

Classification of Avocado Ripeness Levels using Naïve Bayes Method

Ira Nuryani¹, Aldi Muhammad Nur Fadli², Nadila Dwi Saputri³, Alfira Nisa Fadilah⁴, Muhammad Resa Arif Yudianto⁵, Maimunah⁶

^{1,2,3,4,5,6} Department of Informatic Engineering, Universitas Muhammadiyah Magelang Indonesia

 ¹ ira.nuryanii1@gmail.com

 ² almunufadi87@gmail.com

 ³ nadiladwis.12@gmail.com

 ⁴ alfiranisa120617@gmail.com

 ⁵ resamuhammad96@unimma.ac.id

 ⁶ maimunah@unimma.ac.id

Abstract

During this time most people in determining the ripeness of avocados for personal consumption is not difficult because they can distinguish themselves but another case if used for production, which requires a lot of labor to group ripe and raw avocados. One of the innovations in information and communication technology in agriculture and plantations is the use of classification methods with naïve bayes algorithms. The formula of the problem in this study is how to do the classification on the ripeness of avocados and see the accuracy rate of the data. The purpose of this study is to classify avocado ripeness and to acquire intelligent systems, so that it becomes the first step towards the implementation stage. Based on the results and analysis that has been done, it can be concluded that the Naive Bayes method is considered capable in classifying avocado ripeness by using RGB color features. The accuracy in testing using Naïve Bayes method reached 83.34%. The performance obtained from this intelligent system is also effective and efficient so that the classification of avocado ripeness can be implemented.

Keywords : Classification, Fruit Ripeness, Naïve Bayes

Kalsifikasi Kematangan Buah Alpukat Menggunakan Metode Naïve Bayes

Abstrak

Selama ini kebanyakan orang dalam menentukan kematangan buah alpukat untuk konsumsi pribadi tidaklah sulit karena mereka bisa membedakannya sendiri namun lain halnya jika digunakan untuk produksi, yang membutuhkan banyak tenaga kerja untuk mengelompokkan buah alpukat yang matang dan mentah. Salah satu inovasi teknologi informasi dan komunikasi dalam bidang pertanian dan perkebunan adalah penggunaan metode klasifikasi dengan algoritma naïve bayes. Rumusan masalah dalam penelitian ini adalah bagaimana melakukan klasifikasi pada kematangan buah alpukat dan melihat tingkat akurasi data. Tujuan dari penelitian ini adalah untuk melakukan pengklasifikasian pada kematangan buah alpukat dan untuk memperoleh sistem cerdas, sehingga menjadi langkah awal untuk menuju tahap pengimplementasian. Berdasarkan hasil dan analisis yang telah dilakukan, dapat disimpulkan bahwa metode Naive Bayes dinilai mampu dalam melakukan pengklasifikasian kematangan buah alpukat dengan menggunakan fitur warna RGB. Hasil akurasi dalam pengujian menggunakan metode Naïve Bayes mencapai 83,34%. Kinerja yang diperoleh dari sistem cerdas ini juga efektif dan efisien sehingga klasifikasi kematangan buah alpukat dapat di implementasikan.

Kata kunci : Klasifikasi, Kematangan Buah, Naïve Bayes.

