



Penggunaan Algoritma CART untuk Pemilihan Bingkai Kacamata dengan Penerapan Model Morfologi Indeks Wajah untuk Identifikasi Bentuk Wajah

Angga Ayu Retno Hapsari^a, Rachmat Gernowo^b, Catur Edi Widodo^c

^aMagister Sistem Informasi, Universitas Diponegoro, Semarang, Indonesia.

^bDepartemen Fisika, Universitas Diponegoro, Semarang, Indonesia.

Naskah Diterima : 5 November 2019; Diterima Publikasi : 27 Desember 2019

DOI : 10.21456/vol10iss1pp1-9

Abstract

The importance of using the face morphology model is to be able to identify the shape of a person's face. Face shape identification help to determine eyeglass frames. Face shape identification is done based on the morphological facial index by calculating face length and width. The large variety of frame shapes and sizes make it difficult for consumers to choose which one suits their face. The absence of a standard frame style guide between face types against the eyeglass frame complicates the selection of eyeglass frames. The application of the Zen principle (balance) in the selection of the right frame expected to be a consideration in choosing eyeglass frame. Various forms of eyeglass frames that look like a square, round and oval make the Zen principle difficult to apply, so machine learning is needed to be able to create eyeglass frames selection system. The decision tree CART algorithm is chosen as a method for selecting eyeglass frames. The purpose of this study is that the system can provide a reference frame of glasses that matches the shape of the face which the class has chosen by optical. The study uses 109 face data that have been selected by the optical, from 109 data divided into two parts, 100 training data, and 9 test data. The prediction system produces an accuracy value of 93% at max depth 6 for reading glasses and 91% for sunglasses. The implementation of the CART algorithm is proven to be able to predict the selection of eyeglass frames using morphological attributes of face index.

Keywords : Decision Tree; CART; Eyeglasses; Face Shape; Morphological Facial Index

Abstrak

Pentingnya penggunaan model morfologi wajah adalah untuk dapat mengidentifikasi bentuk wajah seseorang. Identifikasi bentuk wajah diperlukan untuk referensi pemilihan bingkai kacamata. Identifikasi bentuk wajah dilakukan berdasarkan model morfologi indeks wajah dengan menghitung panjang dan lebar wajah. Banyaknya macam bentuk dan ukuran bingkai membuat konsumen sulit untuk memilih mana yang cocok dengan bentuk wajah mereka. Tidak adanya panduan gaya bingkai yang baku untuk keserasian antara tipe wajah dan bingkai kacamata mempersulit pemilihan bingkai kacamata. Penerapan prinsip Zen (keseimbangan) dalam pemilihan bingkai yang tepat yaitu jenis bingkai seharusnya dapat menutupi kekurangan wajah sehingga tercapai keseimbangan dapat menjadi kata kunci untuk pemilihan bingkai kacamata. Beragam bentuk bingkai kacamata yang tidak hanya terlihat kotak, bulat dan oval membuat prinsip Zen susah diterapkan, maka dibutuhkan pembelajaran mesin untuk dapat membuat sistem pemilihan bingkai kacamata. *Decision Tree* Algoritma CART dipilih sebagai metode untuk pemilihan bingkai kacamata. Tujuan penelitian ini adalah sistem dapat memberikan satu referensi bingkai kacamata sesuai dengan bentuk wajah yang kelasnya sudah dipilihkan oleh optiker. Penelitian ini menggunakan 109 data wajah, dari 109 data dibagi menjadi dua bagian, 100 data latih dan 9 data uji. Sistem prediksi menghasilkan nilai akurasi sebesar 93% pada *max_depth* 6 untuk kacamata baca dan 91% untuk kacamata hitam. Implementasi Algoritma CART terbukti mampu memprediksi pemilihan bingkai kacamata menggunakan atribut morfologi indeks wajah.

Keywords : Pohon Keputusan; CART; Kacamata; Bentuk Wajah; Index Morfologi Wajah.

*) Penulis korespondensi: anggaayu@gmail.com

1. Pendahuluan

Pentingnya penggunaan model morfologi wajah adalah untuk dapat mengidentifikasi bentuk wajah seseorang. Identifikasi bentuk wajah diperlukan untuk referensi pemilihan bingkai kacamata. Identifikasi bentuk wajah dilakukan berdasarkan model morfologi indeks wajah dengan menghitung panjang dan lebar wajah. Ada ratusan model produk kacamata yang tersedia untuk dipilih, sulit dan memakan waktu untuk memilih model yang terlihat bagus secara estetika di wajah seseorang. Sistem uji coba virtual dapat mengatasi masalah ini (Yuan *et al.*, 2011).

