

Penerapan Metode *Local Binary Pattern* Dan *Random Forest* Pada Klasifikasi Roasted Biji Kopi

YUSUP MIFTAHUDDIN,S.KOM.,M.T. ¹, FIRMAN RESTU FAUZI ^{1*}

¹Institut Teknologi Nasional
Email : firmanrestu1@mhs.itenas.ac.id

Received 26 01 2023 | Revised 02 02 2023 | Accepted 02 02 2023

ABSTRAK

Biji kopi mengalami peningkatan dalam komoditas ekspor. Indonesia memiliki beberapa jenis kopi salah satunya adalah Robusta dan Arabika yang memiliki harga yang berbeda. Proses identifikasi biji kopi secara kasat mata akan sulit dilakukan jika tidak memiliki keahlian khusus. Jenis biji kopi dapat dilakukan dengan melihat bentuk dari biji kopi itu sendiri. Pada penelitian ini dilakukan identifikasi biji kopi Arabika dan Robusta dengan menerapkan metode Local Binary Pattern pada ekstraksi fitur dan Random Forest pada klasifikasi citra. Data citra kopi diambil dari daerah Garut, Jawa Barat Indonesia. Data diambil sebanyak 1000 yang dibagi ke dalam 800 citra latih dan 200 citra uji. Pengujian dilakukan pada citra yang melalui proses augmentasi dan tanpa proses augmentasi. Pada 100 – 200 data yang diuji didapatkan bahwa citra dengan proses augmentasi memiliki hasil akurasi tertinggi sebesar 96.5% dan citra dengan augmentasi mengalami peningkatan nilai akurasi.

Kata kunci: *Machine Learning, Biji Kopi, Ekstraksi Ciri*

ABSTRACT

Coffee beans have increased in export commodities. Indonesia has several types of coffee, one of which is Robusta and Arabica, which have different prices. The process of mentoring coffee beans in plain view will be difficult if you do not have special expertise. The type of coffee bean can be determined by looking at the shape of the coffee bean itself. In this research, we searched for Arabica and Robusta coffee beans by applying the Local Binary Pattern method to feature extraction and Random Forest to image classification. Coffee image data taken from the Garut area, West Java, Indonesia. 1000 data were taken which were divided into 800 training images and 200 test images. Tests were carried out on images that went through the augmentation process and without the augmentation process. At 100 – 200 data tested, it was found that images with augmentation process had the highest accuracy of 96.5% and images with augmentation experienced an increase in accuracy.

Keywords: *Machine Learning, Coffee Bean, Feature Extraction*

1. PENDAHULUAN

Biji kopi telah menjadi salah satu komoditas ekspor yang setiap tahunnya mengalami peningkatan. Pada tahun 2022 berdasarkan data dari bea cukai didapatkan bahwan ekspor biji kopi sebesar 849.373,2 Ton yang dikirim ke beberapa negara di Asia maupun Eropa. (www.bps.go.id). Pada tahun 2020, produk domestik bruto Indonesia mengalami penurunan sebesar -2.07 %, namun perdagangan biji kopi mentah mengalami peningkatan yang dapat mempengaruhi kinerja pasar Indonesia (**Renata, dkk.,2022**).

Indonesia memiliki beragam jenis biji kopi salah satunya Robusta dan Arabika yang tersebar pada setiap daerah. Jenis Robusta dan Arabika memiliki bentuk dan celah tengah yang berbeda seperti pada Gambar 1. Dengan tujuan untuk menjaga kualitas penampilan artikel di dalam jurnal, para penulis diharuskan memperhatikan dengan seksama seluruh ketentuan yang dijelaskan di bawah ini.



