

## EKSTRAKSI FITUR CIRCULARITY UNTUK PENGENALAN VARIETAS KOPI ARABIKA

Maria Mediatix Sebatubun<sup>1</sup>, Muhammad Agung Nugroho<sup>2</sup>

<sup>1,2</sup>Teknik Informatika, STMIK AKAKOM Yogyakarta  
Email: <sup>1</sup>memey@akakom.ac.id, <sup>2</sup>m.agung.n@akakom.ac.id

(Naskah masuk: 3 November 2017, diterima untuk diterbitkan: 30 Desember 2017)

### Abstrak

Kopi merupakan salah satu minuman yang sangat populer di dunia dan digemari oleh banyak orang termasuk di Indonesia. Kopi terdiri dari berbagai varietas, salah satunya adalah varietas arabika. Varietas kopi dapat memiliki kenampakan yang berbeda – beda misalnya seperti perbedaan warna, bentuk, ataupun tekstur. Oleh karena itu, terkadang petani ataupun pemilik coffee shop dapat melakukan kesalahan dalam mengenali varietas kopi arabika yang dijual ataupun yang dibeli. Hal ini juga akan mempengaruhi penentuan harga kopi tersebut, karena masing-masing varietas kopi arabika memiliki harga yang berbeda-beda. Untuk itu, diperlukan sistem yang juga mampu mengenali varietas kopi arabika secara akurat sehingga dapat digunakan sebagai *second opinion* bagi para petani ataupun pemilik *coffee shop* dalam mengenali varietas kopi arabika. Salah satu cara yang dapat dilakukan adalah dengan metode pencitraan. Tahap awal yang dilakukan adalah praproses yaitu *cropping* citra yang dilakukan secara manual, kemudian segmentasi menggunakan metode Otsu. Tahap selanjutnya adalah ekstraksi fitur bentuk menggunakan *circularity* dan klasifikasi menggunakan MultiLayer Perceptron. Hasil klasifikasi menunjukkan tingkat akurasi yang diperoleh sebesar 80%, sensitivitas 83,33% dan spesifisitas 76,7%.

**Kata kunci:** ekstraksi fitur, kopi, klasifikasi, segmentasi

### Abstract

*Coffee is one of the most popular beverages in the world and is favored by many people including Indonesians. Coffee consists of many variety, one of them is arabica. Coffee variety can have different features such as differences in color, shape, or texture. Therefore, sometimes farmers of coffee shop owners can make mistakes in recognizing the variety of arabica coffee that are sold or purchased. This will also affect the determination of coffee's price, because each variety of arabica coffee have different prices. Hence, a capable system which can recognize arabica coffee accurately is required to be used as a second opinion for farmers or coffee shop owners in recognizing these variety. One of the methods that can be done is imaging. The initial stage is pre-processing by cropping the images manually, followed by segmentation using Otsu method. The next stage is shape based feature extraction using circularity and the last is classification using MultiLayer Perceptron. Classification results show 80% level of accuracy, 83.33% sensitivity, and 76.7% specificity.*

**Keywords:** feature extraction, coffee, classification, segmentation

## 1. PENDAHULUAN

Kopi merupakan salah satu komoditas di dunia yang dibudidayakan di berbagai negara, termasuk Indonesia. Pemrosesan kopi sebelum dapat diminum dilakukan melalui proses yang cukup panjang, dimulai dari proses panen biji kopi yang telah matang, kemudian proses pengeringan dilanjutkan proses sangrai dengan menggunakan suhu yang bervariasi. Proses terakhir adalah penggilingan biji kopi yang telah disangrai menjadi bubuk kopi. Jenis kopi yang paling populer adalah arabika dan robusta. Masing-masing jenis kopi ini masih memiliki varietas yang cukup banyak dengan rasa yang berbeda-beda tergantung pada varietas kopi tersebut. Di Indonesia, juga terdapat jenis kopi arabika dengan berbagai varietas yang dibudidayakan di berbagai daerah. Masing – masing varietas kopi arabika tersebut memiliki kenampakan yang berbeda

– beda misalnya seperti perbedaan warna, bentuk, ataupun tekstur. Oleh karena itu, terkadang petani ataupun para pemilik *coffee shop* dapat melakukan kesalahan dalam mengenali varietas kopi arabika yang dijual ataupun yang dibeli. Hal ini juga akan mempengaruhi penentuan harga kopi tersebut, karena masing-masing varietas kopi arabika memiliki harga yang berbeda-beda. Untuk itu, diperlukan sistem yang juga mampu mengenali varietas kopi arabika secara akurat sehingga dapat digunakan sebagai *second opinion* bagi para petani ataupun pemilik *coffee shop* dalam mengenali varietas kopi arabika tersebut.

