

NASKAH PUBLIKASI

**SISTEM DETEKSI JENIS CACAT BIJI KOPI DENGAN
ALGORITMA K-NEAREST NEIGHBOR**

Program Studi Teknik Informatika

Disusun oleh:

ARISKA RESTU GINANJAR

5150411169

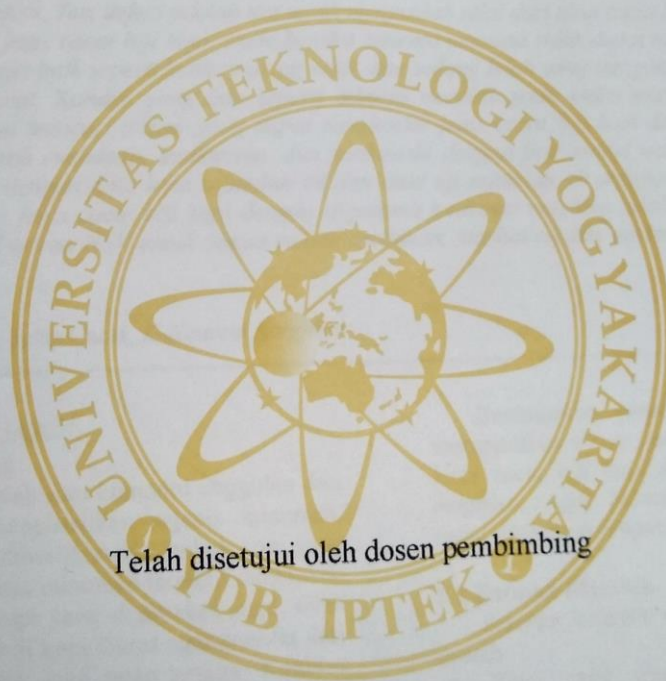
**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN ELEKTRO
UNIVERSITAS TEKNOLOGI YOGYAKARTA**

2019



NASKAH PUBLIKASI

**SISTEM DETEKSI JENIS CACAT BIJI KOPI DENGAN
ALGORITMA K-NEAREST NEIGHBOR**

Disusun oleh:
ARISKA RESTU GINANJAR
5150411169



Telah disetujui oleh dosen pembimbing

Pembimbing



Dr. Enny Itje Sela, S.Si., M.Kom. Tanggal: 12 September 2019

Sistem Deteksi Jenis Cacat Biji Kopi dengan Algoritma K-Nearest Neighbor

Ariska Restu Ginanjar¹, Enny Itje Sela²

¹Program Studi Teknik Informatika, Fakultas Teknologi Informasi dan Elektro

²Program Studi Teknik Informatika, Fakultas Teknologi Informasi dan Elektro
Universitas Teknologi Yogyakarta

Jl. Ringroad Utara Jombor Sleman Yogyakarta

E-mail : ariskarestuginanjar@gmail.com

ABSTRAK

Kopi adalah salah satu komoditi unggulan dari Indonesia yang menghasilkan sejenis minuman. Kualitas minuman tersebut ditentukan dari kualitas biji kopi yang digunakan juga. Kualitas biji kopi salah satunya berdasarkan test defect. Test defect adalah test untuk menjumlah nilai dari nilai cacat biji kopi dan besaran nilai cacat berdasarkan jenis cacat biji kopi. Pada kondisi tertentu manusia tidak dapat melakukan penentuan jenis cacat biji kopi dengan baik seperti ketika sedang sakit atau sedang lelah yang mengakibatkan ketidaktepatan dalam penentuan jenis cacat. Kondisi yang lain seperti adanya masalah pada indra manusia seperti buta warna. Penelitian bertujuan mencari sistem yang dapat mendeteksi jenis cacat biji kopi dengan algoritma k-nearest neighbor rumus jarak euclidean, manhattan, dan minkowski dengan fitur model warna RGB dan HSV. Hasil dari penelitian ini dengan data latih sejumlah 60 dan data uji sejumlah 40 didapatkan bahwa sistem terbaik pada sistem deteksi jenis cacat biji kopi dengan algoritma k-nearest neighbor adalah nilai k sebesar 3 serta dengan fitur model warna HSV untuk semua rumus euclidean, manhattan dan minkowski mendapatkan akurasi sebesar 97%.