Gambar 1 Jenis Kopi Arabika Dan Robusta

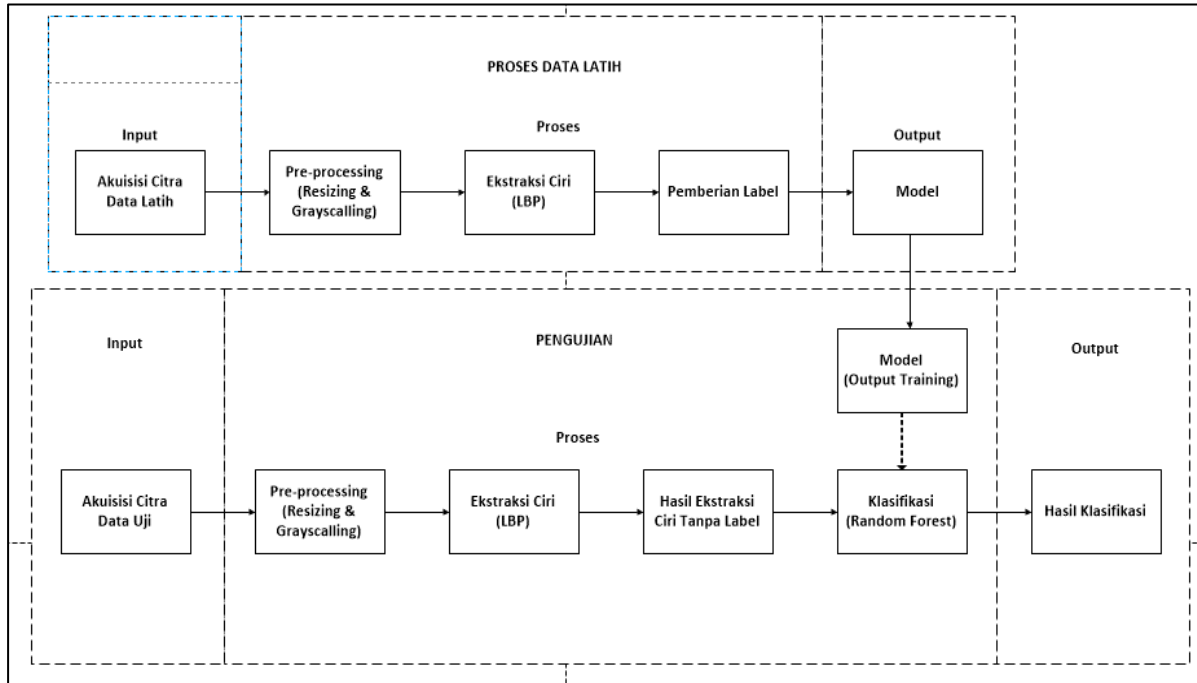
Proses identifikasi biji kopi secara kasat mata akan sulit dilakukan jika tidak memiliki keahlian khusus. Jenis biji kopi dapat dilakukan dengan melihat bentuk dari biji kopi itu sendiri. Arabika memiliki ciri bentuk yang memanjang dan memiliki celah tengah yang tidak lurus. Sedangkan Robusta memiliki ciri bentuk yang membulat dan memiliki celah tengah yang lurus. Jenis kopi Robusta dan Arabika memiliki harga yang berbeda, Arabika lebih mahal dibandingkan Robusta (**BAPPEBTI, 2022**).

Penelitian mengenai biji sudah dilakukan oleh peneliti sebelumnya yaitu identifikasi biji kopi dengan menggunakan pengolahan citra digital untuk mendapatkan ekstraksi fitur. (**Fadjeri, dkk.,2020**). Identifikasi kualitas kopi dengan menggunakan CBIR (*Content Based Information Retrieval*) dengan hasil meneliti tekstur dan warna pada biji kopi (**Prastyaningsih, dkk.,2020**).

Proses ekstraksi ciri dalam citra menggunakan *Local Binary Pattern* memiliki tingkat akurasi sebesar 85,6% dibandingkan dengan GLCM berdasarkan tekstur citra dengan jarak optimum 20 cm (**Neneng, dkk.,2021**). Metode *Random Forest* diterapkan pada proses klasifikasi citra memiliki akurasi 92,98% dibandingkan dengan metode SIFT pada klasifikasi batik (**Devella, dkk.,2020**) dan penerapan *Random Forest* dengan ekstraksi ciri menggunakan GLCM mendapatkan akurasi hanya 65% pada identifikasi penyakit padi (**Miftahuddin, dkk.,2022**). Metode *Random Forest* dapat mengurangi jumlah variabel yang diukur dan merupakan metode klasifikasi yang populer. (**Speiser, dkk.,2019**). *Random Forest* memiliki akurasi lebih baik dibandingkan dengan metode pohon keputusan lainnya (**Schonlau & Zou, 2020**). Penelitian ini menerapkan metode *Local Binary Pattern* dan *Random Forest* pada klasifikasi jenis biji kopi dengan tujuan untuk mengukur tingkat akurasi metode *Local Binary Pattern* (LBP) dan *Random Forest* (RF) dalam mengklasifikasi jenis biji kopi. Citra biji kopi untuk data latih diambil secara manual dengan jarak 20 cm dari kamera dengan format jpg dan ukuran 25 x 25 px. Data latih sebanyak 800 citra biji kopi dengan komposisi 400 biji kopi Robusta dan 400 biji kopi Arabika.

