

SKRIPSI

KLASIFIKASI CITRA BUAH KOPI MENGGUNAKAN
ALGORITMA *K-NEAREST NEIGHBOR*



NATHANAEL NOVANTO

NPM: 6161901078

PROGRAM STUDI MATEMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
2024

FINAL PROJECT

**CLASSIFICATION OF COFFEE FRUIT IMAGES USING
K-NEAREST NEIGHBOR ALGORITHM**



NATHANAEL NOVANTO

NPM: 6161901078

**DEPARTMENT OF MATHEMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
2024**

LEMBAR PENGESAHAN

KLASIFIKASI CITRA BUAH KOPI MENGGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR*

Nathanael Novanto

NPM: 6161901078

Telah lulus ujian skripsi pada 16 Januari 2024 dengan penguji:
Dr. Julius Dharma Lesmono dan Dr. Livia Owen

Bandung, 22 Januari 2024

Menyetujui,

Pembimbing 1

Pembimbing 2

Agus Sukmana, M.Sc.

Liem Chin, M.Si.

Mengetahui,

Ketua Program Studi

Jonathan Hoseana, Ph.D.

PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

KLASIFIKASI CITRA BUAH KOPI MENGGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR*

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
22 Januari 2024

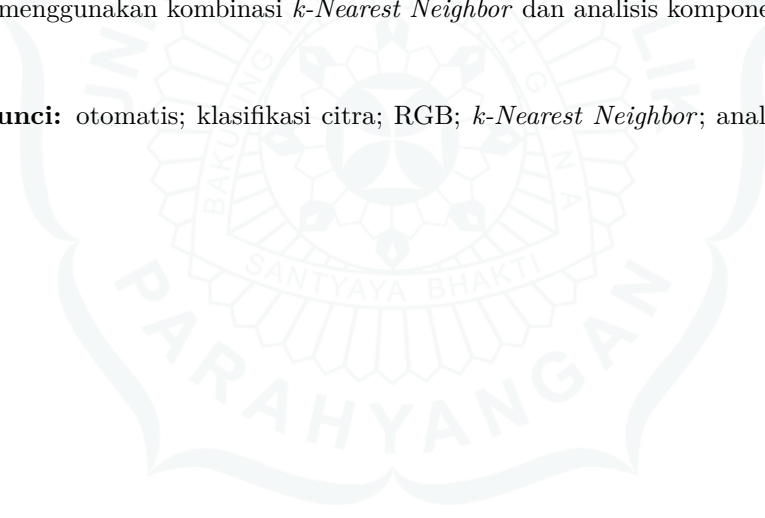


Nathanael Novanto
NPM: 6161901078

ABSTRAK

Proses pemeriksaan tingkat kematangan buah kopi biasanya dilakukan oleh manusia yang ahli di bidangnya. Akan tetapi hal ini memiliki kendala, antara lain terbatasnya tenaga ahli dan mahalnya upah untuk tenaga ahli. Oleh karena itu, akan dikembangkan metode agar dapat melakukan pemeriksaan tingkat kematangan buah secara otomatis. Salah satu penerapannya adalah dengan mengklasifikasi buah kopi yang akan dicek secara otomatis melalui citra buah kopi untuk setiap tingkat kematangan yang berbeda. Transformasi dari citra menjadi bentuk numerik, menggunakan pendekatan intensitas RGB (*Red, Green, Blue*), selanjutnya diklasifikasi menggunakan *k-Nearest Neighbor* dan analisis komponen utama. Terdapat dua klasifikasi yang dibandingkan, yaitu penggunaan algoritma *k-Nearest Neighbor* secara langsung dan kombinasi *k-Nearest Neighbor* dengan analisis komponen utama. Hasil penelitian menunjukkan klasifikasi dengan menggunakan *k-Nearest Neighbor* lebih akurat dan waktu komputasinya lebih singkat dibandingkan menggunakan kombinasi *k-Nearest Neighbor* dan analisis komponen utama.

