



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
- a. Pengutipan hanya untuk keperluan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

IDENTIFIKASI KEMATANGAN TOMAT DENGAN PRINCIPAL COMPONENT ANALYSIS DAN K-NEAREST NEIGHBOUR BERDASARKAN CITRA WARNA

KHAIRANI



**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Identifikasi Kematangan Tomat dengan *Principal Component Analysis* dan *K-Nearest Neighbour* Berdasarkan Citra Warna” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2024

Khairani
G6501202047



KHAIRANI. Identifikasi Kematangan Tomat dengan *Principal Component Analysis* dan *K-Nearest Neighbour* Berdasarkan Citra Warna. Dibimbing oleh IMAS SUKAESI SITANGGANG, TOTO HARYANTO dan AZIZ KUSTIYO.

Tomat (*Lycopersicon esculentum Mill*) merupakan tanaman yang masuk dalam kategori sayuran famili *Solanaceae* yang banyak dijumpai di daerah tropis dan subtropis. Salah satu jenis tomat yang banyak ditanam adalah tomat plum (*Lycopersicon lycopersicum*). Di Indonesia sendiri permintaan kebutuhan buah tomat plum semakin meningkat di pasaran, namun kebutuhan tersebut belum dapat terpenuhi dikarenakan berbagai faktor. Salah satu faktornya adalah penentuan tingkat kematangan buah tomat yang masih dilakukan dengan proses manual. Cara manual ini memiliki beberapa kelemahan, diantaranya proses penentuan kematangan yang cukup memakan waktu, memiliki akurasi rendah dan tidak konsisten. Hal ini dikarenakan standar penentuan tingkat kematangan dilakukan secara subjektif oleh para petani, sehingga dibutuhkan sebuah sistem yang dapat mengklasifikasi kematangan buah tomat dengan baik dan minim kesalahan klasifikasi.

Berdasarkan permasalahan di atas, dalam penelitian ini menggusulkan fitur warna *Hue Saturation Value* (HSV) dan algoritma *Principal Component Analysis* (PCA) untuk mereduksi citra sehingga didapatkan *project image* dari citra buah tomat tersebut. Penelitian ini melakukan proses pembobotan sebelum dilakukan proses klasifikasi. Proses pembobotan pada penelitian ini menggunakan teknik *score calculation*, sehingga didapatkan bobot terbaik untuk proses pemodelan. Selanjutnya dilakukan proses klasifikasi untuk menentukan tingkat kematangan buah tomat yang terdiri dari *green*, *turning*, *pink*, *light red* dan *red* menggunakan algoritma *K-Nearest Neighbor* (KNN) dengan memanfaatkan nilai ketetanggan (*k*).

Penelitian ini menggunakan data sekunder yang didapat pada penelitian sebelumnya. Total 400 citra yang digunakan dengan ukuran 400×400 piksel yang dikelompokkan dalam 5 tingkat kematangan yaitu *green*, *turning*, *pink*, *light red* dan *red*. Penelitian ini menggunakan 20 buah tomat untuk setiap kelas dengan pengambilan gambar dilakukan pada 4 sisi buah, sehingga yang digunakan setiap kelas adalah 80 citra.

Data dibagi menjadi data latih dan data uji yang terbagi atas 70%-30%, 80%-20% dan 90%-10%. Pada penelitian ini ditentukan nilai *k* yaitu 1, 3, 5, 7, 9 dan 11. Nilai PC pada PCA juga ditentukan berdasarkan varian rasio 95% yaitu 5, 10, 15,...,65. Terdapat 7 skenario yang ditetapkan sebagai data dari PCA ke KNN berdasarkan ruang warna yaitu *Hue* (H), *Saturation* (S), *Value* (V), *Hue-Saturation* (HS), *Hue-Value* (HV), *Saturation-Value* (SV) dan HSV.

Berdasarkan hasil pengujian yang dilakukan, dari tiga kelompok pembagian data *Hue-Value* adalah skenario yang mendapat nilai akurasi tertinggi yaitu pada pembagian data 70%-30% akurasi tertinggi adalah 91%, pembagian data 80%-20% akurasi tertinggi adalah 94% dan pada pembagian data 90%-10% akurasi data adalah 93%. Pada penelitian ini hasil varian rasio yang melebihi 50% data pada tiap data HSV pada PC 5, ini membuktikan bahwa PC 5 telah mewakili



dari keseluruhan data. dari keseluruhan pengujian yang dilakukan, hasil tertinggi didapat pada pembagian data 80-20%. dengan akurasi tertinggi 94%.

