

KLASIFIKASI TOMAT BERDASARKAN VARIETAS DENGAN EKSTRAKSI FITUR RGB DAN ALGORITMA NAÏVE BAYES

Evi Widyastuti^{1*}, Arif Hermawan², Donny Avianto³

¹²³Magister Teknologi Informasi, Universitas Teknologi Yogyakarta, DI Yogyakarta, Indonesia

Email: ¹eviwidyast.mti@gmail.com, ²ariefdb@uty.ac.id, ³donny@uty.ac.id

(*: coresponding author)

Abstrak-Tomat merupakan tanaman sayur kategori buah yang mudah dibudidayakan di berbagai tempat. Beragamnya varietas tomat, seperti Red Zebra Tomato, Green Zebra Tomato, dan Kumato, sering kali menyulitkan identifikasi varietas secara akurat. Kesalahan dalam klasifikasi dapat berdampak pada pemilihan kondisi lingkungan dan manajemen hama atau penyakit, yang berujung pada hasil budidaya yang tidak optimal. Saat ini, penelitian lebih banyak berfokus pada bentuk, penyakit, dan tingkat kematangan tomat, sementara klasifikasi kultivar berdasarkan ciri warna masih jarang dilakukan. Penelitian ini bertujuan untuk mengembangkan metode klasifikasi kultivar tomat berdasarkan warna *Red* (R), *Green* (G), dan *Blue* (B) menggunakan algoritma Naïve Bayes. Penelitian dilakukan dengan mengumpulkan 45 citra tomat yang memiliki bentuk serupa tetapi warna berbeda (merah, hijau, dan merah kehitaman). Tahapan penelitian meliputi ekstraksi fitur RGB, pembulatan data, pembagian data latih dan uji dengan rasio 70:30, serta proses klasifikasi menggunakan Naïve Bayes. Pengujian ulang dilakukan dengan menghilangkan atribut warna tertentu untuk mengevaluasi pengaruhnya terhadap akurasi. Penelitian ini diharapkan dapat mendukung identifikasi varietas tomat secara cepat dan akurat, meningkatkan efisiensi pertanian modern, serta memperluas penerapan teknologi dalam industri pertanian guna mewujudkan pertanian yang maju, mandiri, dan modern. Hasil penelitian menunjukkan bahwa metode ekstraksi fitur RGB dan algoritma Naïve Bayes mampu mengklasifikasikan kultivar tomat dengan akurasi hingga 78.57%. Atribut warna RG memiliki pengaruh terbesar terhadap akurasi, mencapai 85.71%.

Kata Kunci: citra tomat, ekstraksi fitur, fitur warna rgb, klasifikasi tomat, naïve bayes.

Abstract- *Tomato is a fruit-category vegetable plant that is easy to cultivate in various locations. The diversity of tomato varieties, such as Red Zebra Tomato, Green Zebra Tomato, and Kumato, often makes rapid and accurate variety identification challenging. Misclassification can impact the selection of environmental conditions and pest or disease management, ultimately leading to suboptimal cultivation results. Currently, research primarily focuses on tomato shape, diseases, and ripeness levels, while cultivar classification based on color characteristics remains limited. This study aims to develop a method for classifying tomato cultivars based on RGB color features using the Naïve Bayes algorithm. The research was conducted by collecting 45 tomato images with similar shapes but different colors (red, green, and dark red). The research stages include RGB feature extraction, data rounding, splitting training and test data with a 70:30 ratio, and classification using Naïve Bayes. A re-evaluation was performed by removing specific color attributes to assess their impact on accuracy. This study is expected to support rapid and accurate tomato variety identification, enhance efficiency in modern agriculture, and expand the application of technology in the agricultural industry to achieve advanced, self-sufficient, and modern farming. The results show that the RGB feature extraction method and the Naïve Bayes algorithm can classify tomato cultivars with an accuracy of up to 78.57%. The RG color attributes have the most significant impact on accuracy, reaching 85.71%.*

Keywords: *feature extraction, naïve bayes, RGB color features, tomato classification, tomato images*

1. PENDAHULUAN

Tomat (*Solanum lycopersicum*) merupakan salah satu tanaman sayur kategori buah yang mudah dibudidayakan di berbagai tempat. Saat ini sudah sangat banyak varietas tanaman tomat yang dibudidayakan dan terus bertambah seiring dengan perkembangan pemuliaan tanaman tomat. Varietas yang beragam membuat masyarakat awam sulit untuk mengenali dan membedakan varietas tomat. Selain itu, industri tomat menghadapi tantangan serius dalam mengelola klasifikasi varietas tomat secara cepat dan akurat. Identifikasi manual varietas tomat memerlukan waktu lama dan sumber daya yang banyak. Salah dalam mengklasifikasikan kultivar tomat akan memiliki beberapa dampak yang menyebabkan budidaya tomat tidak optimal. Dampak-dampak yang dapat antara lain ketidakcocokan lingkungan, manajemen hama dan penyakit yang tidak tepat, budidaya tidak sesuai dengan yang diinginkan oleh petani. Penelitian tentang klasifikasi tomat sudah dilakukan sebelumnya antara lain klasifikasi yang dilakukan berdasarkan bentuk, klasifikasi penyakit tomat, dan klasifikasi tingkat kematangan. Tetapi masih sangat sedikit penelitian yang mengklasifikasikan tomat berdasarkan kemiripan bentuk tetapi memiliki warna yang berbeda.

Kemajuan ilmu pengetahuan dan teknologi pengolahan citra digital memungkinkan mengklasifikasikan hasil pertanian dan perkebunan secara otomatis menggunakan aplikasi pengolahan citra. Aplikasi tersebut dapat dikembangkan di masa depan untuk mendukung pertanian modern melalui identifikasi varietas secara tepat dan cepat sehingga membantu petani untuk melakukan budidaya dengan hasil yang lebih optimal. Selain itu, penelitian ini bertujuan untuk memperluas potensi penerapan teknologi dalam industri pertanian untuk mencapai visi pertanian Indonesia yang maju, mandiri dan modern.