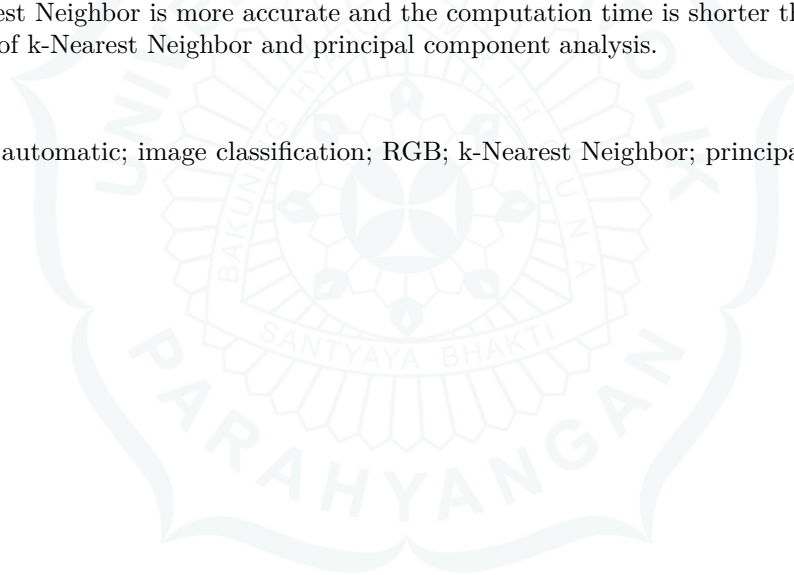
Kata-kata kunci: otomatis; klasifikasi citra; RGB; *k-Nearest Neighbor*; analisis komponen utama.



ABSTRACT

The process of checking the maturity level of coffee fruit is usually carried out by human experts in the field. However, this has obstacles, including the limited number of experts and the high wages for experts. Therefore, a method will be developed to automatically check the maturity level of the fruit. One application is to classify coffee fruit that will be checked automatically through the image of coffee fruit for each different maturity level. The transformation of the image into numerical form, using the RGB (Red, Green, Blue) intensity approach, is then classified using k-Nearest Neighbor and principal component analysis. Two classifications were compared, namely the use of the k-Nearest Neighbor algorithm directly and the combination of k-Nearest Neighbor with principal component analysis. The results show that the classification using k-Nearest Neighbor is more accurate and the computation time is shorter than using the combination of k-Nearest Neighbor and principal component analysis.

Keywords: automatic; image classification; RGB; k-Nearest Neighbor; principal component analysis.



:)



KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa atas rahmat dan anugrah yang diberikan kepada penulis sehingga penulis dapat menyelesaikan skripsi ini. Skripsi yang berjudul “Klasifikasi Citra Buah Kopi Menggunakan Algoritma *k-Nearest Neighbor*” disusun sebagai salah satu syarat untuk menyelesaikan studi Strata-1 Program Studi Matematika, Fakultas Teknologi Informasi dan Sains (FTIS), Universitas Katolik Parahyangan, Bandung. Selama masa perkuliahan dan penulisan skripsi, banyak dukungan, ilmu, dan pengalaman baru yang didapatkan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih sebesar-besarnya kepada:

1. Orang tua dan kakak yang selalu mendukung dan memberikan semangat sehingga dapat menyelesaikan skripsi ini.
2. Bapak Agus Sukmana, M.Sc. dan Bapak Liem Chin, M.Si. selaku dosen pembimbing skripsi atas waktu dan ilmu yang diberikan serta dengan sabar membimbing dan membantu penulisan skripsi ini.
3. Bapak Dr. Julius Dharma Lesmono dan Ibu Dr. Livia Owen selaku dosen penguji skripsi atas kritik dan saran yang sangat membantu penulis dalam menyelesaikan skripsi ini.
4. Seluruh dosen FTIS yang telah memberikan ilmu selama proses perkuliahan dan staff Tata Usaha FTIS yang telah membantu dalam proses administrasi perkuliahan.
5. Teman-teman dan sahabat-sahabat angkatan 2019 dan 2020 yang tidak dapat dituliskan satu per satu yang sudah mendukung, memberikan kenangan indah, dan menemani penulis dalam keadaan suka dan duka.
6. Teman-teman dan sahabat-sahabat sekolah yang tidak dapat dituliskan satu per satu yang sudah mendukung, memberikan kenangan indah, dan menemani penulis dalam keadaan suka dan duka.
7. Semua pihak yang tidak dapat dituliskan satu per satu atas bantuan dan kebaikan yang diberikan. Semoga Tuhan Yang Maha Esa membalas segala kebaikan yang dilakukan kepada penulis.