Salah satu cara yang dapat dilakukan adalah dengan metode pencitraan. Biji kopi yang telah diproses, dipotret menggunakan kamera Canon 750D dengan jarak yang tidak konstan. Tujuannya adalah agar sistem yang dihasilkan lebih fleksibel saat proses pemotretan citra. Kemudian citra hasil

pemotretan tersebut diproses menggunakan teknik-teknik pengolahan citra. Terdapat beberapa penelitian tentang pengenalan varietas kopi arabika yang telah dilakukan sebelumnya menggunakan berbagai teknik pengolahan citra. Salah satu teknik yang sering digunakan adalah ekstraksi fitur yang bertujuan untuk mengambil fitur atau ciri dari sebuah objek. Teknik ini juga dapat digunakan untuk mengambil fitur yang dimiliki biji kopi sehingga sistem dapat mengenali varietas kopi arabika berdasarkan fitur-fitur yang diperoleh.

Salah satu metode ekstraksi fitur yang telah umum digunakan adalah fitur tekstur. Metode ini akan mengambil fitur-fitur tekstur pada permukaan biji kopi dengan menggunakan perhitungan statistik (Kadir and Adhi 2012). Fitur tekstur terdiri dari fitur tekstur orde satu dan fitur tekstur orde dua. Terdapat penelitian untuk pengenalan varietas biji kopi dilakukan dengan menggunakan fitur tekstur orde dua yang dikenal dengan *Gray Level Co-occurrence Matrices* (GLCM) (Condori et al. 2014), (Apaza et al. 2014). Selain fitur tekstur, terdapat fitur bentuk yang juga sering digunakan oleh para petani maupun pemilik *coffee shop* untuk membedakan varietas kopi secara manual. Oleh karena itu, penelitian ini akan melakukan pengenalan varietas kopi arabika yaitu varietas Sigarar Utang dan varietas Lini S-795 berdasarkan fitur bentuk menggunakan teknik pengolahan citra. Setelah mendapat fitur yang diinginkan, tahap selanjutnya yang perlu dilakukan adalah melakukan klasifikasi yang bertujuan untuk mengenali varietas kopi arabika.

## 2. TINJAUAN PUSTAKA

Terdapat beberapa penelitian tentang pengenalan varietas kopi arabika dengan menggunakan teknik pengolahan citra. Tahap-tahap yang dilakukan terdiri dari ekstraksi fitur dan klasifikasi. Salah satunya adalah penelitian (Condori et al. 2014) yang melakukan pengenalan terhadap biji kopi dengan menggunakan citra sampel sebanyak 3367 citra. Tahap awal yang dilakukan adalah segmentasi untuk memisahkan *background* dengan objek pada citra menggunakan *contours algorithm*, kemudian tahap selanjutnya adalah ekstraksi fitur menggunakan GLCM dengan empat sudut arah yang berbeda yaitu  $0^{\circ}$ ,  $45^{\circ}$ ,  $90^{\circ}$  dan  $135^{\circ}$ . Beberapa fitur GLCM yang digunakan adalah *contrast*, *dissimilarity*, *homogeneity*, *energy*, *entropy*, *correlation* dan *variance*. Tahap terakhir adalah klasifikasi dengan metode Support Vector Machine (SVM). Hasil yang diperoleh adalah tingkat akurasi mencapai 86%. Penelitian tentang klasifikasi biji kopi (Apaza et al. 2014), juga dilakukan dengan menggunakan metode lain. Tahap awal yang dilakukan adalah *pre-processing* terhadap citra digital dengan algoritme *Multi-scale Retinex with Color Restoration* (MSRCR). Metode tersebut kemudian dibandingkan dengan metode peningkatan kualitas citra yang berbeda yaitu *Histogram Equalization* dan *Contrast*

*Limited Adaptive Histogram Equalization* (CLAHE). Tahap selanjutnya adalah ekstraksi fitur menggunakan metode *Color GLCM* dengan empat sudut arah yang berbeda yaitu  $0^{\circ}$ ,  $45^{\circ}$ ,  $90^{\circ}$  dan  $135^{\circ}$ . Tahap terakhir adalah klasifikasi menggunakan SVM. Hasil yang diperoleh menunjukkan bahwa kualitas visual dan akurasi lebih baik jika menggunakan metode *pre-processing* MSRCR.