1. Pendahuluan

Alpukat berasal dari Amerika Tengah, yaitu Mexico, Peru dan Venezuela, dan telah menyebar luas ke berbagai negara sampai ke Asia Tenggara, termasuk Indonesia[1]. Alpukat adalah salah satu komoditas buah dengan tingkat produksi cukup tinggi di Indonesia. Tercatat sekitar kurang lebih 609 ribu ton buah alpukat yang diproduksi pada tahun 2020 di Indonesia. Selama ini kebanyakan orang dalam menentukan kematangan buah alpukat untuk konsumsi pribadi tidaklah sulit karena mereka bisa membedakannya sendiri namun lain halnya jika digunakan untuk produksi, yang membutuhkan banyak tenaga kerja untuk mengelompokkan buah alpukat yang matang dan mentah. Kemampuan akan membedakan tingkat kematangan buah alpukat tentu saja dibutuhkan pengamatan yang akurat[2]. Salah satu inovasi teknologi informasi dan komunikasi dalam bidang pertanian dan perkebunan adalah penggunaan metode klasifikasi dengan algoritma Naïve Bayes. Pengolahan citra terkait kematangan buah telah dilakukan oleh beberapa peneliti sebelumnya. Diantaranya adalah peneliti atas nama Rahmi Hidayati dengan judul penelitian Klasifikasi Kematangan Pepaya Menggunakan Ruang Warna Hsv Dan Metode Naive Bayes Classifier. Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, maka diperoleh pengujian berupa sebanyak 20 pepaya terklasifikasi dengan tepat, dan tidak satupun hasil klasifikasi yang salah[3]. Dan Peneliti atas nama Kartika Chandra Pelangi dengan judul Sistem Cerdas Dalam Klasifikasi Kematangan Buah Jeruk Berdasarkan Fitur Ekstraksi Gcm Dengan Metode Naïve Bayes. Dimana tingkat kematangan akan dibagi menjadi dua jenis, yaitu jeruk matang dan jeruk belum matang[4]. Kemudian pada penelitian Klasifikasi Tingkat Kematangan Buah Apel Manalagi Dengan Algoritma Naïve bayes Dan Ekstraksi Citra Digital. Dijelaskan definisi dari Naïve Bayes sendiri dimana dapat menangani kuantitatif dan data diskrit, tidak memerlukan data latih dalam jumlah banyak untuk mengestimasi parameter-parameter yang dibutuhkan untuk klasifikasi, dan cepat dalam melakukan kalkulasi[5]

Penggunaan teknologi pengolahan citra buah alpukat yang terintegrasi dengan komputer dapat membuat sistem cerdas. Dengan menggunakan metode pengolahan citra digital dengan buah alpukat berdasarkan nilai RGB dan mengklasifikasikan buah alpukat ke dalam dua kelas, yaitu matang dan mentah menggunakan algoritma Naïve Bayes. Pengklasifikasian berupa fitur RGB seperti proses pemisahan warna *red* (R), *green* (G), dan *blue* (B), digunakan untuk mengetahui unsur warna yang lebih dominan untuk masing - masing kelas Alpukat. Data yang digunakan untuk mengolah buah alpukat berupa *image* alpukat yang didapatkan dari hasil pemotretan buah alpukat sebanyak 60 buah alpukat.