2. METODOLOGI PENELITIAN

Dalam mencapai tujuan penelitian ini, dilakukan beberapa tahapan yaitu pengambilan citra data latih dan proses data latih yang selanjutnya dilakukan pengujian dalam penerapan metode LBP dan Random Forest. Tahapan penelitian diperlihatkan pada Gambar 2.



Gambar 2 Alur Penelitian

2.1. Grayscale

Citra grayscale adalah sebuah citra yang hanya memiliki satu nilai kanal pada setiap pikselnya, dengan kata lain nilai bagian red = green = blue. (R.F.Gusa, 2013).



Gambar 3 Grayscale

Proses grayscale citra dapat dilakukan dengan perhitungan dengan formula pada persamaan (1).

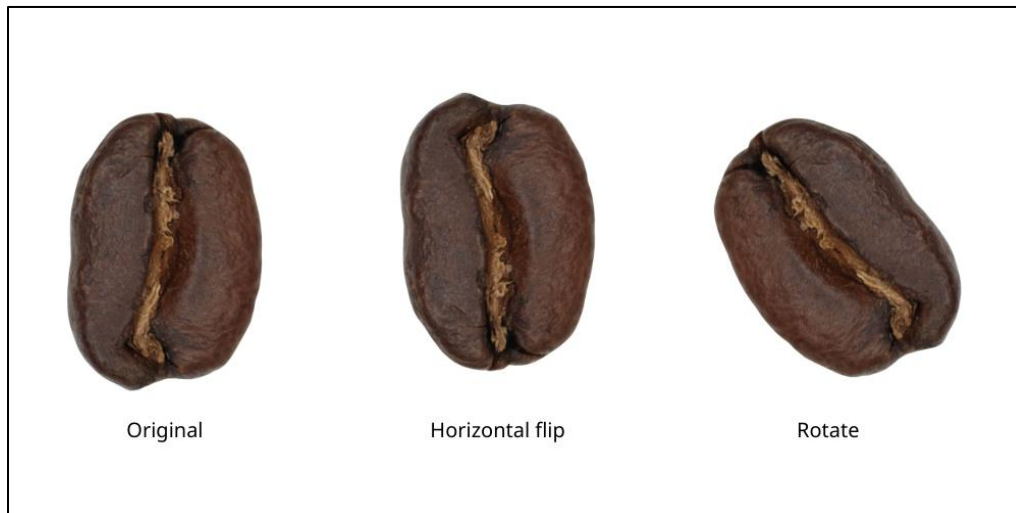
$$G(x, y) = 0.2989 * R(x, y) + 0.5870 * G(x, y) + 0.1141 * B(x, y) \quad (1)$$

Dimana:

- G(x, y) = Nilai dari citra grayscale.
- R(x, y) = Nilai piksel berwarna merah.
- G(x, y) = Nilai piksel berwarna hijau.
- B(x, y) = Nilai piksel berwarna biru.

2.2. Augmentasi Data

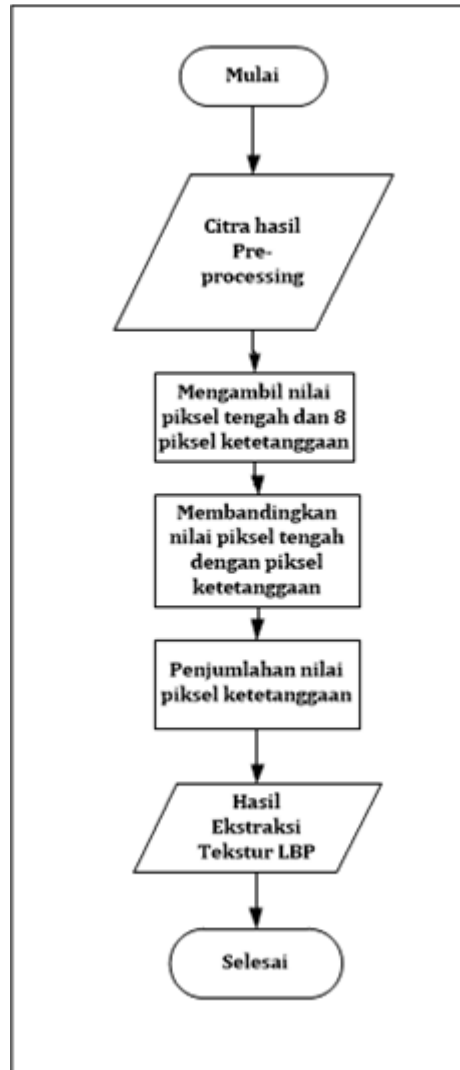
Augmentasi data adalah sebuah proses yang memungkinkan praktisi secara signifikan untuk meningkatkan keragaman data yang tersedia untuk model training, tanpa benar-benar mengumpulkan data baru. Teknik augmentasi data seperti *cropping*, *padding*, dan *flipping* horizontal umumnya digunakan untuk melatih jaringan neural besar. (Sanjaya, dkk, 2020).