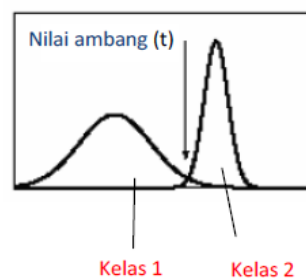
Penelitian lain juga dilakukan (De Oliveira et al. 2016) yaitu pengenalan terhadap biji kopi berdasarkan warna. Metode Artificial Neural Network (ANN) digunakan sebagai model transformasi kemudian NBC digunakan untuk mengenali biji kopi yang terdiri dari empat jenis yaitu : whitish, cane green, green, dan bluish-green. Metode ANN hanya mencapai tingkat eror sebesar 1,15% dan NBC menghasilkan tingkat akurasi sebesar 100%. Untuk memperoleh lebih banyak fitur, maka terdapat penelitian (Radi, Rivai, and Purnomo 2015) yang juga mengusulkan sistem untuk mengenali varietas kopi arabika menggunakan beberapa tahap yaitu ekstraksi fitur dengan menggabungkan metode fitur tekstur orde satu dan orde dua. Tahap klasifikasi dilakukan menggunakan metode ANN dan diperoleh rata-rata akurasi sebesar 80%. Penelitian terdahulu cenderung melakukan ekstraksi fitur berdasarkan pada fitur tekstur. Tetapi penelitian ini akan mengusulkan metode ekstraksi fitur berdasarkan fitur bentuk. Proses klasifikasi akan menggunakan MultiLayer Perceptron (MLP) yang merupakan pengembangan dari ANN dan telah banyak diterapkan pada kasus yang berbeda (Nugroho, Sebatubun, and Adji 2017).

## 3. LANDASAN TEORI

Sebelum ekstraksi fitur, dilakukan proses *cropping* yang dilakukan secara manual sehingga fokus penelitian hanya pada bagian biji kopi. Hasil dari proses *cropping* tersebut digunakan sebagai citra masukan untuk proses selanjutnya.

### 3.1 Tahap Segmentasi

Proses segmentasi dilakukan dengan menggunakan metode Otsu karena metode ini dapat memilih *threshold* paling optimal secara otomatis dan stabil karena didasarkan pada histogram citra (Ratri, Nugroho, and Adji 2015). Metode ini digunakan untuk melakukan konversi citra *grayscale* menjadi citra biner.



Gambar 1 Penentuan nilai ambang

Metode Otsu menentukan nilai ambang dengan cara membedakan dua kelompok yaitu objek dan latar belakang, yang memiliki bagian yang saling bertumpukan berdasarkan histogram seperti pada Gambar 3 (Susanto and Kadir 2012). Prinsip kerja metode Otsu yaitu pertama-tama menghitung probabilitas nilai intensitas  $i$  dalam histogram, dinormalisasikan dan didistribusikan dalam persamaan(1) berikut (Ratri et al. 2015).

$$p_i = \frac{n_i}{N}, p_i \geq 0, \sum_{i=1}^L p_i = 1 \quad (1)$$

Keterangan :

- $p_i$  = normalisasi jumlah piksel dengan intensitas  $i$
- $L$  = level keabuan pada citra
- $n_i$  = jumlah piksel pada level keabuan ke- $i$
- $N$  = jumlah total piksel

Selanjutnya dari persamaan(1) di atas, akan dilakukan pembagian piksel-piksel tersebut menjadi dua kelas  $C_0$  dan  $C_1$  (*background* dan objek dalam citra, atau sebaliknya) dengan sebuah threshold pada level keabuan  $k$ .  $C_0$  akan berisi piksel dengan level keabuan  $(1, \dots, k)$  dan  $C_1$  akan berisi piksel dengan level keabuan  $(k+1, \dots, L)$ . Maka probabilitas dari terjadinya kelas  $\omega_0$  dan  $\omega_1$  (*background* dan objek dalam citra) dan rata-rata level pada kelas  $\mu_0$  dan  $\mu_1$  (*background* dan objek dalam citra) didapatkan dari persamaan(2) hingga persamaan(7) berikut (Ratri et al. 2015).

$$\omega_0 = \Pr(C_0) = \sum_{i=1}^k p_i = \omega(k), \quad (2)$$

$$\omega_1 = \Pr(C_1) = \sum_{i=k+1}^L p_i = 1 - \omega(k), \quad (3)$$

dan 
$$\mu_0 = \sum_{i=1}^k i \Pr(i | C_0) = \sum_{i=1}^k i p_i / \omega_0 = \mu(k) / \omega(k), \quad (4)$$

$$\mu_1 = \sum_{i=k+1}^L i \Pr(i | C_1) = \sum_{i=k+1}^L i p_i / \omega_1 = \frac{\mu_T - \mu(k)}{1 - \omega(k)}, \quad (5)$$

dengan 
$$\omega(k) = \sum_{i=1}^k p_i, \quad (6)$$

dan 
$$\mu(k) = \sum_{i=1}^k i p_i. \quad (7)$$

Persamaan(7) di atas adalah nilai *zeroth* dan untuk mengetahui total rata-rata level keabuan  $\mu_T$  dari citra dengan persamaan(8) berikut (Ratri et al. 2015) :