Suatu perusahaan pasti membuat produk. Produk tersebut perlu dipasarkan dan diperkenalkan kepada konsumen. Perusahaan perlu mengetahui keinginan konsumen. Mengetahui keinginan pelanggan bukanlah hal yang sederhana, diperlukan ketelitian dan kesabaran karena berkaitan dengan teknik pengambilan data, akurasi data dan ketrampilan dalam menganalisa dari data-data yang sudah diperoleh. Strategi digital sangat efektif dan dapat memberikan impresi tersendiri. Dari data produk yang disukai oleh konsumen dapat diolah menjadi informasi yang dapat berguna untuk melihat tren produk ataupun untuk referensi stok barang. Setiap individu mengharapkan bantuan untuk penilaian kesesuaian kacamata, tidak hanya dengan ukuran, tetapi juga dengan aspek estetika bingkai (Clifford *et al.*, 2007). Banyaknya macam bentuk dan ukuran bingkai membuat konsumen sulit untuk memilih mana yang cocok dengan bentuk wajah mereka, bahkan banyak diantara mereka yang tidak tahu bentuk wajah mereka termasuk tipe wajah apa.

Decision Tree merupakan skema klasifikasi terstruktur pohon di mana node mewakili variabel input dan daun sesuai dengan hasil keputusan. *Decision Tree* adalah salah satu metode *Machine Learning* paling awal dan paling menonjol yang telah diterapkan secara luas untuk keperluan klasifikasi. Berdasarkan arsitektur *Decision Tree*, mudah diinterpretasikan dan "cepat" untuk dipelajari (Kourou *et al.*, 2015). Pohon keputusan diakui menjadi pendekatan pembelajaran mesin yang paling efektif dan efisien dan telah berhasil diterapkan untuk memecahkan masalah dunia nyata dalam bidang kecerdasan buatan (Trabelsi *et al.*, 2019). Penggunaan teknologi komputer untuk mengidentifikasi bentuk wajah yang digunakan sebagai dasar pemilihan bingkai kacamata membuat lingkungan interaksi manusia dan komputer yang lebih baik (Huang *et al.*, 2011). Meningkatkan daya tarik emosional yang dihasilkan dari penampilan produk akan mendorong penjualan kacamata (Chuan *et al.*, 2013). Gudang data perlu mendukung data historis yang akurat sehingga penambahan data dapat mengambil data yang bermanfaat untuk kepentingan tren mode (Berry dan Linoff, 2004).

Berdasarkan permasalahan tersebut diperlukan interaksi manusia dan komputer yang inovatif untuk mendapatkan produk fashion yang dalam hal ini kacamata mana yang cocok untuk pengguna.

Tujuan penelitian ini adalah untuk dapat mengidentifikasi bentuk wajah dari citra wajah yang berguna untuk referensi pemilihan bingkai kacamata. Identifikasi bentuk wajah dilakukan berdasarkan model morfologi indeks wajah dengan menghitung panjang dan lebar wajah. Pohon keputusan algoritma CART dipilih sebagai metode untuk pemilihan bingkai kacamata.

2. Kerangka Teori

Skala dapat digunakan untuk mengukur parameter morfologi (tinggi wajah, lebar wajah, dan indeks wajah). Tinggi wajah morfologis diukur dengan dengan skala dari nasion (n) hingga gnathion (gn). Lebar wajah diukur sebagai jarak lurus antara *zygion* kanan dan kiri ($zyzy$) (Yesmin *et al.*, 2014). Pendekatan berbasis fitur dan pendekatan berbasis model dapat digunakan untuk metode identifikasi wajah (Singh *et al.*, 2017). Untuk setiap individu indeks wajah dihitung menggunakan rumus seperti yang diberikan oleh Martin dan Saller dalam tampilan depan untuk menentukan jenis wajah (Mane *et al.*, 2010). Identifikasi wajah dapat dilakukan dengan *Max Margin Object Detection* (MMOD). Metode ini tidak melakukan sub-sampling, tetapi sebaliknya mengoptimalkan semua sub-jendela atau gambar (King, 2015). Algoritma CART digunakan sebagai dasar menentukan atribut terbaik untuk membuat pemisahan dalam simpul (Rutkowski *et al.*, 2014).