Kata kunci : Kopi, Jenis Cacat, K-Nearest Neighbor

1. PENDAHULUAN

1.1 Latar Belakang

Kopi adalah salah satu komoditi unggulan dari Indonesia yang menghasilkan sejenis minuman. Minuman tersebut dihasilkan dari bubuk kopi yang sudah diolah. Kualitas minuman tersebut ditentukan dari kualitas biji kopi yang digunakan juga. Oleh karena itu kualitas biji kopi dijaga agar kualitas dari minuman dari bubuk kopi tetap terjaga. Kualitas biji kopi diukur dengan 2 faktor yaitu mutu fisik dan mutu citarasa. Mutu fisik adalah penentuan kualitas kopi berdasarkan jumlah nilai cacat yang terdapat dalam biji kopi. Sedangkan mutu citarasa adalah penentuan berdasarkan uji organoleptik. Pada mutu fisik terdapat test defect. Test defect adalah test untuk menjumlah nilai dari nilai cacat biji kopi dan besaran nilai cacat berdasarkan jenis cacat biji kopi.

Penentuan jenis cacat biji kopi dilakukan dengan menilai dari ciri – ciri fisik biji kopi dengan menggunakan indra manusia dengan standar yang berlaku. Akan tetapi pada kondisi tertentu manusia tidak dapat melakukan penentuan jenis cacat biji kopi dengan baik seperti ketika sedang sakit atau sedang lelah yang mengakibatkan ketidaktepatan dalam penentuan jenis cacat. Kondisi yang lain seperti adanya masalah pada indra manusia seperti buta warna.

Berdasarkan permasalahan tersebut penulis mengusulkan sistem yang dapat melakukan deteksi jenis cacat biji kopi dengan algoritma k-nearest neighbor. Agar penentuan jenis cacat biji kopi dapat dilakukan dengan baik.

1.2 Batasan Masalah

Adapun batasan masalah dari penelitian ini adalah :

- Kopi yang digunakan adalah kopi jenis robusta.
- Kopi yang diolah yaitu kopi yang sudah siap untuk menentukan mutu kopi.
- Hanya mendeteksi jenis cacat biji coklat penuh, biji hitam penuh, biji muda, dan biji normal.

1.3 Tujuan Penelitian

Adapun tujuan penelitian dari penelitian ini adalah :

- Untuk memanfaatkan algoritma k-nearest neighbor pada penentuan jenis cacat biji kopi.
- Untuk mengetahui berapa nilai akurasi yang dihasilkan oleh sistem.

2. LANDASAN TEORI

2.1 Kopi

Kopi adalah salah satu minuman yang berasal dari proses pengolahan dan ekstraksi biji tanaman

kopi. Kopi merupakan minuman psikostimulant yang akan menyebabkan orang tetap terjaga, sehingga kopi menjadi minuman favorit terutama bagi kaum pria [1].

Kopi terdiri dari dua jenis spesies, yaitu kopi arabika dan kopi robusta. Kopi arabika tumbuh pada ketinggian di atas 1.000 meter di atas permukaan laut, sedangkan kopi robusta tumbuh di bawah ketinggian 1.000 meter di atas permukaan laut. Rasa kopi robusta lebih netral, aroma kopi lebih kuat, dan kadar kafein lebih tinggi daripada kopi arabika [2].

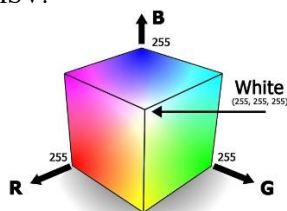
2.2 Model Warna

Model warna atau ruang warna merupakan suatu spesifikasi sistem koordinat dan suatu sub ruang dalam sistem tersebut dengan setiap warna dinyatakan dengan satu titik didalamnya. Tujuan dibentuknya ruang warna adalah untuk memfasilitasi spesifikasi warna dalam bentuk suatu standar. Jenis-jenis ruang warna diantaranya [3].

a. RGB

Ruang warna ini menggunakan tiga komponen dasar yaitu merah (R), hijau (G), dan biru (B). Setiap piksel dibentuk oleh ketiga komponen tersebut. RGB biasa digunakan karena kemudahan dalam perancangan *hardware*, tetapi sebenarnya tidak ideal untuk beberapa aplikasi.

