfusi dari model klasifikasi NB3 dijabarkan dalam Tabel 3.

Tabel 3. Matriks konfusi model NB3

Kelas		Prediksi		
		Matang	Mentah	Setengah Matang
Aktual	Matang	28	0	2
	Mentah	0	33	0
	Setengah Matang	2	0	28

Perhitungan nilai sensitivitas, spesifisitas, dan presisi dari model klasifikasi NB3 terhadap kelas matang adalah sebagai berikut.

$$Sensitivitas = \frac{TP}{TP + FN} \times 100\% = \frac{28}{28 + (0 + 2)} \times 100\% = 93,3\%$$

$$Spesifisitas = \frac{TN}{TN + FP} \times 100\% = \frac{(33 + 28 + 0 + 0)}{(33 + 28 + 0 + 0) + (0 + 2)} \times 100\% = 96,8\%$$

$$Presisi = \frac{TP}{TP + FP} \times 100\% = \frac{28}{28 + (0 + 2)} \times 100\% = 93,3\%$$

Berdasarkan hasil perhitungan, dapat diinterpretasikan bahwa model NB2 memiliki persentase prediksi kelas matang sebesar 93,3%, dengan kemampuan model NB3 mengklasifikasi data kelas matang secara tepat sebesar 93,3%, dan kemampuan untuk menolak data yang tidak sesuai sebesar 96,8%. Perhitungan nilai sensitivitas, spesifisitas, dan presisi dari model klasifikasi NB3 terhadap kelas mentah adalah sebagai berikut.

$$Sensitivitas = \frac{TP}{TP + FN} \times 100\% = \frac{33}{33 + (0 + 0)} \times 100\% = 100\%$$

$$Spesifisitas = \frac{TN}{TN + FP} \times 100\% = \frac{(28 + 28 + 2 + 2)}{(28 + 28 + 2 + 2) + (0 + 0)} \times 100\% = 100\%$$

$$Presisi = \frac{TP}{TP + FP} \times 100\% = \frac{33}{33 + (0 + 0)} \times 100\% = 100\%$$

Berdasarkan hasil perhitungan, dapat diinterpretasikan bahwa model NB3 memiliki persentase prediksi kelas mentah sebesar 100%, dengan kemampuan model NB3 untuk mengklasifikasi data kelas mentah secara tepat sebesar 100%, dan kemampuan untuk menolak data yang tidak sesuai sebesar 100%. Perhitungan nilai sensitivitas, spesifisitas, dan presisi dari model klasifikasi NB3 terhadap kelas setengah matang adalah sebagai berikut.

$$Sensitivitas = \frac{TP}{TP + FN} \times 100\% = \frac{28}{28 + (2 + 0)} \times 100\% = 93,3\%$$

$$Spesifisitas = \frac{TN}{TN + FP} \times 100\% = \frac{(29 + 33 + 0 + 0)}{(29 + 33 + 0 + 0) + (2 + 0)} \times 100\% = 96,9\%$$

$$Presisi = \frac{TP}{TP + FP} \times 100\% = \frac{28}{28 + (2 + 0)} \times 100\% = 93,3\%$$

Berdasarkan hasil perhitungan, dapat diinterpretasikan bahwa model NB3 memiliki persentase prediksi kelas setengah matang sebesar 93,3%, dengan kemampuan model NB3 untuk mengklasifikasi data kelas setengah matang secara tepat sebesar 93,33%, dan kemampuan untuk menolak data yang tidak sesuai sebesar 96,9%.