Beberapa penelitian yang sudah dilakukan sebelumnya antara lain adalah klasifikasi kematangan buah tomat menggunakan fitur RGB dan HSI berbasis backpropagation. Hasil pada penelitian ini didapatkan akurasi yang cukup tinggi yaitu 97.29% untuk mengklasifikasikan kematangan tomat [1]. Penelitian mengenai klasifikasi tomat sebelumnya juga pernah dilakukan oleh Robert G. de Luna, et al. dalam Jurnal “*Size Classification of Tomato Fruit Using Thresholding, Machine Learning and Deep Learning Techniques*”. Penelitian ini dilakukan untuk mengklasifikasikan tomat berdasarkan ukuran yang digunakan untuk penyortiran tomat. Hasil dari penelitian ini adalah KNN merupakan metode yang dapat memberikan akurasi paling tinggi dibanding metode SVM dan ANN [2].

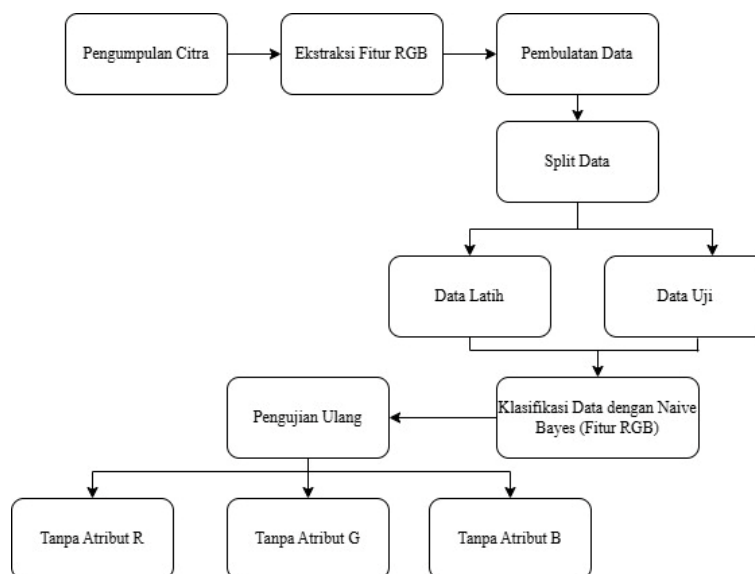
Algoritma Naïve Bayes juga digunakan dalam beberapa penelitian. Salah satu penelitian yang menggunakan algoritma ini adalah klasifikasi dokumen. Algoritma ini disebut sebagai *classifier* terbaik untuk mengklasifikasikan dokumen. Kinerja Naïve Bayes masih tetap unggul ketika pengujian dilakukan pada tipe data kategori. Semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas dan memiliki performa yang baik. [3] Berdasarkan penelitian ini maka penelitian yang akan dilakukan juga akan melakukan pengurangan fitur pada pengujian ulang.

Penelitian lain yang dilakukan untuk klasifikasi tomat yang banyak dilakukan antara lain untuk klasifikasi kematangan [1], [4], klasifikasi jenis penyakit tomat [5], [6], [7], [8], [9], [10] dan klasifikasi jenis tomat berdasarkan bentuk [2], [11]. Maka penelitian yang saat ini dilakukan adalah klasifikasi kultivar tomat berdasarkan bentuk yang sama tetapi memiliki warna yang berbeda dengan algoritma yang pada beberapa penelitian disebutkan sebagai *classifier* terbaik yaitu Algoritma Naïve Bayes. Dalam konteks klasifikasi berbasis warna, penelitian ini menonjol karena menganalisis pengaruh atribut warna RGB secara individual terhadap akurasi klasifikasi, yang belum banyak dibahas sebelumnya. Dengan pendekatan ini, penelitian diharapkan dapat mendukung otomatisasi sistem klasifikasi untuk pertanian modern serta mempercepat pengenalan varietas dengan akurasi yang lebih tinggi. Penelitian ini menemukan metode yang paling sesuai untuk klasifikasi tomat dan pengaruh setiap fitur yang ada dalam ekstraksi Fitur RGB. Manfaat penelitian ini meliputi peningkatan efisiensi identifikasi varietas tomat yang dapat mendukung petani dalam optimalisasi hasil budidaya. Selain itu, penelitian ini membuka potensi penerapan teknologi dalam industri pertanian untuk mendukung terwujudnya visi pertanian Indonesia yang maju, mandiri, dan modern.

2. METODE PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilakukan untuk mengklasifikasikan kultivar tomat yang memiliki kesamaan bentuk tetapi memiliki perbedaan warna. Klasifikasi dilakukan dengan algoritma Naïve Bayes untuk mengklasifikasikan 3 kultivar tomat yaitu Red Zebra Tomato, Green Zebra Tomato dan Kumato. Citra yang digunakan dalam penelitian ini diperoleh dari Kaggle yang berjudul *Tomato Cultivars*. [12] Data yang diperoleh terlebih dahulu dilakukan ekstraksi Fitur RGB sehingga didapatkan nilai *Red* (R), *Green* (G) dan *Blue* (B) dari setiap citra. Setelah itu, dilakukan proses klasifikasi untuk kultivar tomat dengan menggunakan algoritma Naïve Bayes. Tahapan pada penelitian terdapat pada Gambar 1.



Gambar 1. Tahapan Penelitian

Tahapan pertama pada penelitian ini adalah pengumpulan citra tomat yang diperoleh dari Kaggle. Citra-citra ini dipilih berdasarkan kesamaan bentuk tetapi memiliki perbedaan warna, yaitu merah, hijau, dan merah kehitaman. Setelah citra terkumpul, dilakukan proses ekstraksi fitur RGB untuk mendapatkan nilai atribut warna Red (R), Green (G), dan Blue (B) dari masing-masing citra. Nilai-nilai ini menjadi data utama yang akan digunakan untuk klasifikasi. Tahapan berikutnya adalah pembulatan data atribut RGB yang telah diekstraksi. Hal ini bertujuan untuk menyederhanakan data sehingga lebih mudah diolah oleh algoritma klasifikasi. Data yang telah diproses kemudian dibagi menjadi dua kelompok yaitu data latih dan data uji. Model ini bertujuan untuk mengklasifikasikan citra tomat ke dalam kultivar tertentu. Setelah model dikembangkan, dilakukan pengujian ulang untuk mengevaluasi pengaruh atribut warna RGB secara individu terhadap akurasi. Proses ini dilakukan dengan menghilangkan satu atribut pada setiap pengujian.