Penulis menyadari bahwa masih terdapat banyak kekurangan di dalam skripsi ini. Oleh karena itu, penulis menerima segala kritik dan saran yang dapat membuat skripsi ini menjadi lebih baik. Akhir kata, penulis berharap skripsi ini dapat bermanfaat bagi para pembaca.

Bandung, 22 Januari 2024

Penulis

DAFTAR ISI

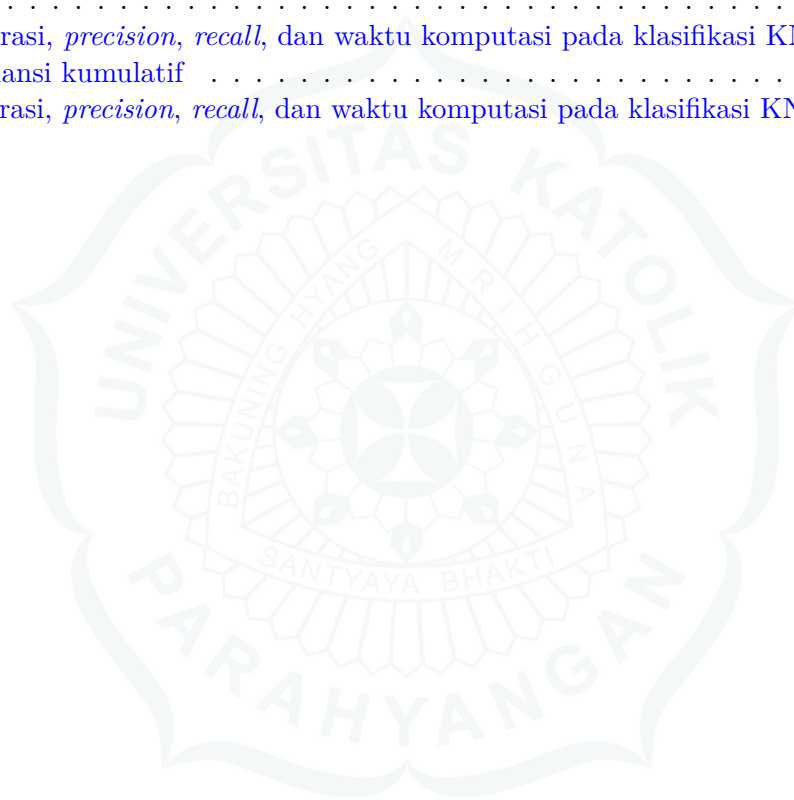
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 <i>State of the Art</i>	2
2 LANDASAN TEORI	3
2.1 Ruang Warna RGB	3
2.2 Pengolahan Data Citra	4
2.3 <i>Machine Learning</i>	6
2.4 Analisis Komponen Utama	7
2.5 <i>k-Nearest Neighbor</i>	12
2.6 Evaluasi Model	14
3 ALUR ALGORITMA KLASIFIKASI CITRA	16
3.1 Pengolahan Data Citra	17
3.2 Klasifikasi dengan Menggunakan <i>k-Nearest Neighbor</i>	18
3.3 Klasifikasi dengan Menggunakan <i>k-Nearest Neighbor</i> dan Analisis Komponen Utama	18
3.4 Evaluasi Model	19
4 ANALISIS HASIL KLASIFIKASI CITRA	20
5 KESIMPULAN DAN SARAN	26
5.1 Kesimpulan	26
5.2 Saran	26
DAFTAR REFERENSI	27