Warna merah, hijau, dan biru sesungguhnya berkorelasi erat, sangat sulit untuk beberapa algoritma pemrosesan citra. Sebagai contoh, kebutuhan untuk memperoleh warna alamiah seperti merah dengan menggunakan RGB menjadi sangat kompleks mengingat komponen R dapat berpasangan dengan G dan B, dengan nilai berapa saja. Hal ini menjadi mudah jika menggunakan ruang warna HSV.



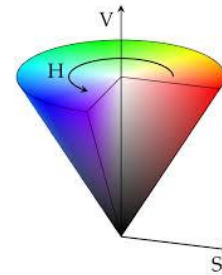
Gambar 1 Model Warna RGB

b. HSV

HSV merupakan contoh ruang warna yang merpresentasikan warna seperti yang dilihat oleh manusia. H berasal dari kata *hue*, S berasal dari *saturation* dan V berasal dari *value*.

HSV (*Hue, Saturation, Value*) merupakan model warna yang diturunkan dari RGB [4]. Model warna HSV mendefinisikan warna dalam terminologi *Hue*, *Saturation*, dan *Value*. *Hue* menyatakan warna sebenarnya, seperti merah, violet, dan kuning. *Hue* berasosiasi dengan panjang gelombang cahaya. *Saturation* menyatakan tingkat kemurnian suatu warna, yaitu mengindikasikan seberapa banyak

warna putih diberikan pada warna. *Value* adalah atribut yang menyatakan banyaknya cahaya yang diterima oleh mata tanpa memperdulikan warna.



Gambar 2 Model Warna HSV

2.3 K-Nearest Neighbor

K-nearest neighbor (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut [5]. *K-nearest neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama dengan berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada yang memiliki kesamaan (*similarity*) [6]. Tujuan dari algoritma ini untuk mengklasifikasikan objek baru berdasarkan atribut dan *training sample*. *Classifier* tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori.

Adapun beberapa rumus jarak yang akan digunakan antara lain:

a. Euclidean Distance

Euclidean distance digunakan untuk menghitung jarak 2 titik. Rumus *euclidean distance* dapat dilihat pada persamaan 1.

$$distance(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

b. Manhattan Distance

Manhattan distance atau *city block* digunakan untuk menghitung jarak berdasarkan penjumlahan jarak selisih antara dua buah obyek. Hasil yang didapatkan dari *manhattan distance* bernilai mutlak. Rumus *manhattan distance* dapat dilihat pada persamaan 2.

$$distance(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (2)$$

c. Minkowski Distance

Minkowski distance merupakan bentuk umum dari *euclidean distance* dan *manhattan distance*. Rumus *minkowski distance* dapat dilihat pada persamaan 3.

$$distance(x, y) = (\sum_{i=1}^n |x_i - y_i|^p)^{1/p} \quad (3)$$

2.4 Metode Pengujian

Confusion matrix merepresentasikan prediksi dan kondisi sebenarnya (aktual) dari data yang dihasilkan oleh algoritma *machine learning*. Berdasarkan *confusion matrix* bisa menentukan *accuracy*, *precision*, *recall* dan *f1-score* [7].

		Actual Values	
		Positive	Negative
Predicted Values	Positive	TP	FP
	Negative	FN	TN

Gambar 3 Confusion Matrix

a. Accuracy

Accuracy merupakan rasio prediksi benar (positif dan negatif) dengan keseluruhan data. Nilai *accuracy* didapatkan dengan persamaan 4.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

b. Precision

Precision merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Nilai *precision* didapatkan dengan persamaan 5.

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

c. Recall

Recall merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Nilai *recall* didapatkan dengan persamaan 6.

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

d. F1-Score

F1-score merupakan perbandingan rata-rata *precision* dan *recall* yang dibobotkan. Nilai *f1-score* didapatkan dengan persamaan 7.

$$F1 - Score = 2x \frac{Recall*Precision}{Recall+Precision} \quad (7)$$

3. METODOLOGI PENELITIAN

3.1 Data Penelitian

Objek penelitian yang berupa biji kopi yang siap untuk menentukan mutu kopi. Adapun beberapa jenis cacat yang akan dilakukan deteksi yaitu coklat penuh, hitam penuh, muda, dan normal. Tiap jenis cacat terdapat 15 gambar data latih dan 10 gambar data uji.