Gambar 4 Augmentasi Data

2.3. Local Binary Pattern

Tahapan pertama adalah pengambilan citra latih biji kopi, pada penelitian ini sebanyak 800 citra biji kopi dengan 400 citra biji kopi Arabika dan 400 citra biji kopi Robusta. Citra latih yang diperoleh kemudian dilakukan proses *resizing* menjadi ukuran 25 x 25 piksel yang kemudian dilakukan *grayscale* yang akan dijadikan sebagai *input* untuk proses ekstraksi fitur dengan LBP. LBP merupakan salah satu fitur yang digunakan untuk klasifikasi tekstur. Operator LBP mengacu pada teknik yang melakukan transformasi gambar menjadi label bilangan *integer* atau *array* gambar yang memberikan deskripsi tampilan mikroskopik citra (Zeebaree, dkk, 2021).



Gambar 5 Flowchart LBP

Proses LBP dengan citra input grayscale diambil nilai piksel tengah dan 8 piksel ketetanggaannya, kemudian nilai piksel pada citra tersebut dibandingkan dengan piksel tetangganya, jika hasil lebih besar dari nol, maka bernilai 1, namun jika hasil pengurangan bernilai lebih kecil atau sama dengan nol maka bernilai 0. Perhitungan LBP dapat dilihat pada persamaan (2).

$$LBP_{P,R} = \sum_{p=0}^{p-1} s(I_{p,R} - I_c) 2^{P-1-p}$$

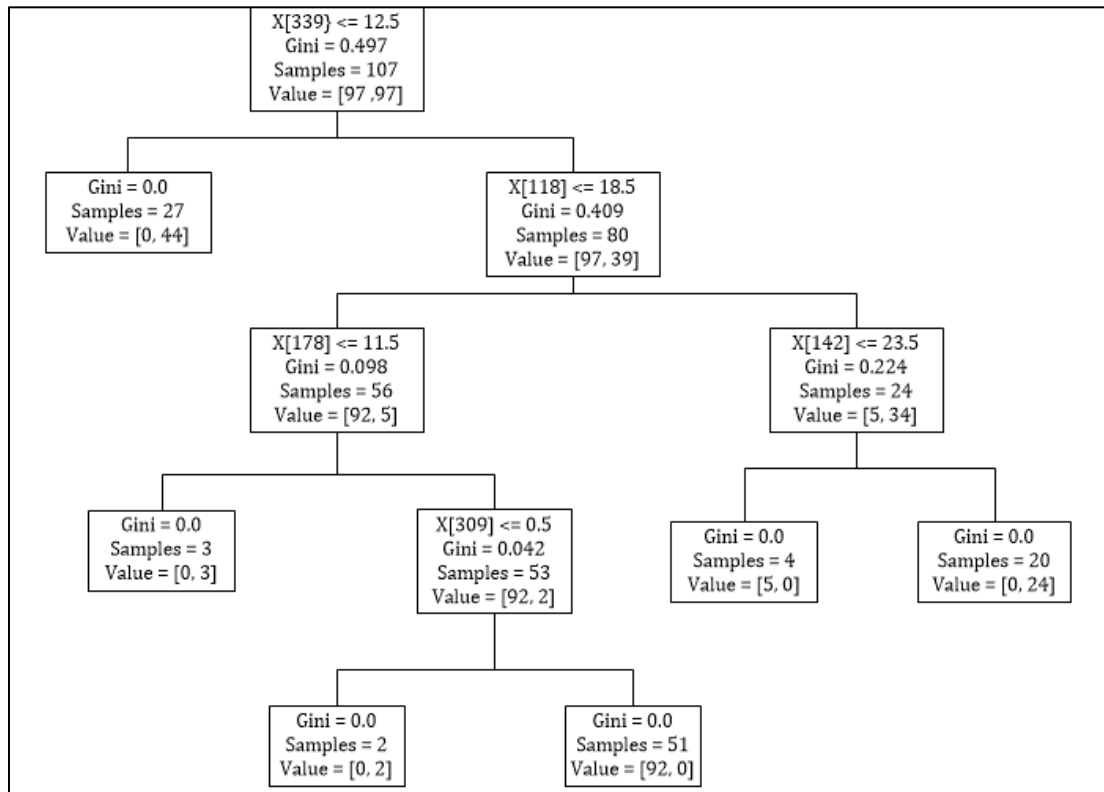
$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