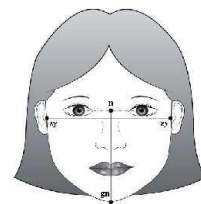
2.1. Morfologi Index Wajah

Untuk mengetahui bentuk (*shape*) diciptakan proporsi antara ukuran-ukuran yang dinamakan indeks (Glinka *et al.*, 2007).

Indeks wajah dihitung menggunakan rumus:

$$I = \frac{\text{Panjang wajah}}{\text{Lebar wajah}} \times 100 \quad (1)$$

Panjang wajah diambil dari jarak antara nasion (n) dan gnathion (gn) dan lebar wajah sebagai jarak antara titik *zygoma* kanan dan kiri. Gambar titik wajah dapat dilihat pada gambar 2.1 :



Gambar 2.1 Titik Wajah

Setelah menghitung indeks wajah, tipe wajah diklasifikasikan seperti yang ditunjukkan pada Tabel 1.




Tabel 1. Klasifikasi jenis wajah (Glinka *et al.*, 2007)

Tipe Wajah	Laki-laki	Perempuan
<i>Hypereuryprosop</i>	≤ 78,9	≤ 76,9
<i>Euryprosop</i>	79,0 – 83,9	77,0 – 80,9
<i>Mesoprosop</i>	84,0 – 87,9	81,0 – 84,9
<i>Leptoprosop</i>	88,0 – 92,9	85,0 – 89,9
<i>Hyperleptoprosop</i>	≥ 93,0	≥ 90,0

Bentuk wajah dalam indeks prosopic subjek *leptoprosop* dan *hyperleptoprosop* memiliki bentuk wajah oval, *mesoprosopic* untuk wajah kotak sedangkan individu dengan bentuk bulat dapat dianggap sebagai *euryprosop* atau *hypereuriprosop*, dengan dimensi horizontal tertinggi (Mane *et al.*, 2010).

Ringkasan kesesuaian antara bentuk wajah dan bingkai kacamata sesuai prinsip Zen dapat dilihat pada Tabel 2.

Tabel 2. Tabel kacamata berdasarkan bentuk wajah

No	Wajah	Ciri		Kacamata
		Tipe wajah	Bentuk Jawline/rahang	
1	Oval	<i>Leptoprosop</i> / <i>Hyperleptoprosop</i>		Semua tipe
2	Bundar (Round)	<i>Euryprosop</i> / <i>Hypereuriprosop</i>		Rectangle Square
3	Kotak (Square)	<i>Mesoprosopic</i>		Round Oval

2.2. Machine Learning

Machine learning merupakan cabang dari kecerdasan buatan yang menghubungkan masalah belajar dari sampel data dengan konsep umum inferensi. Ada dua jenis umum metode machine learning yang dikenal sebagai *supervised learning* dan *unsupervised learning*. Dalam pembelajaran *supervised*, data pelatihan yang digunakan untuk memperkirakan atau memetakan adalah input data ke output yang diinginkan. Sebaliknya, metode pembelajaran *unsupervised* tidak ada contoh berlabel yang disediakan dan tidak ada gagasan tentang output selama proses pembelajaran.

Data akan dibagi menjadi 2 kelompok, yaitu data *training* dan data *testing*. *Data training* akan digunakan untuk melatih algoritma untuk mencari model yang cocok, sementara data *testing* akan dipakai untuk menguji dan mengetahui performa model yang didapatkan pada tahapan *testing*.

Dari model yang didapatkan, dapat dilakukan prediksi yang dibedakan menjadi dua macam, tergantung tipe keluarannya. Jika hasil prediksi bersifat diskrit, maka dinamakan proses klasifikasi. Sedangkan jika kelurannya bersifat kontinyu, maka dinamakan proses regresi.