2.2 Pengumpulan Citra Tomat

Citra yang digunakan merupakan citra berwarna. Sumber data citra diperoleh dari *website* yang menawarkan *dataset* yaitu Kaggle. Gambar 2 menyajikan data citra tomat yang akan digunakan dalam penelitian sebagai data latih dan data uji.



Gambar 2. Data Citra Tomat

Data citra yang diambil hanya data citra kultivar tomat yang memiliki bentuk yang sama tetapi warna yang berbeda sebanyak 45 citra. Data ini akan dibagi menjadi data latih dan data uji dengan ratio 70:30. Data latih sebanyak 70% dari total data (31 citra) digunakan untuk melatih model. Data uji sebanyak 30% dari total data (14 citra) digunakan untuk menguji performa model. Data latih digunakan untuk membangun model klasifikasi menggunakan algoritma Naïve Bayes dengan fitur RGB sebagai atribut input.

2.3 Ekstraksi Fitur

Fitur adalah ciri khas suatu objek pada suatu gambar yang dimaksudkan untuk membedakannya dengan objek lainnya. Informasi yang diekstraksi digunakan sebagai parameter masukan untuk membedakan satu objek dengan objek lainnya selama fase identifikasi atau klasifikasi. Fitur dapat diamati dan dikenali karena setiap objek mempunyai pola-pola tertentu sehingga mudah dibedakan oleh manusia. Fitur warna, fitur bentuk, dan fitur tekstur adalah beberapa fitur yang biasa diamati. Pada fitur warna, objek atau benda dikenali karena adanya perbedaan warna di ruang warna tertentu misal RGB.

Ekstraksi fitur warna RGB banyak digunakan dalam beberapa penelitian di bidang pertanian [13],[14] [15] diantaranya untuk mengklasifikasikan kematangan [16], mengklasifikasikan hama dan penyakit tanaman [17], dan juga untuk mengklasifikasikan varietas ataupun kultivar hasil pertanian.[18] Setiap piksel pada citra warna mewakili warna yang merupakan kombinasi dari tiga warna dasar (RGB = Red Green Blue). Setiap warna dasar menggunakan penyimpanan 8bit, yang berarti setiap warna mempunyai gradasi sebanyak 255 warna. Representasi warna ini terdiri dari tiga unsur utama yaitu merah (*red*), hijau (*green*), dan biru (*blue*). Gabungan tiga warna ini membentuk warna-warna lainnya berdasarkan intensitas dari masing-masing warna tersebut dengan intensitas maksimal, dan warna hitam merupakan gabungan dari ketiga warna tersebut dengan intensitas [1].

2.4 Algoritma Naïve Bayes

Naïve Bayes Classifier merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes. Naive Bayes merupakan metode *supervised document classification* yang berarti membutuhkan data *training* sebelum melakukan proses klasifikasi. [19] Ciri utama dari Naïve Bayes Classifier ini adalah asumsi yang sangat kuat akan independensi

dari masing-masing kondisi atau kejadian. Sebelum menjelaskan Naïve Bayes Classifier ini, akan dijelaskan terlebih dahulu Teorema Bayes yang menjadi dasar dari metoda tersebut. Pada teorema Bayes, bila terdapat dua kejadian yang terpisah (misalkan A dan B), maka teorema Bayes dirumuskan dalam Persamaan (1).

$$P(A|B) = \frac{P(A)}{P(B)} P(B|A) \quad (1)$$

Teorema Bayes sering pula dikembangkan mengingat berlakunya hukum probabilitas total, menjadi seperti pada Persamaan (2).

$$P(A|B) = \frac{P(A)P(B|A)}{\sum_{i=1}^n P(A|B)} \quad (2)$$

dimana $A_1 \cup A_2 \cup \dots \cup A_n = S$

Untuk menjelaskan teorema Naïve Bayes, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema Bayes di atas disesuaikan dalam persamaan (3).

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)} \quad (3)$$

Dimana variabel C merepresentasikan kelas, sementara variabel $F_1 \dots F_n$ merepresentasikan karakteristik karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel dengan karakteristik tertentu dalam kelas C (*posterior*) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik - karakteristik sampel pada kelas C (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga *evidence*). Karena itu, Persamaan (3) dapat pula ditulis secara sederhana dalam Persamaan (4).

$$Posterior = \frac{prior \times likelihood}{evidence} \quad (4)$$

Nilai *evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai dari *posterior* tersebut yang nantinya akan dibandingkan dengan nilai-nilai *posterior* kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan [2].

3. HASIL DAN PEMBAHASAN

3.1 Ekstraksi Fitur

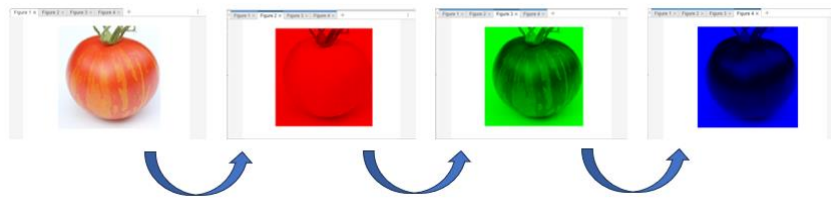
Fitur ruang warna RGB adalah fitur yang digunakan dalam penelitian ini. Ekstraksi fitur ini menghasilkan tiga parameter yaitu *Red*, *Green* dan *Blue*. Algoritma 1 menunjukkan perintah matlab untuk ekstraksi fitur. Setelah dilakukan ekstraksi fitur citra asli berubah ke citra RGB seperti disajikan pada Gambar 3.

```

clc; clear; close all;
I = imread('RZ_013.png');
R = I(:,:,1);
G = I(:,:,2);
B = I(:,:,3);
Red = cat(3,R,G*0,B*0);
Green = cat(3,R*0,G,B*0);
Blue = cat(3,R*0,G*0,B);

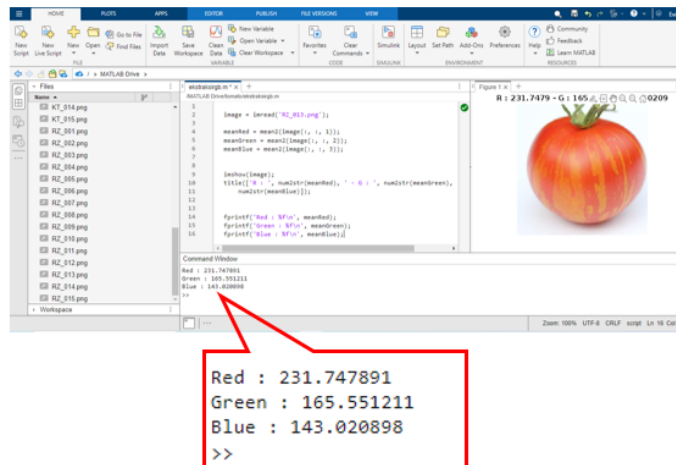
figure, imshow(I);
figure, imshow(Red);
figure, imshow(Green);
figure, imshow(Blue);
    
```