$$\mu_T = \mu(L) = \sum_{i=1}^L i p_i \quad (8)$$

Kemudian untuk memudahkan dalam melakukan verifikasi terhadap beberapa relasi untuk semua pilihan threshold  $k$  yang ada, maka dapat menggunakan persamaan(9) (Ratri et al. 2015) :

$$\omega_0 \mu_0 + \omega_1 \mu_1 = \mu_T, \omega_0 + \omega_1 = 1 \quad (9)$$

Varian dari kedua kelas  $\sigma_0^2$  dan  $\sigma_1^2$  (*background* dan objek dalam citra) yang ada dapat diperoleh dari

persamaan(10) berikut.

$$\sigma_0^2 = \sum_{i=1}^k (i - \mu_0)^2 \Pr(i | C_0) = \sum_{i=1}^k (i - \mu_0)^2 p_i / \omega_0, \quad (10)$$

$$\sigma_1^2 = \sum_{i=k+1}^L (i - \mu_1)^2 \Pr(i | C_1) = \sum_{i=k+1}^L (i - \mu_1)^2 p_i / \omega_1 \quad (11)$$

### 3.2 Tahap Ekstraksi Fitur

Dalam berbagai fitur geometris, *circularity* (sifat bundar) merupakan salah satu fitur penting yang umum digunakan. *Circularity* biasanya didefinisikan dalam persamaan(12) berikut (Fu and Han 2012) :

$$CI = 4\pi * \frac{area}{perimeter^2} \quad (12)$$

$CI$  merupakan keluaran yang bernilai 1 jika objek mendekati bentuk lingkaran. Perimeter atau keliling menyatakan panjang tepi suatu objek yang dapat dihitung dengan persamaan(13) berikut (Fu and Han 2012) :

$$perimeter = \sum_{i=1}^n l_i \quad (13)$$

$l_i$  merupakan piksel-piksel pada tepi objek yang diperoleh dengan menggunakan kode rantai empat arah yang didefinisikan dengan persamaan(14) berikut (Fu and Han 2012) :

$$l_i = \overline{P_i P_{i-1}} = 1, \varepsilon_i = 0, 1, 2, 3, \quad (14)$$

Perimeter terdiri dari  $P_0, P_1, \dots, P_{n-1}, P_0$ , dan  $n$  merupakan total area yang diperoleh dengan menggunakan persamaan(15) berikut (Fu and Han 2012) :

$$area = \sum_{i=1}^n a_{ix} (y_{i-1} \frac{1}{2} a_{iy}), \quad (15)$$

Dengan

$$y_i = \sum_{j=1}^i a_{jy} + y_0 \quad (16)$$

$n$  adalah jumlah kode rantai,  $a_{ix}$  dan  $a_{iy}$  merupakan komponen  $x$  dan  $y$  dari dari rantai arah,  $y_i$  adalah koordinat  $y$  pada masing-masing titik,  $y_0$  adalah koordinat  $y$  untuk titik mulai. Algoritme akan menghitung perimeter dan area terlebih dulu menggunakan metode kode rantai dan diperoleh nilai  $CI$ . Dalam implementasi, sebuah bentuk bundar yang standar sulit didapat sehingga ditetapkan nilai threshold  $s$  di awal. Ketika  $CI$  lebih besar dari  $s$ , lingkaran yang terdeteksi akan ditampilkan berdasarkan kriteria  $s$ .

### 3.3 Tahap Klasifikasi

Multilayer perceptron merupakan pengembangan dari model Perceptron Neural Network yang dikembangkan pada awal 1960-an. Neural Network

memiliki banyak lapisan yang terbatas untuk mengurangi waktu untuk memecahkan masalah yang ada (Ratri et al. 2015). Multilayer Perceptron adalah contoh dari ANN yang biasanya digunakan untuk memberikan solusi untuk masalah yang berbeda, misalnya untuk pengenalan pola dan interpolasi (Noriega 2005).

Diperlukan beberapa langkah untuk menjalankan klasifikasi ANN menggunakan arsitektur MLP, yaitu dimulai dengan pengumpulan data, kemudian membuat dan mengkonfigurasi jaringan. Selanjutnya menginisialisasi bobot dan bias. Setelah jaringan dapat melakukan pelatihan, validasi data dan digunakan selama klasifikasi. Kelemahan perceptron adalah tempat perceptron hanya dapat memecahkan masalah yang dapat dipisahkan secara linear (Ratri et al. 2015).