DAFTAR GAMBAR

2.1	Gambar model warna RGB	3
2.2	Contoh gambar RGB 24-bit	4
2.3	Contoh citra berdasarkan ruang warna RGB	5
2.4	Citra dengan ukuran piksel 3×3	5
2.5	Citra berwarna kuning	6
2.6	Citra berwarna merah	6
2.7	Citra berwarna hijau	6
2.8	Citra berwarna hitam	6
2.9	Plot data bivariat dengan satu data input baru	12
2.10	Plot data bivariat	13
3.1	Alur algoritma klasifikasi citra	16
3.2	Contoh citra buah kopi matang	17
3.3	Contoh citra buah kopi mentah	17
4.1	Contoh gambar dan label buah kopi sebelum direduksi	21
4.2	Perbandingan nilai k dengan nilai akurasi	21
4.3	<i>Confusion matrix</i> pada klasifikasi <i>k-Nearest Neighbor</i>	22
4.4	Variansi kumulatif pada analisis komponen utama	23
4.5	Contoh gambar dan label buah kopi setelah reduksi dimensi	24
4.6	<i>Confusion matrix</i> pada klasifikasi <i>k-Nearest Neighbor</i> dan analisis komponen utama	24

DAFTAR TABEL

2.1	<i>Confusion matrix</i>	14
3.1	<i>Contoh nilai confusion matrix</i>	19
4.1	Data latih	20
4.2	Data uji	20
4.3	Nilai akurasi, <i>precision</i> , <i>recall</i> , dan waktu komputasi pada klasifikasi KNN	22
4.4	Nilai variansi kumulatif	23
4.5	Nilai akurasi, <i>precision</i> , <i>recall</i> , dan waktu komputasi pada klasifikasi KNN dan PCA	25



BAB 1

PENDAHULUAN

1.1 Latar Belakang

Agroindustri adalah industri yang berhubungan dengan agrikultur. Salah satu contoh produk dalam agrikultur adalah kopi. Indonesia merupakan negara penghasil kopi terbesar ketiga di dunia setelah Brazil dan Vietnam.¹ Pada tahun 1920, perusahaan-perusahaan kecil dan menengah di Indonesia mulai menanam kopi sebagai komoditas utama dan perkebunan kopi. Dari perkebunan kopi tersebut akan menghasilkan biji kopi yang berasal dari dalam buah kopi. Buah kopi, seperti buah pada umumnya, tentu akan tumbuh dari awal saat masih mentah, saat mencapai titik matang, dan sampai terlalu matang seiring berjalannya waktu. Pada umumnya, proses pemeriksaan apakah suatu buah masih dalam tahap mentah, matang, atau sudah terlalu matang dilakukan oleh manusia yang ahli di bidangnya. Akan tetapi hal ini memiliki kekurangan karena keterbatasan kemampuan manusia seperti *human error* dan tidak efisien karena selain keterbatasan manusia, pemeriksaan secara manual oleh manusia juga membutuhkan waktu yang cukup lama dan biaya upah yang cukup besar sehingga sulit untuk mendapatkan keuntungan yang maksimal. Di era yang sudah modern ini, kekurangan tersebut dapat diatasi dengan bantuan teknologi yang memungkinkan untuk melakukan pemeriksaan tingkat kematangan buah kopi secara otomatis [1, hlm. 4], [2, hlm. 2]. Pemeriksaan tingkat kematangan buah kopi secara otomatis tentu lebih efisien baik secara waktu maupun biaya karena tidak menggunakan jasa manusia.