2.3. Decision Tree

Pohon keputusan adalah salah satu pendekatan klasifikasi yang paling populer dalam pembelajaran mesin. *Decision Tree* merupakan metode pembelajaran *supervised* non-parametrik yang digunakan untuk klasifikasi dan regresi. Tujuannya adalah untuk membuat model yang memprediksi nilai variabel target dengan mempelajari aturan keputusan sederhana yang disimpulkan dari fitur data. Pohon keputusan terdiri dari "root", "daun", dan *node* internal. *Node* internal menggunakan fitur tertentu untuk membagi ruang *instance* menjadi dua atau lebih subruang. Setiap daun mewakili satu kelas. Daun dapat mewakili nilai target yang paling tepat atau menunjukkan kemungkinan target memiliki nilai spesifik. Pohon keputusan belajar dari data untuk memperkirakan keputusan dengan seperangkat aturan keputusan if-then-else. Semakin dalam pohon, semakin kompleks aturan keputusan.

2.4. Algoritma CART

CART sangat mirip dengan C4.5, tetapi berbeda karena mendukung variabel target numerik (regresi) dan tidak menghitung set aturan. Langkah-langkah penerapan metode CART.

1. Pembentukan pohon

Proses pembentukan pohon klasifikasi terdiri atas 3 tahapan yaitu menghitung total kelas dan menentukan entropy. Formula dasar untuk *entropy* adalah :

$$Entropy(S) = \sum_{i=1}^k - p_i \log_2 p_i \quad (2)$$

Entropy (S) menggambarkan kemungkinan kasus, dan p_i adalah perbandingan ukuran dari kasus. Ini digunakan untuk meramalkan cabang yang akan membagi dataset menjadi beberapa bagian.

Tahapan kedua adalah menghitung gain dari masing-masing atribut/variable. Nilai *Gain* terbesar akan menjadi akar dari pohon. Formula untuk *Gain* adalah :

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{S_i}{S} \times Entropy(S_i) \quad (3)$$

Gain (S, A) menunjukkan seberapa pengaruh terhadap atribut A untuk kondisi S, dan (S_i) adalah angka kemungkinan S_i melawan S. Untuk menghitung SplitInfo masing-masing atribut digunakan perhitungan:

$$SplitInfo(S, A) = - \sum_{i=1}^n \frac{S_i}{S} \log_2 \frac{S_i}{S} \quad (4)$$

Pembagian data latih S menghasilkan *SplitInfo* untuk memperlihatkan potensial informasi *entropy*. itu ditunjukkan dalam variabel yang berbeda yang termasuk atribut A.

Bagian terakhir adalah menghitung *GainRatio* untuk masing-masing atribut dengan formula :

$$GainRatio(A) = \frac{Gain(S,A)}{SplitInfo(S,A)} \quad (5)$$

2. Pemangkasan Pohon Klasifikasi

Pemangkasan dilakukan dengan jalan memangkas bagian tree yang kurang penting sehingga didapatkan pohon optimal. Ukuran pemangkasan yang digunakan untuk memperoleh ukuran tree yang layak adalah cost complexity minimum.

3. Penentuan Pohon Kasifikasi Optimal

Pohon klasifikasi yang berukuran besar akan memberikan nilai penduga pengganti paling kecil, sehingga tree ini cenderung dipilih untuk mendukung nilai respon.

2.5. Confusion Matrix

Confusion Matrix digunakan untuk mengevaluasi kualitas output atau classifier pada set data iris. Elemen-elemen diagonal mewakili jumlah titik di mana label yang diprediksi sama dengan label yang sebenarnya, sedangkan elemen *off-diagonal* adalah mereka yang salah diberi label oleh *classifier*. Semakin tinggi nilai diagonal dari matriks akan semakin baik, hal ini menunjukkan banyak prediksi yang benar. Penggambaran Confusion Matrix dapat dilihat pada Gambar 2 :

	Predicted: NO	Predicted: YES
Actual: NO	TN	FP
Actual: YES	FN	TP

Gambar 2. Confusion Matrix

Ada dua kemungkinan kelas yang diprediksi: "ya" dan "tidak". Hasil dari pengamatan dan prediksi klasifikasi diistilahkan sebagai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). *True Positive* (TP) adalah jumlah data untuk pengamatan positif dan diprediksi positif, *False Negative* (FN) jumlah data untuk pengamatan positif, tetapi diprediksi negatif, *True Negative* (TN) jumlah data untuk pengamatan negatif dan diprediksi negatif, sedangkan *False Positive* (FP) adalah jumlah data untuk pengamatan negatif, tetapi diprediksi positif. Berdasarkan jumlah nilai tersebut dapat diperoleh nilai akurasi, recall dan presisi. Tingkat klasifikasi atau akurasi diberikan dalam persamaan :