Secara umum ANN adalah sebuah unit proses yang memiliki input dan mengeluarkan output, dengan neuron-neuron diorganisasikan sebagai layer. Output dari ANN dapat dikomputasikan dengan persamaan(17) berikut (Ratri et al. 2015).

$$O = f(IW_{io}) \quad (17)$$

Keterangan:

$W_{io}$ = Matriks beban (*weight matrix*), ukuran  $i \times o$

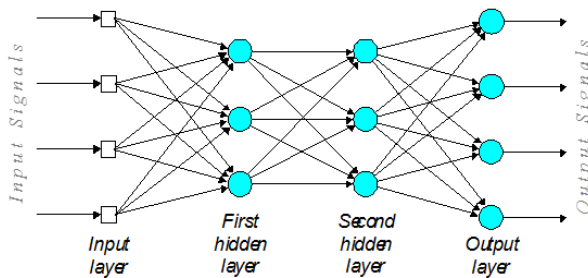
$i$  = jumlah *input* node

$o$  = jumlah *output* node

$I$  = input vector

$O$  = output vector

Secara umum data dipresentasikan dalam input layer, kemudian network akan melakukan proses input dengan mengalikan input dengan layer *weight* (beban). Prinsip kerja ANN ini hampir sama dengan MLP yang dapat ditunjukkan dalam Gambar 2.



Gambar 2 Arsitektur MLP dengan 2 hidden layer

Adapun untuk mempermudah dalam memahami cara kerja MLP, dapat digunakan algoritme yang telah disampaikan L.Noriega (Noriega 2005), sebagai berikut.

1. Melakukan inisialisasi network, dengan semua beban diset secara random antara angka -1 sampai dengan +1.
2. Mempresentasikan pola pelatihan pertama pada network yang ada, dan menyimpan hasil output.
3. Membandingkan output network tersebut dengan

output target yang ada.

4. Memperbaiki eror secara *backward*.

- a. Memperbaiki layer beban dari output dengan formula.

$$\omega_{ho} = \omega_{ho} + (\eta \delta_o o_h) \quad (18)$$

Keterangan :

$\omega_{ho}$ =nilai beban dari unit h yang tersembunyi

dengan output unit o,

$\eta$  = rasio pelatihan,

$o_h$  = output dari unit h yang tersembunyi. Dimana

$$\delta_o = o_o(1 - o_o)(t_o - o_o) \quad (19)$$

Keterangan :

$O_o$  = node o dari output layer

$t_o$  = target output untuk node tersebut

- b. Memperbaiki beban input dengan formula.

$$\omega_{ih} = \omega_{ih} + (\eta \delta_h o_i) \quad (20)$$

Keterangan:

$\omega_{ih}$ =nilai beban dari unit h yang tersembunyi dengan

input unit i,

$\eta$  = rasio pelatihan,

$o_i$  = input dari node i. Dimana

$$\delta_h = o_h(1 - o_h) \sum_o (\delta_o \omega_{ho}) \quad (21)$$

5. Melakukan perhitungan eror, dengan menghitung rata-rata dari nilai target dan output vector. Fungsi berikut dapat digunakan untuk menghitung eror tersebut.

$$E = \frac{\sqrt{\sum_{n=1}^p (t_o - o_o)^2}}{p} \quad (22)$$

Keterangan :

$P$  = jumlah unit pada output layer

$n$  = unit input layer

6. Mengulangi langkah no.2 untuk setiap pola pada dataset pelatihan untuk melengkapi satu *epoch*.
7. Melakukan pertukaran dataset pelatihan secara random. Hal ini untuk mengurangi kemungkinan network dipengaruhi oleh urutan pada data.
8. Mengulangi langkah 2 untuk sejumlah epochs atau hingga eror mulai berubah.

### 3.4 Indeks Pengukuran

Pengukuran yang dilakukan dalam penelitian ini adalah untuk mengetahui tingkat kesuksesan terhadap proses-proses yang telah dilakukan. Pengukuran yang akan dilakukan yaitu pengukuran terhadap kinerja dari metode ekstraksi fitur maupun metode klasifikasi. Pengukuran dari proses klasifikasi ditentukan dengan nilai-nilai berikut:

1. Akurasi

Nilai akurasi dari hasil klasifikasi dapat diperoleh dengan menghitung jumlah klasifikasi yang benar

dan sesuai target dibagi dengan jumlah klasifikasi yang berbeda dengan target dari semua kelas. Akurasi dirumuskan dalam persamaan(23) berikut

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \quad (23)$$

dengan TP (*True Positive*) adalah jumlah data benar pada target yang terklasifikasi benar pada sistem, TN (*True Negative*) adalah jumlah data salah pada target yang terklasifikasi salah pada sistem, FP (*False Positive*) merupakan representasi jumlah data salah pada target yang terklasifikasi benar pada sistem dan FN (*False Negative*) merupakan representasi jumlah data benar pada target yang terklasifikasi salah pada sistem. Nilai-nilai tersebut akan tampil dalam bentuk *confusion matrix*.