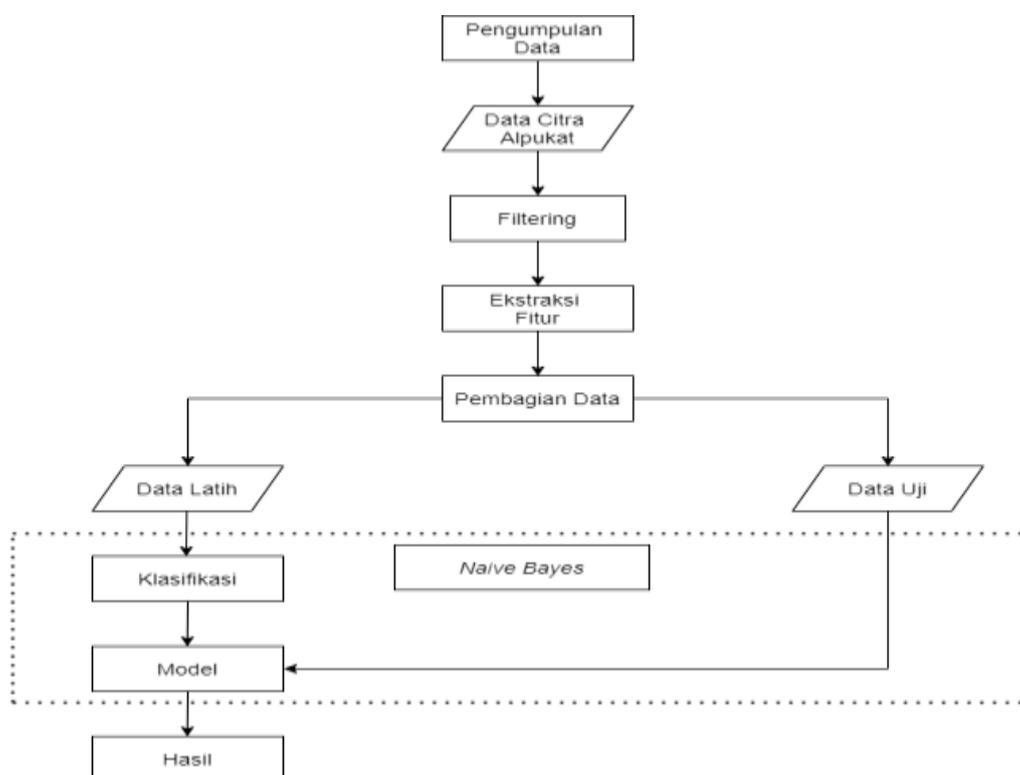
Membedakan buah alpukat yang mentah dan matang terbilang sedikit mudah bagi kebanyakan orang yang ingin mengonsumsi untuk pribadi. Lain halnya jika mengklasifikasikan tingkat kematangan buah alpukat yang digunakan untuk produksi, tentu buah alpukat yang digunakan tidak hanya seratus atau dua ratus, bisa saja sampai ribuan. Sedangkan tenaga kerja manusia tidak memungkinkan untuk melakukan pekerjaan mengklasifikasi kematangan buah alpukat dengan beribu-ribu buah. Perlu tenaga mesin yang menggunakan sistem kecerdasan buatan yang dalam hal ini adalah sistem klasifikasi kematangan buah alpukat dengan mengambil citra masing-masing buah alpukat untuk dibedakan tingkat kematangannya. Untuk itulah dirumuskan masalah

mengenai, bagaimana cara mengklasifikasi tingkat kematangan buah alpukat berdasar warna menggunakan metode Naive Bayes?

Tujuan dibuatnya penelitian ini adalah untuk mempermudah mengklasifikasi tingkat kematangan buah alpukat yang digunakan untuk konsumsi pribadi atau untuk kepentingan produksi.

2. Metode

Metode penelitian yang dilakukan meliputi pengumpulan data citra alpukat, *preprocessing*, ekstraksi ciri fitur warna RGB, pembagian data latih dan data uji, model klasifikasi dengan metode Naive Bayes kemudian membaca hasilnya. Tahapan penelitian tersebut dapat dilihat dalam [Gambar 1](#).



Gambar 1. *Flowchart Penelitian buah Alpukat*

2.1. Pengambilan Dataset

Data yang digunakan adalah citra buah alpukat jenis miki berjumlah 30 buah. Pengambilan citra buah alpukat dilakukan dengan membuat studio mini. Buah alpukat diletakkan dalam studio mini, dilakukan proses pengambilan citra. Posisi objek dengan citra berjarak 20 cm dari kamera. Proses pengambilan citra dilakukan dengan mengambil 1 buah data objek secara bergantian, diletakkan dalam studio mini dan dilakukan proses pengambilan citra. Total citra yang diambil ada 60 citra, yang terdiri dari 30 citra buah alpukat mentah dan 30 citra buah alpukat matang dengan resolusi 3120 piksel atau ukuran 1:1.

2.2. Pengolahan Citra

Pengolahan citra adalah suatu teknik yang digunakan untuk memproses citra atau gambar dengan komputer menjadi citra yang kualitasnya lebih baik. Caranya yaitu dengan memanipulasinya dengan data gambar yang diinginkan untuk mendapatkan informasi

tertentu dari gambar yang diamati. Dengan kata lain pengolahan citra adalah kegiatan memperbaiki kualitas citra agar mudah diinterpretasi oleh manusia atau mesin.