Kata kunci: *hue saturation value, klasifikasi, k-nearest neighbor, principal component analysis, tomat.*

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk keperluan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



KHAIRANI. *Identification on Tomato Ripeness using Principal Component Alalysis and K-Nearest Neighbor Based on Color Image. Supervised by IMAS SUKAESIH SITANGGANG, TOTO HARYANTO and AZIZ KUSTIYO.*

SUMMARY

KHAIRANI. Identification on Tomato Ripeness using Principal Component Alalysis and K-Nearest Neighbor Based on Color Image. Supervised by IMAS SUKAESIH SITANGGANG, TOTO HARYANTO and AZIZ KUSTIYO.

*Tomato (*Lycopersicon esculentum* Mill) is a plant belonging to the vegetable category within the Solanaceae family, often found in tropical and subtropical areas. One widely grown type of tomato is the plum tomato (*Lycopersicon lycopersicum*). In Indonesia, there is a growing demand for plum tomatoes in the market, but it is currently not being met due to various factors. One of these factors is the manual method of determining the ripeness of tomatoes. This manual process has several weaknesses, including being time-consuming, having low accuracy, and lacking consistency in determining maturity levels. This is because the standards for determining maturity levels are carried out subjectively by farmers. Therefore, a system is needed to accurately classify the ripeness of tomatoes and minimize misclassification.*

This study utilized the Hue Saturation Value (HSV) color feature and the Principal Component Analysis (PCA) algorithm to compress tomato images. The process involved assigning weights through score calculation techniques and using the K-Nearest Neighbor (KNN) algorithm to classify tomato ripeness levels. We used 400 images of tomatoes categorized into five maturity levels: green, turning, pink, light red, and red, with each category containing 20 tomatoes and 80 images.

The data in this study split were into training and test sets, with divisions of 70%-30%, 80%-20%, and 90%-10%. This study tested k values of 1, 3, 5, 7, 9, and 11, and PC values ranging from 5 to 65 based on a 95% Variant Ratio. Seven scenarios combining PCA and KNN data are defined based on color space: Hue (H), Saturation (S), Value (V), Hue-Saturation (HS), Hue-Value (HV), Saturation-Value (SV), and HSV.

The results revealed the highest accuracy levels in the 70%-30%, 80%-20%, and 90%-10% data divisions. Notably, the data division of 80-20% yielded the highest accuracy of 94%. Moreover, PCA demonstrated that PC 5 was representative of the entire data as it exceeded 50% of the data for each HSV data.

Keywords: Classification, Hue Saturation Value, K-Nearest Neighbor, Principal Component Analysis, Tomato.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta Milik IPB, tahun 2024
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



IDENTIFIKASI KEMATANGAN TOMAT DENGAN PRINCIPAL COMPONENT ANALYSIS DAN K-NEAREST NEIGHBOUR BERDASARKAN CITRA WARNA

KHAIRANI

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister pada
Program Studi Ilmu Komputer

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**

- Hak Cipta Dilindungi Undang-undang
1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

@Hak cipta milik IPB University

IPB University

@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk keperluan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Tesis : Identifikasi Kematangan Tomat dengan *Principal Component Analysis* dan *K-Nearest Neighbour* Berdasarkan Citra Warna.
Nama : Khairani
NIM : G6501202047

Disetujui oleh



Pembimbing 1:

Prof. Dr. Imas Sukaesih Sitanggang, S.Si., M.Kom.



Pembimbing 2:

Dr. Toto Haryanto, S.Kom., M.Si.



Pembimbing 3:

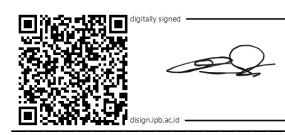
Dr. Aziz Kustiyo, S.Si, M.Kom.