Algoritma 1. Perintah Matlab untuk Ekstraksi Fitur



Gambar 3. Transformasi dari Citra Asli ke RGB

Data citra diproses dengan Matlab dan diperoleh nilai RGB dari citra tersebut seperti ditunjukkan Gambar 4.



Gambar 4. Nilai RGB dari Citra

3.2 Pre-Processing Data

Berdasarkan hasil dari ekstraksi diperoleh nilai R, G, dan B yang ditunjukkan dengan angka desimal diikuti 6 angka di belakang koma seperti ditunjukkan dalam Gambar 4. Tahapan selanjutnya adalah dilakukan pembulatan 2 angka di belakang koma. Hasil data citra yang diperoleh dari ekstraksi fitur disajikan pada Tabel 1.

Tabel 1. Hasil Ekstraksi Fitur

Citra	R	G	B	Jenis
GZ_001.png	167.02	182.27	115.46	Green Zebra
GZ_002.png	149.10	161.16	87.23	Green Zebra
GZ_003.png	133.30	163.21	89.80	Green Zebra
GZ_004.png	106.13	141.75	84.69	Green Zebra
GZ_005.png	181.51	178.46	116.12	Green Zebra
GZ_006.png	93.16	102.04	38.68	Green Zebra
GZ_007.png	119.76	139.69	96.68	Green Zebra
GZ_008.png	115.64	118.17	63.12	Green Zebra
GZ_009.png	98.18	122.89	64.05	Green Zebra
GZ_010.png	129.82	142.32	53.06	Green Zebra
GZ_011.png	111.30	134.01	78.92	Green Zebra
GZ_012.png	115.09	133.55	78.43	Green Zebra
GZ_013.png	121.45	117.15	57.06	Green Zebra
GZ_014.png	118.65	132.44	43.68	Green Zebra
GZ_015.png	92.55	128.54	75.46	Green Zebra
KT_001.png	151.17	106.78	87.45	Kumato
KT_002.png	175.88	162.24	152.94	Kumato
KT_003.png	184.35	166.20	158.21	Kumato
KT_004.png	185.08	168.41	150.02	Kumato
KT_005.png	139.77	124.82	115.36	Kumato
KT_006.png	185.92	171.60	157.81	Kumato

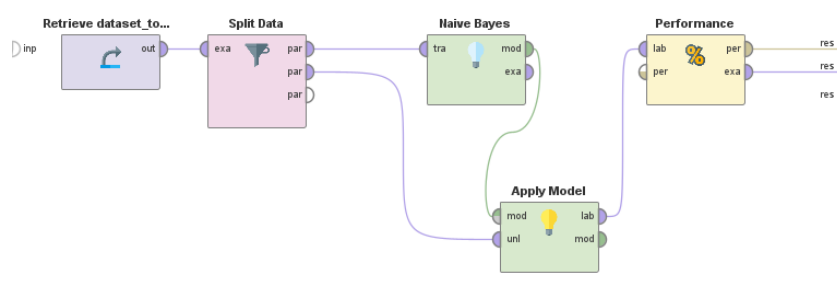
Citra	R	G	B	Jenis
KT_007.png	186.45	178.81	167.10	Kumato
KT_008.png	192.53	165.74	142.79	Kumato
KT_009.png	191.58	181.84	176.59	Kumato
KT_010.png	180.42	170.36	160.78	Kumato
KT_011.png	111.04	89.22	63.23	Kumato
KT_012.png	194.35	159.62	138.50	Kumato
KT_013.png	172.92	157.05	148.45	Kumato
KT_014.png	187.68	172.07	155.63	Kumato
KT_015.png	196.95	182.02	168.15	Kumato
RZ_001.png	177.76	122.66	96.63	Red Zebra
RZ_002.png	190.19	93.87	49.31	Red Zebra
RZ_003.png	136.58	83.48	55.67	Red Zebra
RZ_004.png	125.57	73.89	38.25	Red Zebra
RZ_005.png	152.59	58.97	54.70	Red Zebra
RZ_006.png	136.46	105.71	57.76	Red Zebra
RZ_007.png	164.36	73.57	52.98	Red Zebra
RZ_008.png	192.85	112.00	87.85	Red Zebra
RZ_009.png	153.68	84.17	65.57	Red Zebra
RZ_010.png	220.67	64.74	33.01	Red Zebra
RZ_011.png	146.75	105.65	52.37	Red Zebra
RZ_012.png	161.30	91.37	46.88	Red Zebra
RZ_013.png	231.75	165.55	143.02	Red Zebra
RZ_014.png	157.84	85.90	31.66	Red Zebra
RZ_015.png	202.39	101.48	50.74	Red Zebra

3.3 Pemrosesan Data

Setelah diperoleh hasil ekstraksi fitur seperti ditunjukkan dalam Tabel 1 dilakukan pemrosesan data. Proses klasifikasi dilakukan pada tahap pemrosesan data dengan algoritma Naïve Bayes. Validitas data diuji dengan mengevaluasi fitur yang dipilih secara teoritis sesuai dengan konsep yang diukur. Sedangkan pengukuran berulang (*repeated measurements*) dengan cara melakukan cek hasil pengukuran yang sama pada waktu yang berbeda untuk memastikan reabilitas data. Data yang sudah diproses sebelumnya dibagi menjadi data latih dan data uji sebelum ke prosedur klasifikasi. Operator *split data* digunakan untuk proses pembagian data. Rasio yang digunakan adalah 70%:30%.