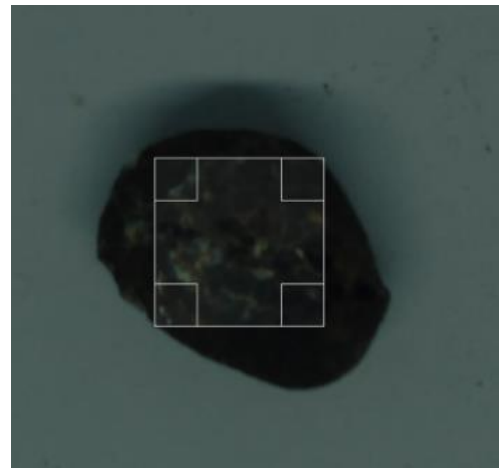
3.2 Metode Penelitian

a. Alur Pembuatan Data Latih dan Data Uji

Pengambilan gambar data latih dan data uji menggunakan kamera *Canon EOS 1600D* dengan ISO 125. Pengambilan gambar dilakukan dengan ketinggian kurang lebih 18 cm.

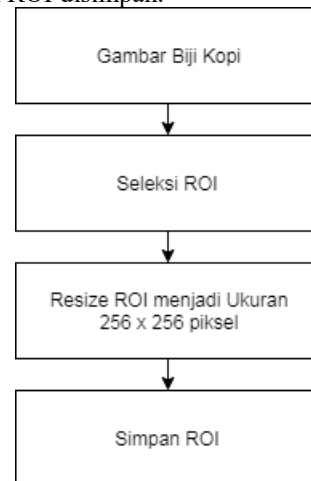


Gambar 4 Pengambilan Gambar Biji Kopi



Gambar 5 Pengambilan ROI

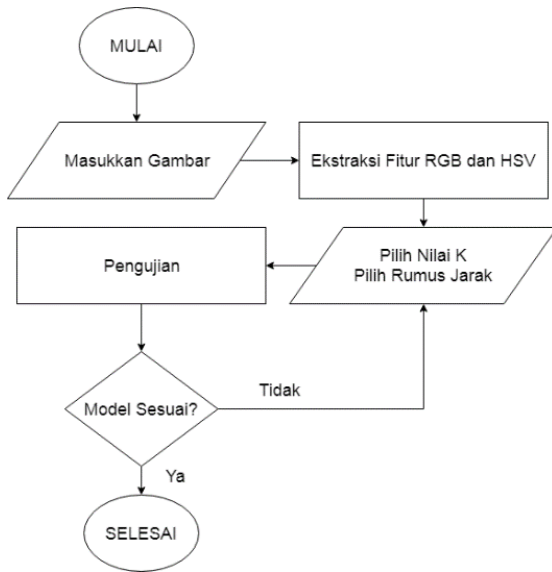
Gambar kemudian dilakukan proses pengambilan ROI (*Region of Interest*). Selanjutnya ROI di *resize* menjadi ukuran 256x256 piksel. Kemudian ROI disimpan.



Gambar 6 Alur Pembuatan Data Latih dan Data Uji

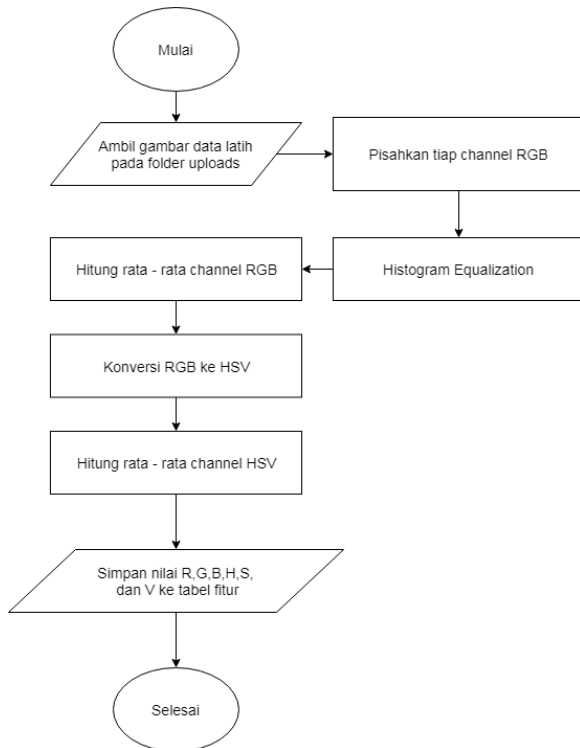
b. Perancangan Sistem

Perancangan sistem merupakan proses membuat gambaran sistem yang akan dibangun. Gambaran sistem ditampilkan dengan *flowchart*