2.3. Ekstraksi Ciri

Ekstraksi ciri merupakan tahapan yang mengekstrak informasi yang tergantung dalam suatu objek dalam citra digital. Informasi tersebut digunakan untuk membedakan antara objek yang satu dengan objek lainnya pada tahapan pengenalan atau identifikasi citra.

Untuk melakukan pencarian menggunakan gambar maka perlu dilakukan pengolahan citra digital untuk memperoleh ciri atau fitur dari gambar tersebut. Fitur yang digunakan adalah *red, green, dan blue* (RGB) *color channel* dan *simple morphological shape descriptors*.

2.4. Naïve Bayes

Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes [3] yang memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal dengan Teorema Naive Bayes. Klasifikasi Naive Bayes ini diasumsikan bahwa kelas satu tidak terpengaruh dengan kondisi kelas lainnya, yang artinya bersifat independen. Langkah-langkah untuk pelatihan data:

- Hitung probabilitas prior tiap kelas yang ada
- Lalu hitung rata-rata mean tiap fitur dan tiap kelas
- Hitung nilai standar deviasi tiap fitur dan tiap kelas
- Hitung densitas probabilitas
- Hitung probabilitas masing-masing kelas

Rumus Teorema Naive Bayes dilihat pada persamaan(1).

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$

Keterangan:

- X : Data dengan kelas yang belum diketahui
 H : hipotesis data X merupakan suatu kelas spesifik
 P(H | X) : Probabilitas hipotesis H berdasar kondisi X
 P(H) : Probabilitas hipotesis H
 P(X | H) : Probabilitas X berdasarkan kondisi pada hipotesis H
 P(X) : Probabilitas X.

3. Hasil dan Pembahasan

Pengujian sistem klasifikasi Naive Bayes ini bertujuan untuk mengetahui keakuratan sistem dalam mengklasifikasi tingkat kematangan buah alpukat, apakah termasuk matang atau mentah. Uji coba dilakukan dengan menggunakan 30 gambar buah alpukat matang dan 30 buah alpukat mentah dipaparkan dalam [Gambar 2](#). Alpukat Mentah dan Matang.



Gambar 2. Alpukat Mentah dan Matang

3.1. Hasil Ekstraksi Fitur

Preprocessing dilakukan terlebih dahulu untuk memisahkan objek citra alpukat dengan latar belakangnya. Kemudian dilakukan pengolahan citra seperti *filtering* untuk memperbaiki kualitas citra. Setelahnya melakukan deteksi kontur agar dapat menentukan nilai piksel pada objek yang akan diekstraksi ciri. Pada tahap ekstraksi fitur dilakukan ekstraksi ciri atau fitur warna berupa R, G, B. Untuk satu buah alpukat, nilai ciri tersebut diperoleh dengan meratakan semua piksel yang ada, data tersebutlah yang kemudian dilakukan pengenalan. Berikut contoh nilai hasil ekstraksi fitur yang dapat dilihat pada [Tabel 1](#);

Tabel 1. Nilai Ekstraksi Ciri Warna RGB dan Jenis Kematangan Buah Alpukat

Jenis	Nilai R	Nilai G	Nilai B
Mentah	60	77	37
Mentah	68	83	36
Mentah	54	70	35
Mentah	59	74	35
Mentah	59	72	36
Matang	36	29	26
Matang	46	31	28
Matang	35	31	29
Matang	42	36	34
Matang	48	35	31

3.2. Hasil Klasifikasi

Citra buah alpukat dibagi menjadi dua data, yaitu data uji dan data training. Pembagian data training sebesar 80%, dan data uji sebesar 20%. Setelah mendapat pemodelan data dari data training, dilakukan pengujian dengan data uji yang kemudian diperoleh hasil akurasi sebesar 83,34%. Hal ini dapat dilihat pada nilai prediksi dan akurasi data yang diperoleh pada [Gambar 3](#);

```
[12] #Nilai akurasi
from sklearn.metrics import accuracy_score
accuracy= accuracy_score(y_test, y_pred)
accuracy

0.8333333333333334

[13] y_pred=modelnb.predict(X_test)
print("The predicted Data is :")
print(y_pred)
print("The actual data is:")
print(np.array(y_test))
print(f"The model is {accuracy_score(y_pred,y_test)*100}% accurate")

The predicted Data is :
[33 35 37 35 31 34 32 35 38 35 31 35]
The actual data is:
[33 35 37 35 28 34 32 35 40 35 31 35]
The model is 83.33333333333334% accurate
```