## 2. Sensitivitas

Sensitivitas merupakan ukuran kemampuan sistem untuk melakukan prediksi terhadap data yang dianggap benar sesuai dengan TPR (*True Positive Rate*). Sensitivitas dapat dirumuskan dalam persamaan(24) berikut.

$$sensitivity = \frac{TP}{TP + FN} \times 100 \quad (24)$$

## 3. Spesifisitas

Spesifisitas berkebalikan dengan sensitivitas yaitu kemampuan sistem untuk melakukan prediksi terhadap data yang dianggap salah sesuai dengan TNR (*True Negative Rate*). Spesifisitas dapat dirumuskan dalam persamaan(25) berikut.

$$specificity = \frac{TN}{TN + FP} \times 100 \quad (25)$$

## 4. METODE PENELITIAN

### 4.1 Lokasi Penelitian

Penelitian ini bekerja sama dengan black java coffee roaster Yogyakarta yang merupakan tempat untuk melakukan pengambilan sampel green bean sekaligus sebagai tempat wawancara narasumber untuk klasifikasi varietas kopi arabika. Selanjutnya, penelitian dilakukan di STMIK AKAKOM Yogyakarta.

### 4.2 Teknik Pengumpulan dan Analisis Data

Data yang digunakan dalam penelitian ini merupakan citra green bean coffee yang difoto secara manual menggunakan kamera Canon 750D dengan jumlah data sebanyak 60 citra. Green bean coffee ini merupakan kopi arabika yang terdiri dari dua jenis yaitu mandailing dan toraja. Dari 60 citra tersebut, terdapat 30 citra jenis kopi mandailing dan 30 citra jenis kopi toraja. Citra tersebut merupakan citra RGB yang diambil dari dua sisi sehingga masing-masing biji kopi memiliki dua citra yang berbeda. Salah satu sisi dari jenis kopi cenderung hampir sama sehingga cukup sulit untuk dibedakan.

Sementara sisi yang lain secara visual dapat dibedakan berdasarkan bentuk permukaan dari biji kopi tersebut. Sementara itu, terdapat fitur lain yang sering digunakan sebagai acuan oleh petani ataupun pemilik coffee shop yaitu fitur bentuk. Bagi kaum awam, fitur ini tidak bisa terlihat dengan jelas, karena bentuknya cenderung hampir sama.

### 4.3 Tahap Penelitian

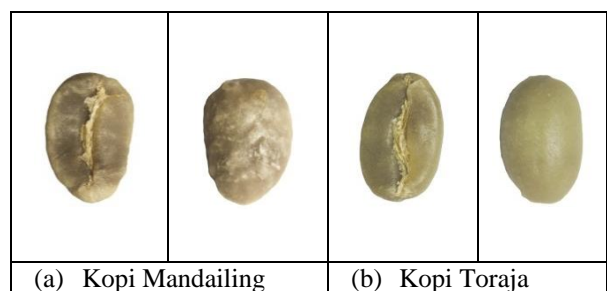


Gambar 3 Tahap Penelitian

Penelitian ini terdiri dari beberapa tahapan yang diawali dengan proses pengumpulan data citra, selanjutnya citra tersebut digunakan sebagai citra input dalam tahap pre-processing yaitu *cropping* citra yang bertujuan untuk menyiapkan citra agar dapat diolah. Proses ini hanya akan mengambil biji kopi dan mengabaikan *background* citra serta dilakukan secara manual karena mengikuti bentuk dan ukuran dari masing-masing biji kopi. Tahap selanjutnya adalah segmentasi citra yang bertujuan untuk mengambil bentuk biji kopi sebelum diekstraksi. Setelah segmentasi, tahap selanjutnya adalah ekstraksi fitur untuk memperoleh fitur bentuk. Tahap terakhir adalah proses klasifikasi menggunakan MLP. Proses klasifikasi dilakukan menggunakan WEKA 3.6 dengan test mode 10-fold cross-validation yang berarti data akan dibagi menjadi 10 bagian secara acak, kemudian dilakukan 10 kali eksperimen dimana masing-masing eksperimen menggunakan 10 data sebagai data uji dan sisanya sebagai data latih.