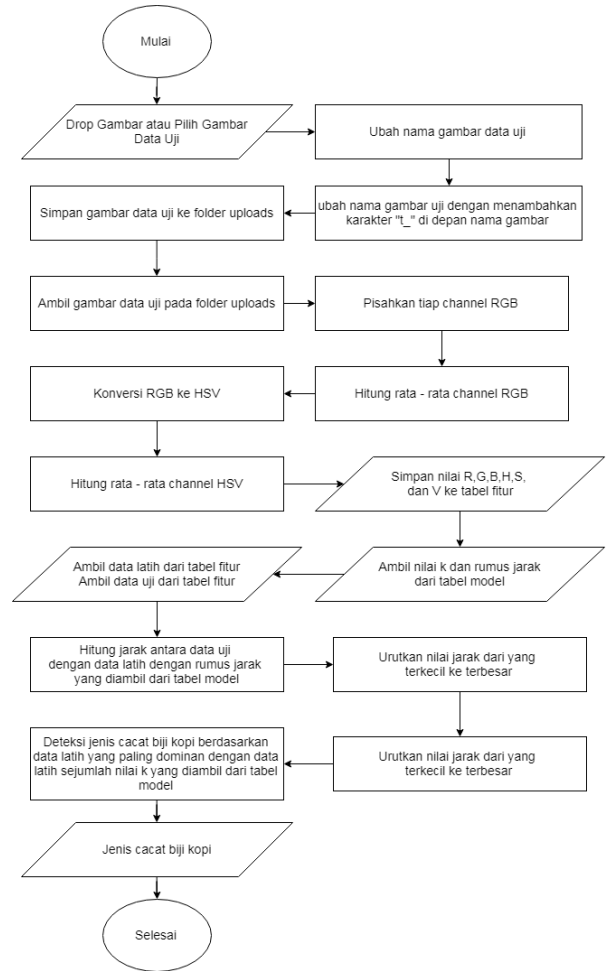


Gambar 7 Perancangan Sistem

Berikut adalah alur dari proses ekstraksi fitur.



Gambar 8 Ekstraksi Fitur



Gambar 9 Proses Pengujian


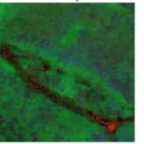
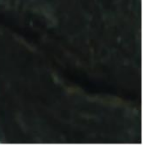
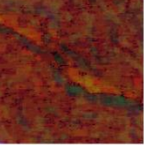

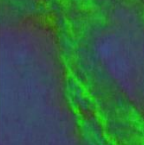
4. HASIL DAN PEMBAHASAN

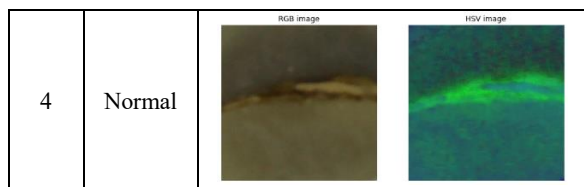
4.1 Hasil

Berikut adalah hasil dari penelitian ini antara lain

a. Konversi RGB ke HSV

Tabel 1 Konversi RGB ke HSV

No	Label	Gambar	
1	Coklat Penuh		
2	Hitam Penuh		
3	Muda		



b. Confusion Matrix

Tabel 2 Confusion Matrix

Aktual \ Prediksi	Prediksi			
	Coklat Penuh	Hitam Penuh	Muda	Normal
Coklat Penuh	10	0	0	0
Hitam Penuh	1	9	0	0
Muda	0	0	10	0
Normal	0	0	0	10

c. Precision, Recall dan F1-Score

Tabel 3 Precision, Recall dan F1-Score

Label	Precision	Recall	F1-Score
Coklat Penuh	0.91	1.0	0.95
Hitam Penuh	1.0	0.9	0.95
Muda	1.0	1.0	1.0
Normal	1.0	1.0	1.0

d. Accuracy

Tabel 4 Akurasi yang Dihasilkan

Jarak	K	Akurasi (%)		
		RGB	HSV	RGB dan HSV
Euclidean Distance	3	95	97	95
	5	90	95	90
	7	85	95	85
Manhattan Distance	3	95	97	95
	5	82	95	82
	7	78	95	78
Minkowski Distance	3	95	97	95
	5	90	95	90
	7	85	95	85