3.4 Pengujian

Pembagian data uji dan data latih secara proporsional untuk set pelatihan dan set pengujian merupakan cara untuk menguji hasil klasifikasi. Penelitian ini menggunakan metode pengujian *percentage split* dengan pembagian data sebesar 70% untuk data latih dan 30% untuk data uji. Hasil pengujian menggunakan Algoritma Naïve Bayes untuk Klasifikasi Tomat menunjukkan hasil akurasi 78.57%.



Gambar 5. Tampilan Rapid Miner menggunakan Klasifikasi Naïve Bayes dengan Atribut R, G dan B

Gambar 5 menyajikan Rancangan Naïve Bayes dengan Atribut R, G, dan B. Rancangan tersebut menghasilkan Data Pengujian dengan atribut R, G, dan B seperti ditunjukkan pada Tabel 2.

Tabel 2. Data Hasil Pengujian dengan Atribut R, G dan B

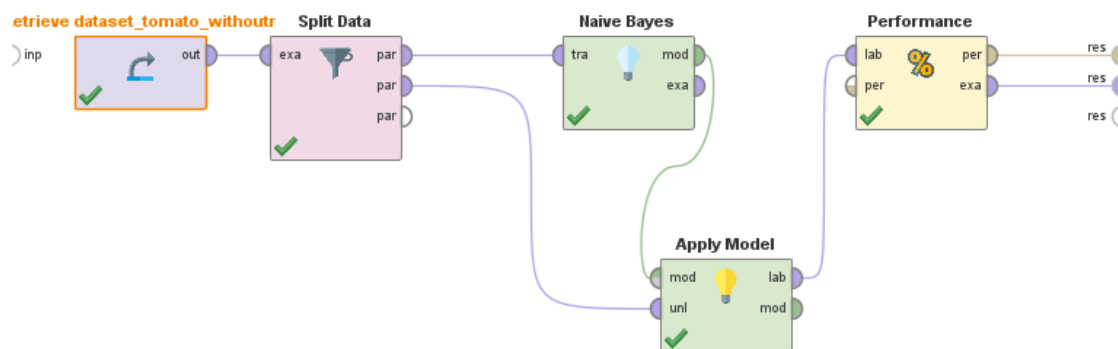
Row No.	Jenis	Prediksi	Confidence	Confidence	Confidence	Citra	R	G	B
1	Green Zebra	Kumato	0.013	0.974	0.013	GZ_001.png	167.023	182.266	115.464
2	Green Zebra	Red Zebra	0.444	0.000	0.556	GZ_006.png	93.158	102.036	38.677
3	Green Zebra	Green Zebra	0.944	0.020	0.036	GZ_007.png	119.762	139.691	96.682
4	Green Zebra	Green Zebra	0.906	0.0002	0.092	GZ_010.png	129.822	142.324	53.060
5	Kumato	Red Zebra	0.427	0.035	0.539	KT_001.png	151.171	106.784	87.450
6	Kumato	Kumato	0.000	0.997	0.003	KT_002.png	175.876	162.239	152.944
7	Kumato	Kumato	0.000	1.000	0.000	KT_009.png	191.577	181.844	176.594
8	Kumato	Kumato	0.000	0.999	0.001	KT_010.png	180.423	170.361	160.782
9	Red Zebra	Red Zebra	0.173	0.303	0.524	RZ_001.png	177.757	122.661	96.626
10	Red Zebra	Red Zebra	0.004	0.001	0.995	RZ_002.png	190.195	93.868	49.314
11	Red Zebra	Red Zebra	0.071	0.000	0.929	RZ_003.png	136.580	83.480	55.674
12	Red Zebra	Red Zebra	0.040	0.110	0.851	RZ_008.png	192.849	111.999	87.851
13	Red Zebra	Red Zebra	0.045	0.002	0.953	RZ_009.png	153.677	84.170	65.570
14	Red Zebra	Red Zebra	0.000	0.000	1.000	RZ_015.png	202.386	101.476	50.742

Tabel 3. Hasil Akurasi dengan Atribut R, G dan B

Accuracy: 78.57%

	true Green Zebra	true Kumato	true Red Zebra	class precision
pred. Green Zebra	2	0	0	100.00%
pred. Kumato	1	3	0	75%
pred. Red Zebra	1	1	6	75%
class recall	50.00%	75.00%	100.00%	

Berdasarkan Pengujian dengan atribut R, G dan B, akurasi paling tinggi adalah pada kultivar tomat Red Zebra dengan nilai akurasi 100%. Untuk kultivar Kumato memiliki nilai akurasi 75% dan paling rendah untuk kultivar Green Zebra dengan nilai akurasi 50% seperti disajikan pada Tabel 3. Setelah dilakukan pengujian dengan ketiga atribut R, G dan B secara menyeluruh maka dilakukan pengujian ulang dengan menghilangkan atribut R, kemudian menghilangkan atribut G dan terakhir menghilangkan atribut B untuk melihat pengaruh setiap atribut dalam akurasi hasil. Percobaan pertama yang dilakukan adalah menghilangkan atribut Red (R) dengan rancangan disajikan pada Gambar 6, dan didapatkan nilai akurasi menurun menjadi 64.29%. Jika atribut yang digunakan G dan B maka akurasi paling tinggi adalah untuk kultivar tomat Kumato sebesar 75% seperti ditunjukkan pada Tabel 5. Jenis kultivar ini memiliki warna merah keunguan dan tidak mengandung warna hijau. Rancangan Gambar 6 menghasilkan data pengujian tanpa atribut R seperti ditampilkan dalam Tabel 4.