Dimana:

- S = fungsi *thresholding*
- I = nilai intensitas piksel tetangga ke p
- P = jumlah banyak piksel tetangga
- R = radius antara titik pusat dan titik tetangga

2.4. Random Forest

Matriks hasil LBP dijadikan sebagai input untuk RF yang akan dijadikan sebagai klasifikasi. Hasil Random Forest berupa nilai 0 dan 1 dimana 0 untuk kopi jenis Robusta dan 1 untuk kopi jenis Arabika. Berikut hasil RF.



Gambar 6 Hasil *Random Forest*

Gambar 6 menampilkan salah satu hasil dari proses *Random Forest* yang diambil dari potongan matriks citra. Terdapat nilai *Gini*, *Samples* dan *Value*. Dari data latih yang sudah melalui proses *Random Forest* kemudian dibandingkan dengan data uji sehingga didapatkan sebuah prediksi probabilitas dalam bentuk *confusion matrix* yang dapat menghasilkan perhitungan *Accuracy*, *Precision*, *Recall*, dan *F-Measure*.

Tabel 1 Confusion Matrix

	Actual Yes	Actual No
Predicted Yes	True Positive (TP)	False Negative (FN)
Predicted No	False Posiitive (FP)	True Negative (TN)

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F - Measure = \frac{2 \times Recall \times Precision}{Recall+Precision} \quad (6)$$

3. DATA PENGAMATAN DAN ANALISA

Pengujian dilakukan 2 kali yaitu menggunakan data latih yang tidak melalui proses augmentasi data dan menggunakan data latih yang melalui proses augmentasi data rotasi 1 – 10°.

3.1 Pengujian Tanpa Augmentasi Data

Dalam pengujian citra menggunakan data latih yang tidak melalui proses augmentasi data dilakukan beberapa kali pengujian menggunakan jumlah data uji yang berbeda, yaitu 100, 120, 140, 160, 180, dan 200 data citra uji yang dapat dilihat pada tabel 2 sampai tabel 7.

Tabel 2 Hasil Pengujian 100 Data Tanpa Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	51,00%	100,00%	68,00%	52,00%
0	Arabika	100,00%	4,00%	8,00%	
Rata - rata		75,50%	52,00%	38,00%	

Tabel 2 didapatkan rata-rata nilai presisi 75,5%, *recall* 52%, *F-Measure* 38%, dan akurasi 52% dengan penggunaan data latih yang melalui augmentasi data dan 100 data citra uji.

Tabel 3 Hasil Pengujian 120 Data Tanpa Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	51,00%	100,00%	67,00%	51,67%
0	Arabika	100,00%	3,00%	6,00%	
Rata - rata		75,50%	51,50%	36,50%	

Tabel 3 didapatkan rata-rata nilai presisi 75,5%, *recall* 51,5%, *F-Measure* 36,5%, dan akurasi 51,67% dengan penggunaan data latih yang melalui augmentasi data dan 120 data citra uji.

Tabel 4 Hasil Pengujian 140 Data Tanpa Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	51,00%	100,00%	67,00%	51,43%
0	Arabika	100,00%	3,00%	6,00%	
Rata - rata		75,50%	51,50%	36,50%	

Tabel 4 didapatkan rata-rata nilai presisi 75,5%, *recall* 51,5%, *F-Measure* 36,5%, dan akurasi 51,43% dengan penggunaan data latih yang melalui augmentasi data dan 140 data citra uji.

Tabel 5 Hasil Pengujian 160 Data Tanpa Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	51,00%	100,00%	67,00%	51,25%
0	Arabika	100,00%	3,00%	5,00%	
Rata - rata		75,50%	51,50%	36,00%	

Tabel 5 didapatkan rata-rata nilai presisi 75,5%, *recall* 51,5%, *F-Measure* 36%, dan akurasi 51,25% dengan penggunaan data latih yang melalui augmentasi data dan 160 data citra uji.