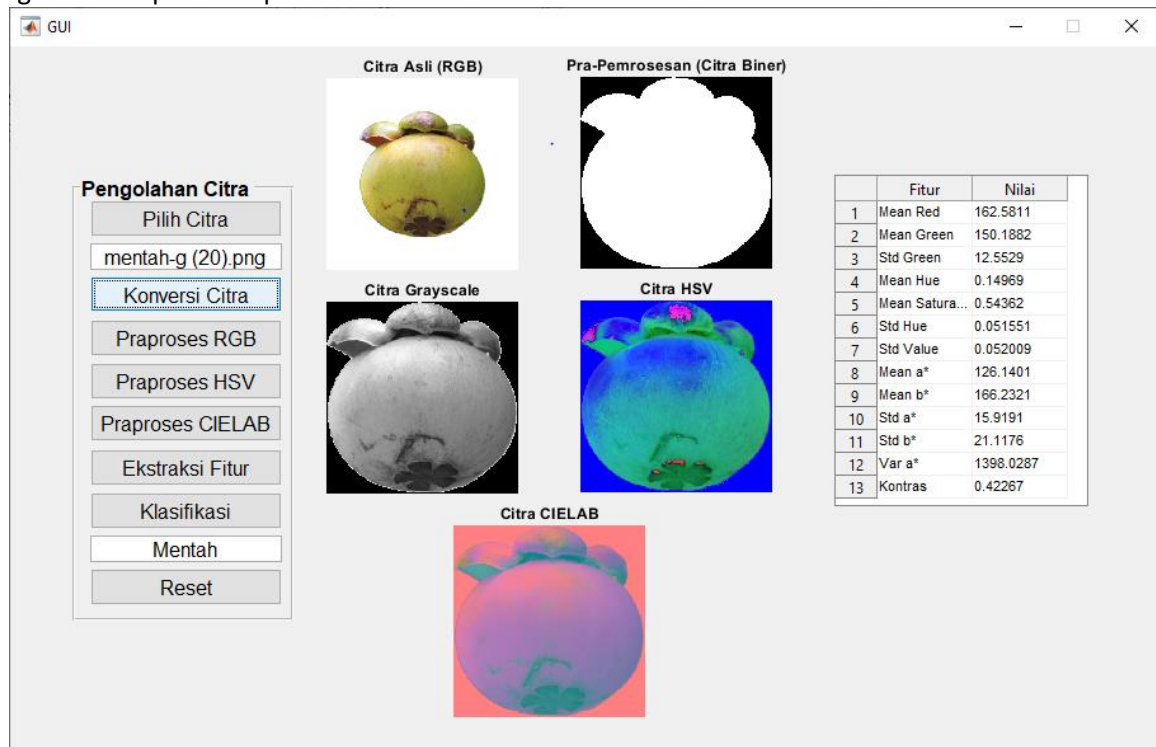
Secara singkat, hasil nilai akurasi, sensitivitas, spesifisitas, dan presisi dari model klasifikasi NB3 dapat dilihat pada [Tabel 4](#).

Tabel 4. Evaluasi kinerja model NB3

Kelas	Akurasi	Sensitivitas	Spesifisitas	Presisi
Matang	95,7%	93,3%	96,8%	93,3%
Mentah		100%	100%	100%
Setengah Matang		93,3%	96,9%	93,3%

Pembuatan GUI

Program GUI dirancang dengan menggunakan GUIDE Matlab. Pada rancangan GUI klasifikasi kematangan buah manggis ini, digunakan komponen dari GUIDE Matlab, diantaranya yaitu *axes*, *panel*, *edit text*, *push button*, dan *table*. Model klasifikasi yang digunakan dalam program GUI ini adalah model NB3, dengan pemilihan parameter fitur yang digunakan dalam pembangunan model diantaranya yaitu mean R, mean G, std G, mean Saturation, mean Hue, std Hue, std Value, mean a*, mean b*, std a*, std b*, var a*, dan kontras. Antarmuka program GUI dapat dilihat pada [Gambar 5](#).



Gambar 5. Tampilan Program GUI

Rincian waktu komputasi dari program GUI klasifikasi kematangan buah manggis dapat dilihat pada [Tabel 5](#).

Tabel 5. Waktu komputasi program GUI

Proses	Waktu (Detik)
Memuat citra	0,2933
Konversi & pra-pemrosesan citra	0,5462
Ekstraksi fitur	0,0499
Klasifikasi	0,0201
Total	0,9095

Jadi, diperoleh waktu komputasi program GUI untuk mengklasifikasi kematangan buah manggis adalah selama 0,9095 detik.

PEMBAHASAN

Berdasarkan hasil penelitian yang telah dilakukan, metode Naive Bayes terbukti mampu mengklasifikasi tingkat kematangan buah manggis dengan sangat baik. Hasil penelitian menunjukkan bahwa tingkat akurasi yang dicapai mencapai 95,7%. Selain itu, penelitian juga menunjukkan bahwa model klasifikasi Naive Bayes dengan seleksi fitur MRMR menghasilkan tingkat akurasi yang lebih tinggi dibandingkan dengan model tanpa seleksi fitur MRMR. Model dengan seleksi fitur MRMR mencapai akurasi pelatihan sebesar 95,6% dan akurasi pengujian sebesar 95,7%, sementara model tanpa seleksi fitur MRMR hanya mencapai akurasi pelatihan sebesar 94,7% dan akurasi pengujian sebesar 91,4%.