Gambar 6. Tampilan Rapid Miner menggunakan Naïve Bayes Tanpa Atribut R

Tabel 4. Data Hasil Pengujian Tanpa Atribut R

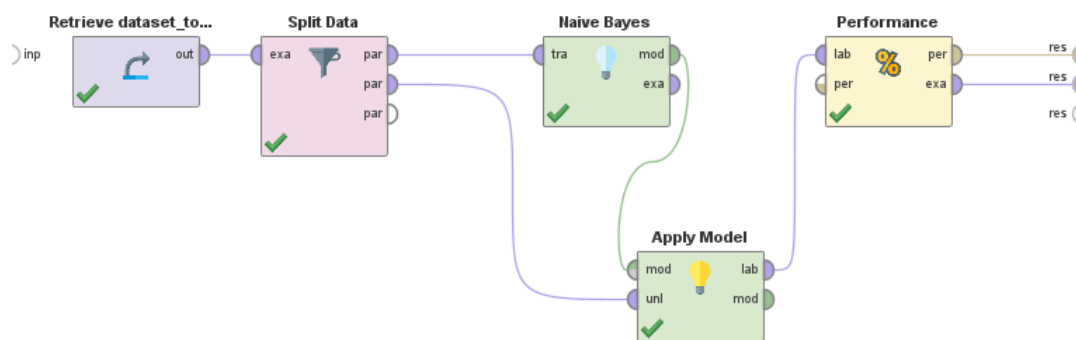
Row No.	Jenis	Prediksi	Confidence	Confidence	Confidence	Citra	R	G
1	Green Zebra	Kumato	0.055	0.929	0.015	GZ_001.png	182.266	115.464
2	Green Zebra	Red Zebra	0.087	0.000	0.913	GZ_006.png	102.036	38.677
3	Green Zebra	Green Zebra	0.733	0.154	0.113	GZ_007.png	139.691	96.682
4	Green Zebra	Green Zebra	0.781	0.008	0.211	GZ_010.png	142.324	53.069
5	Kumato	Red Zebra	0.468	0.032	0.499	KT_001.png	106.784	87.450
6	Kumato	Kumato	0.000	0.996	0.004	KT_002.png	162.239	152.944
7	Kumato	Kumato	0.000	1.000	0.000	KT_009.png	181.844	176.594
8	Kumato	Kumato	0.000	0.998	0.002	KT_010.png	170.361	160.782
9	Red Zebra	Green Zebra	0.653	0.104	0.243	RZ_001.png	122.661	96.626
10	Red Zebra	Red Zebra	0.084	0.001	0.916	RZ_002.png	93.868	49.314
11	Red Zebra	Red Zebra	0.038	0.001	0.962	RZ_003.png	83.480	55.674
12	Red Zebra	Green Zebra	0.570	0.039	0.392	RZ_008.png	111.999	87.851
13	Red Zebra	Red Zebra	0.061	0.001	0.937	RZ_009.png	84.170	65.570
14	Red Zebra	Red Zebra	0.006	0.000	0.994	RZ_015.png	101.476	50.742

Tabel 5. Hasil Akurasi Tanpa Atribut R

Accuracy: 64.29 %

	true Green Zebra	true Kumato	true Red Zebra	class precision
pred. Green Zebra	2	0	2	50.00%
pred. Kumato	1	3	0	75%
pred. Red Zebra	1	1	4	66.67%
class recall	50.00%	75.00%	66.67%	

Percobaan selanjutnya dilakukan dengan menghilangkan atribut *Green* (G) dengan rancangan naïve bayes ditunjukkan pada Gambar 7. dan hasil nilai akurasi adalah 71.43%. Data hasil pengujian tanpa atribut G tersaji pada Tabel 6. Pada percobaan atribut yang digunakan R dan B akurasi paling tinggi pada kultivar Red Zebra nilai akurasi mencapai 100% seperti ditampilkan pada Tabel 7. Kultivar ini memiliki warna dominan merah dengan sedikit semburat kuning.



Gambar 7. Tampilan Rapid Miner menggunakan Naïve Bayes Tanpa Atribut G

Tabel 6. Data Hasil Pengujian Tanpa Atribut G

Row No.	Jenis	Prediction	confidence	confidence	confidence	Citra	R	B
1	Green Zebra	Kumato	0.048	0.727	0.224	GZ_001.png	167.023	115.464
2	Green Zebra	Green Zebra	0.733	0.000	0.267	GZ_006.png	93.158	38.677
3	Green Zebra	Green Zebra	0.851	0.031	0.119	GZ_007.png	119.762	96.682
4	Green Zebra	Green Zebra	0.707	0.002	0.291	GZ_010.png	129.822	53.060
5	Kumato	Green Zebra	0.575	0.099	0.327	KT_001.png	151.171	87.450
6	Kumato	Kumato	0.000	0.975	0.025	KT_002.png	175.876	152.944
7	Kumato	Kumato	0.000	0.993	0.007	KT_009.png	191.577	176.594
8	Kumato	Kumato	0.000	0.984	0.016	KT_010.png	180.423	160.782

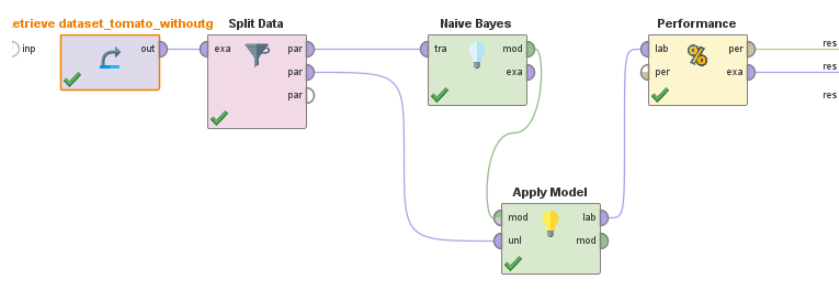
Row No.	Jenis	Prediction	confidence	confidence	confidence	Citra	R	B
9	Red Zebra	Red Zebra	0.105	0.409	0.486	RZ_001.png	177.757	96.626
10	Red Zebra	Red Zebra	0.029	0.013	0.958	RZ_002.png	190.195	49.314
11	Red Zebra	Green Zebra	0.667	0.004	0.329	RZ_003.png	136.580	55.674
12	Red Zebra	Red Zebra	0.045	0.276	0.679	RZ_008.png	192.849	87.851
13	Red Zebra	Green Zebra	0.524	0.022	0.454	RZ_009.png	153.677	65.570
14	Red Zebra	Red Zebra	0.000	0.000	0.999	RZ_015.png	202.386	50.742