Tabel 6 Hasil Pengujian 180 Data Tanpa Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	51,00%	100,00%	67,00%	51,67%
0	Arabika	100,00%	3,00%	6,00%	
Rata - rata		75,50%	51,50%	36,50%	

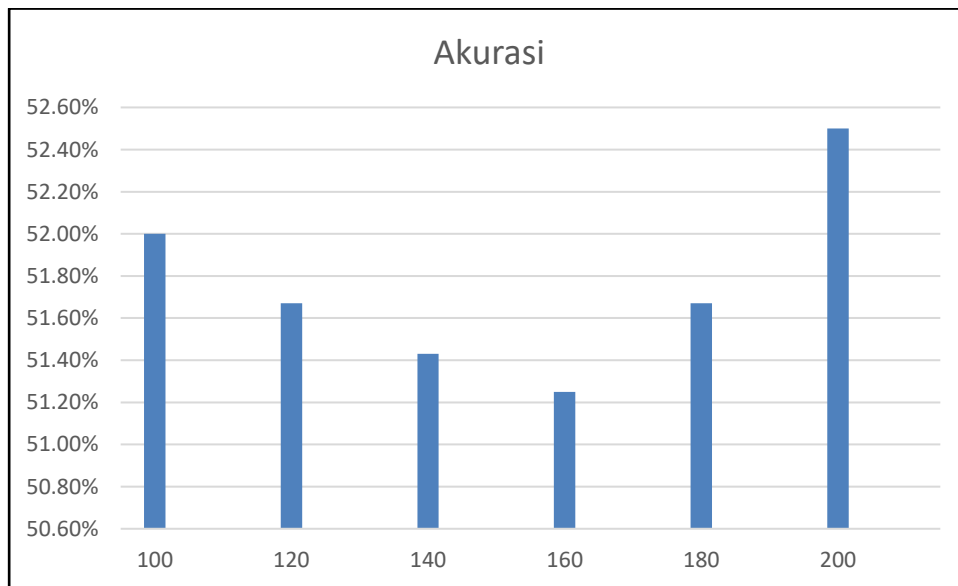
Tabel 6 didapatkan rata-rata nilai presisi 75,5%, *recall* 51,5%, *F-Measure* 36,5%, dan akurasi 51,67% dengan penggunaan data latih yang melalui augmentasi data dan 180 data citra uji.

Tabel 7 Hasil Pengujian 200 Data Tanpa Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	51,00%	100,00%	68,00%	52,50%
0	Arabika	100,00%	5,00%	10,00%	
Rata - rata		75,50%	52,50%	39,00%	

Tabel 7 didapatkan rata-rata nilai presisi 75,5%, *recall* 52,5%, *F-Measure* 39%, dan akurasi 52,5% dengan penggunaan data latih yang melalui augmentasi data dan 200 data citra uji.

Dari semua pengujian yang dilakukan menggunakan data latih yang tidak melalui proses augmentasi data dapat dibuat sebuah grafik pengujian dari 100, 120, 140, 160, 180, dan 200 data citra uji yang dapat dilihat pada gambar 7.



Gambar 7. Grafik Pengujian Tanpa Augmentasi Data.

3.2 Pengujian Dengan Augmentasi Data

Dalam pengujian citra menggunakan data latih yang melalui proses augmentasi data rotasi 1-10° dilakukan beberapa kali pengujian menggunakan jumlah data uji yang berbeda, yaitu 100, 120, 140, 160, 180, dan 200 data citra uji yang dapat dilihat pada tabel 8 sampai tabel 13.

Tabel 8 Hasil Pengujian 100 Data Dengan Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	88,00%	98,00%	92,00%	92,00%
0	Arabika	100,00%	86,00%	91,00%	
Rata - rata		94,00%	92,00%	91,50%	

Tabel 8 didapatkan rata-rata nilai presisi 94%, *recall* 92%, *F-Measure* 91,5%, dan akurasi 92% dengan penggunaan data latih yang melalui augmentasi data dan 100 data citra uji.

Tabel 9 Hasil Pengujian 120 Data Dengan Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	90,00%	100,00%	94,00%	94,17%
0	Arabika	100,00%	88,00%	94,00%	
Rata - rata		95,00%	94,00%	94,00%	

Tabel 9 didapatkan rata-rata nilai presisi 95%, *recall* 94%, *F-Measure* 94%, dan akurasi 94,17% dengan penggunaan data latih yang melalui augmentasi data dan 120 data citra uji.