Pengklasifikasian tingkat kematangan manggis dengan menggunakan Naive Bayes merupakan sebuah penelitian yang masih jarang dilakukan, sehingga belum terdapat perbandingan langsung dengan penelitian sebelumnya. Namun, terdapat penelitian lain yang juga menggunakan pendekatan Naive Bayes untuk mengklasifikasikan tingkat kematangan buah, yaitu penelitian yang dilakukan oleh Kusuma dan Putra (2018) tentang klasifikasi kematangan tomat menggunakan algoritma *Naive Bayes* dan ekstraksi fitur histogram memperoleh akurasi sebesar 76%. Penelitian tersebut hanya menggunakan fitur dari citra RGB dan grayscale, dengan pra-pemrosesan yang lebih sederhana yaitu *resizing* dan *cropping*. Dalam penelitian klasifikasi kematangan manggis ini, akurasi yang lebih tinggi berhasil dicapai karena opsi fitur yang dapat digunakan lebih banyak dan juga pra-pemrosesan citra lebih komprehensif, dengan pra-pemrosesan citra meliputi proses *resizing*, segmentasi, komplement, morfologi, pengurangan noise, dan *cropping*.

Teknik-teknik pra-pemrosesan yang digunakan dalam penelitian ini terinspirasi oleh penelitian yang dilakukan sebelumnya oleh Susanto et al. (2020) tentang klasifikasi mutu bawang merah berdasarkan citra HSV dan identifikasi ukuran. Penelitian tersebut berhasil mencapai tingkat akurasi sebesar 91,67%. Pada penelitian tersebut, pengolahan citra yang dilakukan meliputi proses konversi, proses ekstraksi, proses segmentasi, proses morfologi, dan proses pelabelan objek.

Faktor utama dalam mengklasifikasi kematangan buah manggis adalah ekstraksi fitur warna dan tekstur kulit buah. Maka, dalam pengklasifikasian dengan data berupa citra buah, citra perlu melewati tahapan pra-pemrosesan agar objek dalam citra dapat dikenali dengan baik pada saat dilakukan pengklasifikasian. Penggunaan algoritma MRMR (Minimum Redundancy Maximum Relevance) juga membantu dalam pemilihan parameter fitur yang optimal, dengan fitur-fitur optimal yang diperoleh adalah sebanyak 13 fitur, diantaranya yaitu mean R, mean G, std G, mean Saturation, mean Hue, std Hue, std Value, mean a*, mean b*, std a*, std b*, var a*, dan kontras. Seluruh proses ini berkontribusi pada tingkat akurasi yang sangat baik dalam klasifikasi kematangan buah manggis menggunakan metode Naive Bayes.

SIMPULAN

Fitur yang diekstraksi dalam penelitian ini berjumlah 31 fitur. Fitur tersebut diekstraksi dari citra RGB, citra *grayscale*, citra HSV, dan citra CIELAB. Data hasil ekstraksi fitur melalui tahap seleksi fitur untuk mengetahui fitur-fitur apa saja yang berpengaruh dalam penentuan kelas. Klasifikasi kematangan buah manggis dengan menggunakan algoritma *Naive Bayes* memperoleh hasil klasifikasi yang maksimal dengan pemilihan fitur sebanyak 13 fitur, diantaranya yaitu mean G, std G, mean *Saturation*, mean *Hue*, std *Hue*, std *Value*, mean a*, mean b*, std a*, std b*, var a*, dan kontras. Pemilihan parameter fitur berdasarkan pada urutan pengaruh fitur terhadap penentuan kelas yang diperoleh melalui seleksi fitur menggunakan algoritma *Minimum Redundancy Maximum Relevance* (MRMR). Klasifikasi kematangan buah manggis berdasarkan fitur warna dan tekstur dengan

menggunakan algoritma *Naive Bayes* memperoleh tingkat akurasi sebesar 95,7% dengan sensitivitas, spesifisitas, dan presisi untuk kelas matang sebesar 93,3%, 96,8%, dan 93,3%. Sensitivitas, spesifisitas, dan presisi untuk kelas mentah masing-masing sebesar 100%. Sensitivitas, spesifisitas, dan presisi untuk kelas setengah matang sebesar 93,3%, 96,9%, dan 93,3%.