Tabel 7. Hasil Akurasi Tanpa Atribut G

Accuracy: 71.43 %

	true Green Zebra	true Kumato	true Red Zebra	class precision
pred. Green Zebra	3	1	2	50.00%
pred. Kumato	1	3	0	75.00%
pred. Red Zebra	0	0	4	100.00%
class recall	75.00%	75.00%	66.67%	

Percobaan selanjutnya dilakukan dengan menghilangkan atribut *Blue* (B) sesuai dengan rancangan yang ditampilkan pada Gambar 8 dengan hasil pengujian yang disajikan pada Tabel 8. Hasil nilai akurasi adalah paling tinggi dibanding jika tanpa atribut R dan G, yaitu dengan nilai akurasi 85.71%. Pada percobaan atribut yang digunakan R dan G akurasi paling tinggi pada kultivar Red Zebra nilai akurasi mencapai 100% seperti yang tersaji pada Tabel 9. Kultivar ini memiliki warna dominan hijau dengan sedikit semburat kuning.



Gambar 8. Rancangan Naive Bayes Tanpa Atribut B

Tabel 8. Data Hasil Pengujian Tanpa Atribut B

Row No.	Jenis	Prediksi	confidence	confidence	confidence	Citra	R	B
1	Green Zebra	Kumato	0.043	0.926	0.032	GZ_001.png	167.023	182.266
2	Green Zebra	Green Zebra	0.700	0.005	0.295	GZ_006.png	93.158	102.036
3	Green Zebra	Green Zebra	0.886	0.053	0.061	GZ_007.png	119.762	139.691
4	Green Zebra	Green Zebra	0.811	0.112	0.077	GZ_010.png	129.822	142.324
5	Kumato	Red Zebra	0.235	0.133	0.632	KT_001.png	151.171	106.784
6	Kumato	Kumato	0.065	0.856	0.078	KT_002.png	175.876	162.239
7	Kumato	Kumato	0.005	0.962	0.032	KT_009.png	191,577	181,844
8	Kumato	Kumato	0.030	0.917	0.053	KT_010.png	180.423	170.361
9	Red Zebra	Red Zebra	0.087	0.434	0.479	RZ_001.png	177,757	122,661
10	Red Zebra	Red Zebra	0.005	0.091	0.905	RZ_002.png	190.195	93,868
11	Red Zebra	Red Zebra	0.065	0.021	0.915	RZ_003.png	136,580	83,430
12	Red Zebra	Red Zebra	0.015	0.285	0.700	RZ_008.png	192,849	111,999
13	Red Zebra	Red Zebra	0.027	0.033	0.940	RZ_009.png	153,677	84,170
14	Red Zebra	Red Zebra	0.000	0.004	0.995	RZ_015.png	202.386	101.476

Tabel 9. Hasil Akurasi Tanpa Atribut B

Accuracy: 85.71%

	true Green Zebra	true Kumato	true Red Zebra	class precision
pred. Green Zebra	3	0	0	100.00%
pred. Kumato	1	3	0	75.00%
pred. Red Zebra	0	1	6	85.71%
class recall	75.00%	75.00%	100.00%	

Hasil Klasifikasi Tomat dengan Ekstraksi Fitur RGB dan Algoritma Naïve Bayes memiliki nilai akurasi yang cukup yaitu 78.57%. Nilai akurasi ini berubah ketika beberapa atribut dihilangkan. Atribut yang paling berpengaruh dalam akurasi klasifikasi tomat adalah atribut *Red* (R). Jika atribut ini dihilangkan nilai akurasi menjadi sangat rendah dengan nilai 64.29%. Atribut *Green* (G) juga memiliki pengaruh jika atribut ini dihilangkan. Atribut yang memberikan pengaruh terbesar terhadap peningkatan akurasi adalah kombinasi atribut R dan G. Ketika atribut B dihilangkan, nilai akurasi justru meningkat menjadi 85.71%. Atribut R, G dan B memiliki pengaruh dalam meningkatkan dan menurunkan nilai akurasi. Percobaan dengan menghilangkan salah satu atribut juga mempengaruhi hasil pada setiap kultivar. Tabel 10 menunjukkan hasil percobaan pengaruh dari setiap atribut pada akurasi klasifikasi kultivar tomat.

Tabel 10. Hasil Percobaan Pengaruh Atribut R, G dan B dalam Klasifikasi Kultivar Tomat

Atribut	Kultivar dengan Nilai Akurasi Tertinggi	Nilai Akurasi
RGB	Red Zebra	100%
GB	Kumato	75%
RB	Red Zebra	100%
RG	Red Zebra	100%

4. KESIMPULAN

Penelitian ini bertujuan untuk mengatasi permasalahan dalam mengklasifikasikan kultivar tomat yang memiliki kesamaan bentuk tetapi warna yang berbeda. Dengan menggunakan Ekstraksi Fitur RGB dan Algoritma Naïve Bayes, penelitian ini berhasil mengidentifikasi pengaruh atribut warna terhadap akurasi klasifikasi. Hasil penelitian menunjukkan bahwa atribut warna merah (R) memiliki pengaruh paling besar terhadap akurasi klasifikasi, sementara atribut warna biru (B) dapat meningkatkan akurasi hingga 85.71% ketika dikeluarkan. Metode klasifikasi yang digunakan berhasil mencapai akurasi hingga 78.57% dengan atribut penuh (R, G, B). Perlu dilakukan penelitian lebih lanjut dengan menambah data latih dan data uji untuk menguji kebenaran pengaruh atribut warna dalam klasifikasi tomat. Selain itu meskipun dalam beberapa penelitian algoritma Naïve Bayes memiliki akurasi yang tinggi dan disebut sebagai *classifier* terbaik, tetapi dalam penelitian ini algoritma Naïve Bayes tidak memiliki nilai akurasi yang tinggi sehingga perlu dilakukan penelitian lebih lanjut untuk menguji algoritma terbaik untuk klasifikasi kultivar tomat.