## 5. HASIL PENELITIAN

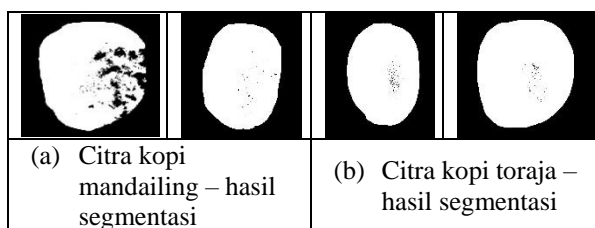
Berdasarkan penelitian-penelitian sebelumnya, metode yang sering digunakan adalah GLCM dan memberikan hasil yang akurat. Sementara penelitian ini mengusulkan metode *circularity* yang digunakan untuk melakukan ekstraksi fitur bentuk dan hasil yang diperoleh menunjukkan bahwa metode ini juga mampu memberikan hasil yang akurat. Hasil dari ekstraksi ini berupa angka-angka sesuai dengan fitur yang diukur. Gambar 4 berikut merupakan salah satu contoh citra kopi mandailing dan toraja.



Gambar 4 Contoh sampel dataset



Gambar 4(a) adalah varietas kopi mandailing yang diambil dari sisi depan maupun sisi belakang, selanjutnya hal yang sama juga dilakukan untuk Gambar 4(b) yang merupakan varietas kopi toraja. Sebelum diproses, citra tersebut *dicropping* secara manual dengan tujuan agar fokus penelitian hanya pada objeknya dan mempercepat proses komputasi karena ukuran citra menjadi lebih kecil. Sebelum diekstraksi, proses yang perlu dilakukan adalah mengambil objek yang terdapat dalam citra dan mengabaikan bagian yang bukan objek. Salah satu cara yang dapat dilakukan adalah melalui proses segmentasi. Gambar 5 berikut merupakan salah satu contoh hasil segmentasi citra menggunakan metode Otsu.



Gambar 5 Citra hasil segmentasi

Tahap selanjutnya adalah ekstraksi fitur menggunakan *circularity*. Tabel 1 berikut menunjukkan hasil ekstraksi fitur menggunakan *circularity*.

<i>Circularity</i>	Jenis	<i>Circularity</i>	Jenis
0.26	Sigarar Utang	0.8	Lini S-795
0.76	Sigarar Utang	0.83	Lini S-795
0.54	Sigarar Utang	0.83	Lini S-795
0.78	Sigarar Utang	0.79	Lini S-795
0.11	Sigarar Utang	0.78	Lini S-795
0.71	Sigarar Utang	0.76	Lini S-795
0.62	Sigarar Utang	0.79	Lini S-795
0.67	Sigarar Utang	0.8	Lini S-795
0.72	Sigarar Utang	0.85	Lini S-795
0.56	Sigarar Utang	0.82	Lini S-795
0.71	Sigarar Utang	0.78	Lini S-795
0.6	Sigarar Utang	0.7	Lini S-795
0.37	Sigarar Utang	0.8	Lini S-795
0.78	Sigarar Utang	0.83	Lini S-795

0.68	Sigarar Utang	0.82	Lini S-795
0.73	Sigarar Utang	0.82	Lini S-795
0.77	Sigarar Utang	0.5	Lini S-795
0.79	Sigarar Utang	0.79	Lini S-795
0.74	Sigarar Utang	0.78	Lini S-795
0.73	Sigarar Utang	0.79	Lini S-795
0.67	Sigarar Utang	0.78	Lini S-795
0.75	Sigarar Utang	0.75	Lini S-795
0.48	Sigarar Utang	0.87	Lini S-795
0.34	Sigarar Utang	0.81	Lini S-795
0.74	Sigarar Utang	0.79	Lini S-795
0.69	Sigarar Utang	0.82	Lini S-795
0.68	Sigarar Utang	0.8	Lini S-795
0.71	Sigarar Utang	0.75	Lini S-795
0.75	Sigarar Utang	0.72	Lini S-795
0.73	Sigarar Utang	0.72	Lini S-795

Hasil ekstraksi digunakan sebagai masukan dalam proses klasifikasi. Tabel 1 berikut merupakan *confusion matrix* dari hasil klasifikasi fitur GLCM menggunakan WEKA 3.6.

Tabel 1 Hasil klasifikasi dengan fitur GLCM

		<i>Target</i>	
		<i>Sigarar urang</i>	<i>Lini S-795</i>
<i>Prediksi</i>	<i>Sigarar utang</i>	25	5
	<i>Lini S-795</i>	7	23

Tabel 1 merupakan *confusion matrix* yang menggambarkan hasil klasifikasi menggunakan MLP dan diperoleh nilai *True Positive* (TP)=25, *True Negative*(TN)=23, *False Negative* (FN)=5 dan *False Positive* (FP)=7. Artinya dari 30 citra kopi mandailing, MLP mampu mengenali sebagai citra kopi sigarar utang sebanyak 25 citra sedangkan 5 citra dikenali sebagai citra kopi lini S-795. Selanjutnya dari 30 citra kopi lini S-795, MLP mampu mengenali sebagai citra kopi lini S-795 sebanyak 23 citra sedangkan 7 citra dikenali sebagai

citra kopi sigarar utang. Berdasarkan *confusion matrix* maka dapat dihitung tingkat akurasi, sensitivitas dan spesifisitas. Untuk pengenalan varietas kopi tersebut, diperoleh nilai akurasi sebesar 80% dengan sensitivitas sebesar 83,33% dan spesifisitas sebesar 76,7%. Berdasarkan hasil klasifikasi yang diperoleh, terdapat kesalahan pengenalan sebanyak 13 citra. Kesalahan terbanyak adalah ketika pengenalan varietas lini S-795. Hal ini mungkin disebabkan karena terdapat beberapa citra yang memiliki rentang nilai yang mirip ataupun sama seperti varietas sigarar utang, dan begitu juga sebaliknya.