DAFTAR PUSTAKA

- Azima, A. S., Noriham, A., & Manshoor, N. (2017). Phenolics, antioxidants and color properties of aqueous pigmented plant extracts: *Ardisia colorata* var. *elliptica*, *Clitoria ternatea*, *Garcinia mangostana* and *Syzygium cumini*. *Journal of Functional Foods*, 38, 232-241. <https://doi.org/10.1016/j.jff.2017.09.018>
- Badan Pusat Statistik. (2021). Produksi Tanaman Buah-buahan (2017-2021). <https://www.bps.go.id/indicator/55/62/1/produksi-tanaman-buah-buahan.html>
- Departemen Pertanian. (2004). Standar Prosedur Operasi. Jakarta: Direktorat Tanaman Buah Departemen Pertanian.
- Ding, C., and H. Peng. (2005). Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology*, 3(2), 185–205. <https://doi.org/10.1142/S0219720005001004>
- Kamel, H., Abdulah, D., & Al-Tuwaijari, J. M. (2019). Cancer classification using gaussian naive bayes algorithm. In *2019 International Engineering Conference (IEC)*, 165-170. IEEE. <https://doi.org/10.1109/IEC47844.2019.8950650>
- Khisnudin, I. S., dan Murinto. (2020). Dragon fruit maturity detection based- HSV space color using naive bayes classifier method. *IOP Conference Series: Materials Science and Engineering*, 771(1). <https://doi.org/10.1088/1757-899X/771/1/012022>
- Kusuma, A., & Putra, M. D. M. (2018). Tomato maturity classification using naive bayes algorithm and histogram feature extraction. *Journal of Applied Intelligent System* 3(1), 39-48. <https://doi.org/10.33633/jais.v3i1.1988>
- Li, G., Thomas, S., & Johnson, J. J. (2013). Polyphenols from the mangosteen (*Garcinia mangostana*) fruit for breast and prostate cancer. *Frontiers in pharmacology*, 4, 80. <https://doi.org/10.3389/fphar.2013.00080>
- Liantoni, F., dan Nugroho, H. (2015). Klasifikasi daun herbal menggunakan metode naive bayes classifier dan k-nearest neighbor. *Jurnal Simantec*, 5(1). <https://journal.trunojoyo.ac.id/simantec/article/view/1009>
- Manik, F. Y, Saragih, K. S. (2017). Klasifikasi belimbing menggunakan naive bayes berdasarkan fitur warna RGB. *IJCCS*, 11(1), 99. 1978-1520. <https://doi.org/10.22146/ijccs.17838>
- Minut, M. (2021). Fruits-262 dataset: A dataset containing a vast majority of the popular and known fruits. *Kaggle*. <https://www.kaggle.com/datasets/aelchimminut/fruits262>
- Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta: ANDI Yogyakarta.
- Ray, S. (2019). A Quick Review of Machine Learning Algorithms. *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 35-39. <https://doi.org/10.1109/COMITCon.2019.8862451>
- Sahagun, M. A. M. (2020). Determination of Sweetness Level of Mutingia Calabura using HSV Colorspace. *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, 1-5. <https://doi.org/10.1109/ICDABI51230.2020.9325684>
- Sari, C. A., et al. (2020). Papaya fruit type classification using LBP features extraction and naive bayes classifier. *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 28-33. <https://doi.org/10.1109/iSemantic50169.2020.9234240>
- Sugianto, S., & Wibowo, F. (2015). Klasifikasi tingkat kematangan buah pepaya (*Carica papaya* L) California (Callina-IPB 9) dalam ruang warna hsv dan algoritma k-nearest neighbors. *prosiding SENATEK*, 335-341.

- Susanto, A., Dewantoro, Z. H., Sari, C. A., Setiadi, D. R. I. M., Rachmawanto, E. H., dan Mulyono, I. U. W. (2020). Shallot quality classification using HSV color models and size identification based on naive bayes classifier. *Journal of Physics: Conference Series*, 1577(1). <https://doi.org/10.1088/1742-6596/1577/1/012020>
- Vembandasamy, K., Sasipriya, R., & Deepa, E. (2015). Heart diseases detection using Naive Bayes algorithm. *International Journal of Innovative Science, Engineering & Technology*, 2(9), 441-444. https://www.ijiset.com/vol2/v2s9/IJSET_V2_I9_54.pdf
- Whidhiasih, R. N., Guritman, S., & Suprio, P. T. (2012). Klasifikasi kematangan buah manggis ekspor dan lokal berdasarkan warna dan tekstur menggunakan fuzzy neural network. *Jurnal Ilmu Komputer dan Agri-Informatika*, 1(2), 71-77. <https://doi.org/10.29244/jika.1.2.71-77>
- Widiastuti, A., Sobir, S., & Suhartanto, M. R. (2013). Analisis keragaman genetik manggis (*Garcinia mangostana*) diiradiasi dengan sinar gamma berdasarkan penanda ISSR. *Asian Journal of Tropical Biotechnology*, 10(1), 15-22. <https://doi.org/10.13057/biotek/c100103>