4.2 Pembahasan

Model sistem dihasilkan pada penelitian ini berjumlah 27 model. Model tersebut dikelompokkan berdasarkan rumus jarak yang digunakan pada algoritma *k-nearest neighbor* untuk memudahkan dalam mendeskripsikan model yang telah dibuat. Berdasarkan informasi yang terdapat pada tabel 5 didapatkan bahwa dari semua rumus jarak yang digunakan baik *euclidean distance*, *manhattan distance*, dan *minkowski distance* dan nilai k sebesar 3 dengan fitur HSV mendapatkan akurasi tertinggi yaitu 97%.

- [8] T. Y. Prahudaya dan A. Harjoko, "Metode Klasifikasi Mutu Jambu Biji Menggunakan KNN Berdasarkan Fitur Warna Dan Tekstur," *J. Teknosains*, vol. 6, no. 2, hal. 113–123, 2017.
- [9] M. Rizal, "Klasifikasi Mutu Biji Kopi Menggunakan Metode K-Nearest Neighbor Berdasarkan Warna dan Tekstur," Universitas Teknologi Yogyakarta, 2019.
- [10] E. I. Sela dan M. Ihsan, "Deteksi Kualitas Telur Menggunakan Analisis Tekstur," *IJCCS (Indonesian J.*

5. PENUTUP

5.1. Kesimpulan

Setelah melakukan penelitian ini dapat didapatkan beberapa kesimpulan antara lain:

1. Sistem deteksi jenis cacat biji dengan algoritma *k-nearest neighbor* mampu mendeteksi jenis cacat dengan baik pada nilai k sebesar 3 fitur model warna HSV dan berlaku ketika menggunakan jarak *euclidean*, *manhattan*, maupun *minkowski*.
2. Akurasi dari sistem yang diusulkan sebesar 97%.

5.2. Saran

Adapun saran untuk mengembangkan penelitian ini adalah sistem dapat melakukan deteksi jenis cacat beberapa kopi sekaligus dalam satu foto.

DAFTAR PUSTAKA

- [1] E. Saputra, *Kopi*. Yogyakarta: Harmoni, 2008.
- [2] N. Sofiana, *1001 Fakta tentang Kopi*. Jakarta: Cahaya Atma, 2011.
- [3] A. Kadir, *Dasar Pengolahan Citra dengan Delphi*. Yogyakarta: ANDI Publisher, 2013.
- [4] Karmilasari dan A. Sumarna, "Temu Kenali Citra Berbasis Konten Warna," *Semin. Nas. Apl. Teknol. Inf. 2011*, 2011.
- [5] P. G. I. Widiarsana, O., Putra, N.W., Budiyasa dan S. N. Bismantara, A.N.I., Mahajaya, "Data Mining: Metode Clasification K-Nearest Neighbor (KNN)," 2011.
- [6] E. T. L. Kusrini, *Algoritma Data Mining*. Yogyakarta: ANDI Publisher, 2009.
- [7] R. Arthana, "Mengenal Accuracy, Precision, Recall dan Specificity serta yang diprioritaskan dalam Machine Learning," *Medium*, 2019. [Daring]. Tersedia pada: <https://medium.com/@rey1024/mengenal-accuracy-precision-recall-dan-specificity-septa-yang-diprioritaskan-b79ff4d77de8>. [Diakses: 30-Jul-2019].

- Comput. Cybern. Syst.*, vol. 11, no. 2, hal. 199–208, 2017.
- [11] M. Sholihin dan M. G. Rohman, “Klasifikasi Mutu Telur Berdasarkan Fitur Warna Dengan Menggunakan Metode K-Nearest Neighbor,” *Seminar Nasional Sistem Informasi (SENASIF)*, 2018, vol. 2, no. 1, hal. 1188–1193.