## 6. KESIMPULAN

Proses pengenalan varietas kopi selama ini dilakukan secara manual dengan memanfaatkan pengalaman seorang petani. Ciri yang sering dijadikan acuan dalam mengenali varietas kopi ini adalah warna, ukuran dan bentuk maupun bentuk pohon. Namun demikian, petani dapat melakukan kesalahan ataupun pengenalan bisa menjadi tidak konsisten karena perbedaan pendapat antar petani. Oleh karena itu, penelitian ini mengusulkan teknik pengolahan citra yang terdiri dari proses segmentasi menggunakan metode Otsu dan ekstraksi fitur menggunakan *circularity*. Tahap terakhir adalah proses klasifikasi menggunakan MultiLayer Perceptron. Nilai akurasi yang diperoleh sebesar 80% dengan sensitivitas sebesar 83,33% dan spesifisitas sebesar 76,7%.

## 7. DAFTAR PUSTAKA

- APAZA, REL GUZMÁN, CHRISTIAN E. PORTUGAL-ZAMBRANO, JUAN C. GUTIÉRREZ-CÁCERES, & CÉSAR A. BELTRÁN-CASTAÑÓN. 2014. An Approach for Improve the Recognition of Defects in Coffee Beans Using Retinex Algorithms. *Proceedings of the 2014 Latin American Computing Conference, CLEI 2014*.
- CONDORI, RAYNER H.MONTES, JUAN H.CHUCTAYA HUMARI, CHRISTIAN E. PORTUGAL-ZAMBRANO, JUAN C. GUTIÉRREZ-CÁCERES, & CÉSAR A. BELTRÁN-CASTAÑÓN. 2014. Automatic Classification of Physical Defects in Green Coffee Beans Using CGLCM and SVM. *Proceedings of the 2014 Latin American Computing Conference, CLEI 2014*.
- FU, ZHAOXIA & YAN HAN. 2012. A Circle Detection Algorithm Based on Mathematical Morphology and Chain Code. Pp. 253–56 in *2012 International Conference on Computing, Measurement, Control and Sensor Network*. Retrieved (<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6245828>).
- KADIR, ABDUL & SUSANTO ADHI. 2012. Ekstraksi Fitur Tekstur. Pp. 575–640 in *Teori dan Aplikasi Pengolahan Citra*.
- NORIEGA, LEONARDO. 2005. Multilayer Perceptron Tutorial.
- NUGROHO, H. A., M. M. SEBATUBUN, & T. B. ADJI. 2017. Ground Glass Opacity Lesion Morphology Extraction in Primary Lung Cancer. *International Journal of Medical Engineering and Informatics* 9(4):398–411.
- DE OLIVEIRA, EMANUELLE MORAIS, DIMAS SAMID LEME, BRUNO HENRIQUE GROENNER BARBOSA, MIRIAN PEREIRA RODARTE, AND ROSEMARY GUALBERTO FONSECA ALVARENGA PEREIRA. 2016. A Computer Vision System for Coffee Beans Classification Based on Computational Intelligence Techniques. *Journal of Food Engineering* 171:22–27. Retrieved (<http://dx.doi.org/10.1016/j.jfoodeng.2015.10.009>).
- RADI, MUHAMMAD RIVAI, & MAURIDHI HERY PURNOMO. 2015. Combination of First and Second Order Statistical Features of Bulk Grain Image for Quality Grade Estimation of Green Coffee Bean. *ARPN Journal of Engineering and Applied Sciences* 10(18):8165–74.
- RATRI, IGNATIAN DHIAN ESTU KARISMA, HANUNG ADI NUGROHO, & TEGUH BHARATA ADJI. 2015. Pengenalan Keaslian Tanda Tangan Dengan Menggunakan Kombinasi Dynamic Time Warping (DTW) Dan Polar Fourier Transform. Universitas Gadjah Mada.
- SUSANTO, ADHI & ABDUL KADIR. 2012. Ekstraksi Fitur Bentuk Dan Kontur. Pp. 575–640 in *Pengolahan Citra*.