Tabel 10 Hasil Pengujian 140 Data Dengan Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	90,00%	100,00%	67,00%	94,29%
0	Arabika	100,00%	89,00%	6,00%	
Rata - rata		95,00%	94,50%	94,50%	

Tabel 10 didapatkan rata-rata nilai presisi 95%, *recall* 94,5%, *F-Measure* 94,5%, dan akurasi 94,29% dengan penggunaan data latih yang melalui augmentasi data dan 140 data citra uji.

Tabel 11 Hasil Pengujian 160 Data Dengan Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	92,00%	100,00%	96,00%	95,63%
0	Arabika	100,00%	91,00%	95,00%	
Rata - rata		96,00%	95,50%	95,50%	

Tabel 11 didapatkan rata-rata nilai presisi 96%, *recall* 95,5%, *F-Measure* 95,5%, dan akurasi 95,63% dengan penggunaan data latih yang melalui augmentasi data dan 160 data citra uji.

Tabel 12 Hasil Pengujian 180 Data Dengan Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	94,00%	100,00%	97,00%	96,27%
0	Arabika	100,00%	92,00%	96,00%	
Rata - rata		97,00%	96,00%	96,50%	

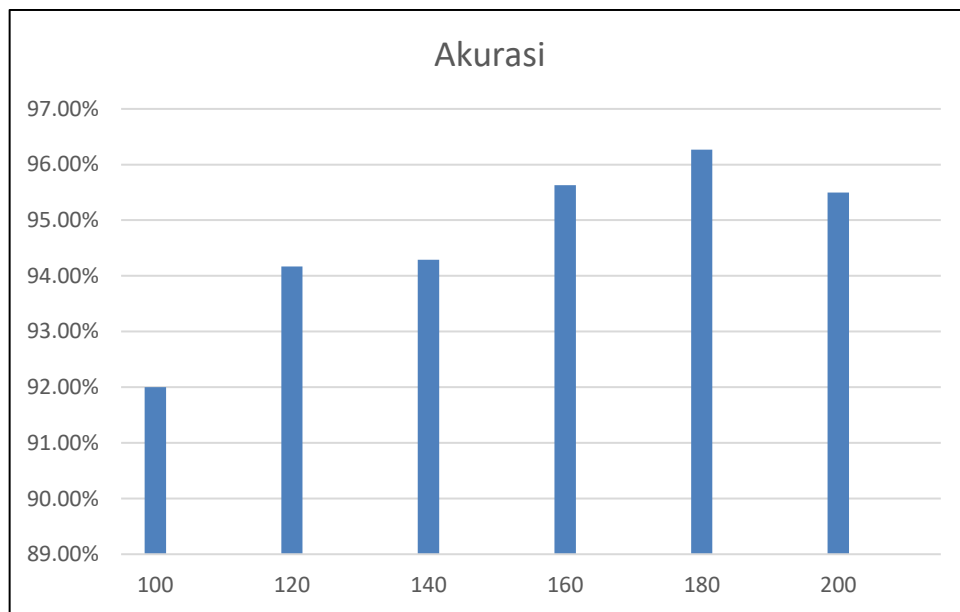
Tabel 12 didapatkan rata-rata nilai presisi 97%, *recall* 96%, *F-Measure* 96,5%, dan akurasi 96,27% dengan penggunaan data latih yang melalui augmentasi data dan 180 data citra uji.

Tabel 13 Hasil Pengujian 200 Data Dengan Augmentasi Citra

Value	Kelas	Presisi	Recall	F-Measure	Akurasi
1	Robusta	93,00%	100,00%	97,00%	96,50%
0	Arabika	100,00%	93,00%	96,00%	
Rata - rata		96,50%	96,50%	96,50%	

Tabel 13 didapatkan rata-rata nilai presisi 96,5%, *recall* 96,5%, *F-Measure* 96,5%, dan akurasi 96,5% dengan penggunaan data latih yang melalui augmentasi data dan 200 data citra uji.

Dari semua pengujian yang dilakukan menggunakan data latih yang melalui proses augmentasi data rotasi 1-10° dapat dibuat sebuah grafik pengujian dari 100, 120, 140, 160, 180, dan 200 data citra uji yang dapat dilihat pada gambar 8.



Gambar 8. Grafik Pengujian Dengan Augmentasi Data.