Kesimpulan ini menunjukkan bahwa metode ini dapat mendukung pengembangan sistem klasifikasi otomatis untuk mendukung pertanian modern, khususnya dalam mempercepat identifikasi varietas tomat secara akurat. Penerapan teknologi ini dapat membantu petani mengoptimalkan hasil budidaya melalui pengelolaan varietas yang lebih tepat.

DAFTAR PUSTAKA

- [1] F. Moh, M. B. Sulthan, and I. Wahyudi, "Klasifikasi Kematangan Buah Tomat Menggunakan Fitur RGB dan HSI Berbasis Backpropagation," *Jurnal Aplikasi Teknologi Informasi dan Manajemen (JATIM)*, vol. 4, no. 1, pp. 84–95, 2023.
- [2] R. G. de Luna, E. P. Dadios, A. A. Bandala, and R. R. P. Vicerra, "Size classification of tomato fruit using thresholding, machine learning and deep learning techniques," *Agrivita*, vol. 41, no. 3, pp. 586–596, 2019, doi: 10.17503/agrivita.v4i3.2435.
- [3] A. P. Wibawa, M. Guntur, A. Purnama, M. Fathony Akbar, and F. A. Dwiyanto, "Metode-metode Klasifikasi," *Prosiding Seminar Ilmu Komputer dan Teknologi Informasi*, vol. 3, no. 1, pp. 134–138, 2018.
- [4] S. Aprilisa and Sukemi, "Klasifikasi Tingkat Kematangan Buah Tomat Berdasarkan Fitur Warna Menggunakan K-Nearest Neighbor," in *Prosiding Annual Research Seminar*, 2019, pp. 170–173.
- [5] J. Basavaiah and A. Arlene Anthony, "Tomato Leaf Disease Classification using Multiple Feature Extraction Techniques," *Wirel Pers Commun*, vol. 115, no. 1, pp. 633–651, Nov. 2020, doi: 10.1007/s11277-020-07590-x.
- [6] R. Soekarta, N. Nurdjan, and A. Syah, "Klasifikasi Penyakit Tanaman Tomat Menggunakan Metode Convolutional Neural Network (CNN)," *INSECT*, vol. 8, no. 2, pp. 143–151, 2023.
- [7] U. Khultsum and A. Subekti, "Penerapan Algoritma Random Forest dengan Kombinasi Ekstraksi Fitur Untuk Klasifikasi Penyakit Daun Tomat," *Jurnal Media Informatika Budidarma*, vol. 5, no. 1, pp. 186–193, Jan. 2021, doi: 10.30865/mib.v5i1.2624.
- [8] L. Tan, J. Lu, and H. Jiang, "Tomato Leaf Diseases Classification Based on Leaf Images: A Comparison between Classical Machine Learning and Deep Learning Methods," *AgriEngineering*, vol. 3, no. 3, pp. 542–558, Sep. 2021, doi: 10.3390/agriengineering3030035.
- [9] V. Maeda-Gutiérrez *et al.*, "Comparison of convolutional neural network architectures for classification of tomato plant diseases," *Applied Sciences (Switzerland)*, vol. 10, no. 4, Feb. 2020, doi: 10.3390/app10041245.

- [10] A. J. Bastari and A. Cherid, “Klasifikasi Penyakit Tanaman Tomat Menggunakan Convolutional Neural Network dan Implementasi Model H5 Pada Aplikasi Desktop,” *SIMKOM*, vol. 8, no. 2, pp. 199–207, Aug. 2023, doi: 10.51717/simkom.v8i2.194.
- [11] M. A. Alajrami and S. S. Abu-Naser, “Type of Tomato Classification Using Deep Learning,” *International Journal of Academic Pedagogical Research*, vol. 3, no. 12, pp. 21–25, 2019.
- [12] O. Belitskaya, “Tomato Cultivars,” *Kaggle*. Accessed: Jan. 15, 2024. [Online]. Available: <https://www.kaggle.com/datasets/olgabelitskaya/tomato-cultivars/data>
- [13] L. Farokhah and P. Korespondensi, “Implementasi K-Nearest Neighbor untuk Klasifikasi Bunga dengan Ekstraksi Fitur WarnaRGB,” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 7, no. 6, pp. 1129–1135, 2020, doi: 10.25126/jtiik.202072608.
- [14] A. Susanto and I. U. Wahyu Mulyono, “A Good Accuracy in Apple Fruits Quality Based on Back Propagation Neural Network and Feature Extraction,” *Journal of Informatics and Telecommunication Engineering*, vol. 6, no. 1, pp. 38–48, Jul. 2022, doi: 10.31289/jite.v6i1.6938.
- [15] N. P. Batubara, D. Widiyanto, and N. Chamidah, “Klasifikasi Rempah Rimpang Berdasarkan Ciri Warna RGB dan Tekstur GLCM Menggunakan Algoritma Naive Bayes,” *Jurnal INFORMATIK*, vol. 16, no. 3, pp. 156–163, 2020.
- [16] R. Maneno, B. Baso, P. G. Manek, and K. Fallo, “Deteksi Tingkat Kematangan Buah Pinang Menggunakan Metode Support Vector Machine Berdasarkan Warna Dan Tekstur,” *Journal of Information and Technology*, vol. 3, no. 2, pp. 60–66, Sep. 2023, doi: 10.32938/jitu.v3i2.5323.
- [17] M. Astiningrum, P. Prima Arhandi, and N. Aqmarina Ariditya, “Identifikasi Penyakit Pada Daun Tomat Berdasarkan Fitur Warna dan Tekstur,” *JIP (Jurnal Informatika Polinema)*, vol. 6, no. 2, pp. 47–50, 2020.
- [18] S. Sanjaya, M. L. Pura, S. K. Gusti, F. Yanto, and F. Syafria, “K-Nearest Neighbor for Classification of Tomato Maturity Level Based on Hue, Saturation, and Value Colors,” *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 2, no. 2, pp. 101–106, Nov. 2019, doi: 10.24014/ijaidm.v2i2.7975.
- [19] R. N. Devita, H. W. Herwanto, and A. P. Wibawa, “Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa Indonesia,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 4, pp. 427–434, Oct. 2018, doi: 10.25126/jtiik.201854773.