Gambar 3. Hasil Tingkat Akurasi Data

Akurasi adalah perbandingan jumlah prediksi yang benar [6]. Untuk menghitung akurasi menggunakan formula sebagaimana pada persamaan(2) [7].

$$Accuracy = \frac{TP + TN}{TP + TN + TN + FP + FN} \quad (2)$$

Berdasarkan Gambar 3., diketahui metode Naive Bayes untuk mengklasifikasikan tingkat kematangan buah alpukat telah berhasil memperoleh nilai akurasi 83,3% . Hal ini dikarenakan fitur warna buah alpukat belum dapat sepenuhnya membedakan secara baik buah alpukat matang dan mentah, karena warna buah alpukat mentah dan matang ada yang hampir mirip. Untuk itu, perlu penambahan ekstraksi ciri lain seperti tekstur yang dapat membantu memperoleh ciri yang lebih khusus.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa ekstraksi fitur warna RGB dapat digunakan dalam mengekstraksi ciri pada citra buah alpukat. Hasil ekstraksi ciri tersebut digunakan sebagai masukan dalam metode pengklasifikasian Naive Bayes yang dilakukan untuk mengenali ciri citra dan mengklasifikasikan tingkat kematangan buah alpukat. Setelah dilakukan proses klasifikasi, hasil klasifikasi menunjukkan tingkat akurasi sebesar 83,34%. Untuk pengembangan lebih lanjut, klasifikasi kematangan buah alpukat dapat dilakukan dengan menambahkan fitur lain seperti tekstur dengan menggunakan algoritma lain seperti decision tree, neural network atau K-Nearest Neighbor.

Referensi

- [1] L. Sadwiyanti, S. Djoko, and T. Budiyan, *Budiaya Alpukat*. 2009.
- [2] M. H. Hanafi, N. Fadillah, and A. Insan, "Optimasi Algoritma K-Nearest Neighbor untuk Klasifikasi Tingkat Kematangan Buah Alpukat Berdasarkan Warna," *It J. Res. Dev.*, vol. 4, no. 1, pp. 10–18, 2019, doi: 10.25299/itjrd.2019.vol4(1).2477.
- [3] R. H. Ellif, Sampe Hotlan Sitorus, "Klasifikasi Kematangan Pepaya Menggunakan Ruang Warna HSV dan Metode Naive Bayes Classifier," *J. Komput. dan Apl.*, vol. 09, no. 01, pp. 66–75, 2021.
- [4] A. Rahmat, K. Haba, and K. C. Pelangi, "Ruk Berdasarkan Fitur Ekstraksiglem Dengan Metode," *J. Teknol. Dan Manaj. Inform.*, vol. 6, no. 1, 2020.
- [5] A. Ciputra, D. R. I. M. Setiadi, E. H. Rachmawanto, and A. Susanto, "Klasifikasi Tingkat Kematangan Buah Apel Manalagi Dengan Algoritma Naive Bayes Dan Ekstraksi Fitur Citra Digital," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 9, no. 1, pp. 465–472, 2018, doi: 10.24176/simet.v9i1.2000.
- [6] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognit.*, vol. 91, pp. 216–231, 2019, doi: 10.1016/j.patcog.2019.02.023.
- [7] A. P. Taftazani Ghazi Pratama, Achmad Ridwan, "Penerapan Algoritma C4.5 untuk Klasifikasi Kanker Serviks Tingkat Awal," vol. 1, no. 1, pp. 1–6, 2021.



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)