Salah satu penerapan dalam pemeriksaan tingkat kematangan buah secara otomatis yaitu dengan bantuan klasifikasi citra. Klasifikasi citra akan mengubah bentuk citra menjadi bentuk numerik agar dapat digunakan untuk proses klasifikasi selanjutnya. Klasifikasi citra ini sudah banyak dilakukan oleh para peneliti sebelumnya dalam mengetahui tingkat kematangan buah secara otomatis. Klasifikasi citra menggunakan algoritma *k-Nearest Neighbor* menghasilkan performa yang sangat tinggi mencapai lebih dari 96% berdasarkan [3, hlm. 748]. Kemudian penelitian yang dilakukan oleh Bouteldja dan Kourgli pada tahun 2020 menunjukkan bahwa algoritma *k-Nearest Neighbor* memiliki tingkat akurasi yang lebih baik jika dibandingkan dengan algoritma *decision tree* serta lebih sederhana daripada algoritma *SVM (support vector machine)* [4, hlm. 6]. Algoritma *k-Nearest Neighbor* cenderung mudah dipelajari namun tetap memiliki tingkat akurasi yang baik [5, hlm. 1]. Oleh karena itu, pada penelitian ini akan digunakan algoritma *k-Nearest Neighbor* dan algoritma kombinasi antara *k-Nearest Neighbor* dan analisis komponen utama untuk mengklasifikasi citra buah kopi. Citra buah kopi yang dimaksud berasal dari hasil foto buah kopi.

Selain algoritma *k-Nearest Neighbor*, akan dipakai bantuan algoritma yang mereduksi dimensi

¹<https://indonesiabaik.id/infografis/negara-penghasil-kopi-terbesar>, diakses pada 12 Oktober 2023.

data pada citra buah kopi. Algoritma analisis komponen utama akan digunakan pada penelitian ini karena analisis komponen utama dapat mengurangi jumlah dimensi data yang signifikan namun tetap dapat mempertahankan informasi penting sebanyak mungkin. Selain itu, analisis komponen utama cenderung lebih unggul jika dibandingkan algoritma reduksi dimensi lainnya seperti *linear discriminant analysis* [6, hlm. 40], [7, hlm. 233].

Pada penelitian ini akan digunakan perbedaan warna RGB pada citra untuk setiap tingkat kematangan buah kopi dengan menggunakan algoritma *k-Nearest Neighbor* dan kombinasi antara algoritma analisis komponen utama dan algoritma *k-Nearest Neighbor*. Algoritma analisis komponen utama dan algoritma *k-Nearest Neighbor* merupakan algoritma yang sering digunakan pada bidang *machine learning*. *Machine learning* menggunakan data-data yang tersedia untuk menciptakan algoritma prediksi yang efisien dan akurat. Perbandingan antara klasifikasi citra kematangan buah kopi yang hanya memakai algoritma *k-Nearest Neighbor* dengan dan tanpa analisis komponen utama berdasarkan lama waktu komputasi dan tingkat akurasi pada klasifikasi.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, berikut rumusan masalah-masalah yang akan diselesaikan:

1. Bagaimana langkah-langkah penerapan algoritma *k-Nearest Neighbor* dan kombinasinya dengan analisis komponen utama untuk mengklasifikasi citra kematangan buah kopi?
2. Bagaimana hasil evaluasi dari klasifikasi citra kematangan buah kopi menggunakan algoritma *k-Nearest Neighbor* dengan dan tanpa analisis komponen utama?

1.3 Tujuan

Tujuan pada penelitian adalah:

1. Menerapkan algoritma *k-Nearest Neighbor* dan kombinasinya dengan analisis komponen utama untuk mengklasifikasi citra kematangan buah kopi.
2. Mengetahui hasil evaluasi dari klasifikasi citra kematangan buah kopi menggunakan algoritma *k-Nearest Neighbor* dengan dan tanpa analisis komponen utama.

1.4 *State of the Art*

Penelitian yang dilakukan oleh Bharate dan Shirdhonkar menggunakan algoritma *support vector machine* untuk mengklasifikasi citra [3]. Pada penelitian ini, diterapkan klasifikasi citra kematangan buah kopi menggunakan algoritma yang berbeda, yaitu *k-Nearest Neighbor* dengan dan tanpa analisis komponen utama. Penggunaan analisis komponen utama membantu mereduksi dimensi data yang memiliki skala besar. Penelitian yang dilakukan oleh Andrian, dkk. menggunakan ruang warna *grayscale* untuk proses klasifikasi citra [8]. Pada penelitian ini, proses klasifikasi citra dilakukan berdasarkan ruang warna RGB.