Diketahui oleh



Ketua Program Studi:

Dr. Toto Haryanto, S.Kom., M.Si.
NIP 198211172014041001



Dekan Fakultas Matematika dan Ilmu Pengetahuan
Alam:

Dr. Berry Juliandi, S.Si., M.Si.
NIP 19780723 200701 1 001

Tanggal Ujian:

29 Juli 2024

Tanggal Lulus:



Puji dan syukur penulis panjatkan kepada Allah *Subhanaahu wa Ta'ala* atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Desember 2023 sampai bulan Januari 2024 ini adalah klasifikasi ruang warna pada buah tomat dengan judul “Identifikasi Kematangan Tomat dengan *Principal Component Analysis* dan *K-Nearest Neighbour* Berdasarkan Citra Warna”

Terima kasih penulis ucapkan kepada para pembimbing, Prof. Dr. Imas Sukaesih Sitanggang, S.Si., M.Kom., Dr. Toto Haryanto, S.Kom., M.Si. dan Dr. Aziz Kustiyo, S.Si, M.Kom. yang telah membimbing dan banyak memberi saran. Ucapan terima kasih juga disampaikan kepada moderator seminar Dr. Erfiani, M.Si., moderator ujian tesis Dr. Sony Hartono Wijaya, S.Kom., M.Kom., dan penguji luar komisi pembimbing Hari Agung Adrianto, S.Kom., M.Si., Ph.D. Ungkapan terima kasih juga penulis sampaikan kepada Ibunda Nurchairati, adik kandung Aulia Rahman Fahmi, Fajrin Fahmi, dan Sandy Rizki Fahmi, ananda Audrey Abint Khairan, serta seluruh keluarga dan teman yang telah memberikan dukungan, doa, dan kasih sayang.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juli 2024

Khairani

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR TABEL

xiii

DAFTAR GAMBAR

xiii

DAFTAR LAMPIRAN

xiii

PENDAHULUAN

1

- 1.1 Latar Belakang
- 1.2 Rumusan Masalah
- 1.3 Tujuan
- 1.4 Manfaaat
- 1.5 Ruang Lingkup

1

2

3

3

3

TINJAUAN PUSTAKA

4

- 2.1 Tomat
- 2.2 Ruang Warna RGB
- 2.3 Ruang Warna Hue HSV
- 2.4 *Principal Component Analysis* (PCA)
- 2.5 K-Nearst Neighbor (KNN)

4

4

4

5

6

8

METODE

10

- 3.1 Praproses Data
 - 3.1.1 *Cropping*
 - 3.1.2 *Resize*
- 3.2 Pembagian Data
 - 3.2.1 Data latih
 - 3.2.2 Data Uji
- 3.3 *Hue Saturation Value* (HSV)
- 3.5 *Principal Component Analysis* (PCA)
- 3.5 Klasifikasi
- 3.6 Pengujian

10

10

11

11

11

11

11

12

13

15

HASIL DAN PEMBAHASAN

16

- 4.1 Persiapan Data
- 4.2 Ekstrasi warna HSV
- 4.3 Ekstrasi ciri PCA
- 4.4 Klasifikasi KNN
- 4.5 Pengujian

16

16

17

18

20

SIMPULAN DAN SARAN

23

- 5.1 Simpulan
- 5.2 Saran

23

23

DAFTAR PUSTAKA

24

RIWAYAT HIDUP

38

DAFTAR ISI



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR TABEL

1 Jumlah data latih yang digunakan pada setiap kelas.	11
2 Jumlah data uji yang digunakan pada setiap kelas.	11
3 <i>Confusion matrix</i>	15
4 Contoh nilai komponen H, S dan V	17
5 Nilai varian rasio dari jumlah PC dan proporsi data latih yang berbeda	17
6 <i>Confusion matrix</i> pada data latih 70% dengan akurasi tertinggi 91%	21
7 <i>Confusion matrix</i> pada data latih 90% dengan akurasi tertinggi 93%	21
8 <i>Confusion matrix</i> pada data latih 90% dengan akurasi tertinggi 93%.	21

DAFTAR GAMBAR

1 Klasifikasi data buah tomat (a) <i>Green</i> , (b) <i>Turning</i> , (c) <i>Pink</i> , (d) <i>Light Red</i> , (e) <i>Red</i> .	4
2 Konsep warna RGB	5
3 Ruang Warna HSV	6
4 Tahapan penelitian	10
5 Proses <i>cropping</i> buah tomat	10
6 Proses resize citra buah tomat	11
7 Perbandingan foto citra original dan citra HSV (a) original dan (b) HSV	12
8 <i>Flowchart</i> algoritma HSV	12
9 <i>Flowchart</i> ekstraksi data latih algoritma PCA	13
10 <i>Flowchart</i> data uji algoritma PCA	13

DAFTAR LAMPIRAN

1 Data Sekunder Tomat	28
2 Tabel Hasil Pengujian Klasifikasi	33