4. KESIMPULAN

Penelitian ini bertujuan untuk mengidentifikasi jenis biji kopi arabika atau robusta dan mengklasifikasi citra biji kopi berdasarkan hasil identifikasinya. Dengan tahapan mulai dari *preprocessing*, ekstraksi fitur dengan LBP (*Local Binary Pattern*) proses identifikasi dan *Random Forest* untuk klasifikasi. Metode ekstraksi ciri LBP dapat digunakan untuk mengidentifikasi biji kopi dibuktikan berdasarkan hasil pengujian yang ditunjukkan pada pengujian latih menggunakan augmentasi data dan tanpa augmentasi data serta pengujian ekstraksi ciri, lalu metode klasifikasi *Random Forest* dalam mengklasifikasi data uji berdasarkan hasil identifikasi dapat digunakan hal ini dibuktikan pada pengujian kinerja sistem yang mendapatkan nilai akurasi 96,50% dengan data latih yang melalui augmentasi data yang diambil dengan intensitas cahaya 4107 lux dan jarak pengambilan 20cm dan nilai akurasi 52,50% dengan data latih yang tidak melalui augmentasi data yang diambil dengan intensitas cahaya 4107 lux dan jarak pengambilah 20cm. Akurasi merupakan gambaran dari data uji yang terklasifikasi dengan benar oleh *Random Forest*.

DAFTAR PUSTAKA

- Ahmasetyosari, A. S., & Fatimah, T. (2018). Aplikasi Presensi Siswa Pada PT.Samudera Anugrah Menggunakan Metode Geofencing dan Perhitungan Jarak Menggunakan Algoritma Euclidean Distance Berbasis Android. *Skanika Vol. 1 No.2*, 481-485.
- Aldya, A. P., Rahmatullah, A., & Fachurroji. (2019). Haversine Formula Untuk Membatasi Jarak Pada Aplikasi Presensi Online. *Jurnal Informasi Sains Dan Teknologi, Vol.4, No.2*, 171-180.
- Buaton, R., Sundari, Y., & Maulita, Y. (2016). Clustering Tindak Kekerasan Pada Anak Menggunakan Algoritma K-Means Dengan Perbandingan Jarak Kedekatan Manhattan City Dan Euclidean. *Media Informasi Analisa dan Sistem, Vol.1, No.2*, 47-53.
- Derisma, Firdaus, & Yusya, R. P. (2016). Perancangan Ikat Pinggang Elektronik Untuk Tunanetra Menggunakan Mikrokontroller Dan Global Positioning System (GPS) Pada Smartphone Android. *Jurnal Teknik Elektro ITP, Volume 5, No.2*, 130-136.
- Farid, & Yunus, Y. (2017). Analisa Algoritma Haversine Formula Untuk Pencarian Lokasi Terdekat Rumah Sakit Dan Puskesmas . *ILKOM Jurnal Ilmiah Vol.9 No.3*, 353-355.
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma KMeans Clustering berbasis Chi-Square. *Jurnal Pengebangan IT, Vol.04, No.01*, 20-24.
- Okilas, A. F., Siswanti, S. D., & M Dieka, R. (2015). Akurasi Pembacaan GPS pada Android untuk Location Based Service (Studi Kasus: Informasi Lokasi SMA di Palembang). *IKAI Vol. 4 No. 1*, 1-5.
- Pamungkas, C. A. (2019). Aplikasi Penghitung Jarak Koordinat Berdasarkan Latitude Dan Longitude Dengan Metode Euclidean Distance Dan Metode Haversine. *INFORMA Politeknik Indonusa Surakarta Vol.5 No.2*, 8-13.
- Susilo, Y. S., Pranjoto, H., & Gunadhi, A. (2014). Sistem Pelacakan Dan Pengamanan Kendaraan Berbasis Gps Dengan Menggunakan Komunikasi Gprs. *Jurnal Ilmiah Widya Teknik*, 22-25.

- Yulianto, Ramadiani, & Kridalaksana, A. H. (2018). Penerapan Formula Haversine Pada Sistem Informasi Geografis Pencarian Jarak Terdekat Lokasi Lapangan Futsal. *Jurnal Ilmiah Ilmu Kmputer, Vol.13, No.1*.
- Yushar, I., Purnama, I. P., Sutardi, & Aksara, L. B. (2019). Pengenalan Wajah Berbasis Perhitungan Jarak Fitur LBP Menggunakan Euclidean, Manhattan, Chi Square Distance. *SEMNASITIK*, 386-393.
- "Assignment_13_ravide.ipynb." BC-ML-IGLO_Assigment. 2020. Github. 11 oct. 2020.<https://github.com/ravide-lubis/BC-ML_IGLO_Assigment/Assignment_13_ravide.ipynb>.