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Recall dapat didefinisikan sebagai rasio dari jumlah total positif yang diklasifikasikan dengan benar dibagi dengan jumlah total positif. *High Recall* menunjukkan kelas dikenali dengan benar. Persamaan untuk menghitung recall adalah sebagai berikut :

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

Sedangkan untuk mendapatkan nilai presisi, dapat dilakukan dengan membagi jumlah total positif yang diklasifikasikan dengan benar dengan jumlah total positif yang diprediksi. Presisi Tinggi menunjukkan contoh berlabel positif memang positif. Persamaan untuk menghitung presisi adalah sebagai berikut ;

$$Presisi = \frac{TP}{TP+FP} \quad (8)$$

2.6. Fold Cross Validation

Cross-Validation merupakan salah satu teknik untuk menilai/memvalidasi keakuratan sebuah model yang dibangun berdasarkan dataset tertentu . Dalam teknik *Cross-Validation* ini dataset dibagi menjadi sejumlah K-buah partisi secara acak. Kemudian dilakukan sejumlah K-kali eksperimen, dimana masing-masing eksperimen menggunakan data partisi ke-K sebagai data testing dan memanfaatkan sisa partisi lainnya sebagai data training. Pada penelitian ini *K-fold cross-validation* menggunakan 10 *fold* yang artinya data dibagi menjadi 10 *fold* berukuran kira-kira sama, sehingga kita memiliki 10 subset data untuk mengevaluasi kinerja model atau algoritma. Dalam setiap subset, *cross-validation* memisahkan data set menjadi 2 bagian kemudian menggunakan satu bagian *fold* sebagai data uji dan 9 *fold* lainnya sebagai data latihan. Untuk mendapatkan nilai akurasi ataupun ukuran penilaian lainnya dari hasil eksperimen yang kita lakukan, dapat diambil nilai rata-rata dari seluruh eksperimen tersebut. Sebagai gambaran jika kita melakukan *10-Fold Cross-Validation* maka skema desain data eksperimennya adalah seperti Gambar 3:

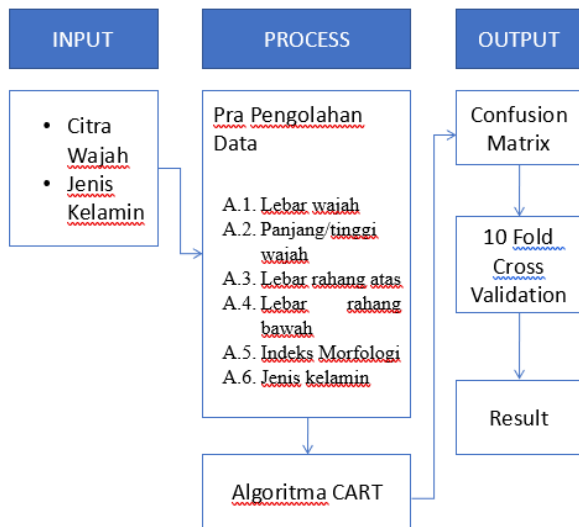
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10

□ Training Data □ Testing Data

Gambar 3. Ilustrasi 10 Fold Cross Validation

3. Metode

Data wajah pada penelitian ini menggunakan data citra natural. Citra natural adalah citra yang berasal dari hasil pengambilan gambar melalui perangkat elektronik misal dengan menggunakan kamera pada objek nyata yang ada di dalam kehidupan sehari-hari. Langkah-langkah penelitian dapat dilihat seperti pada Gambar 4 :



Gambar 4. Alur proses analisis

Penelitian ini menggunakan hasil survei dari optiker pada Optik Melawai untuk mengklasifikasi pemilihan bingkaiacamata. Klasifikasi yang akan diteliti yaitu pemilihan bingkaiacamata berdasarkan morfologi indeks wajah. Kacamata secara umum ada 2 macam yaitu kacamata hitam/sunglasses dan kacamata baca. Formulir survei yang diisikan oleh optiker berisi gambar citra wajah yang akan langsung dipilhkan kelas model bingkai kacamataanya.

Data penelitian diambil dari formulir survei dan data wajah dari citra gambar. Dari formulir survei didapatkan kelas untuk output pilihan bingkaiacamata. Tabel variabel penelitian ditunjukkan pada Tabel 3.

Tabel 3. Variabel Penelitian

Variabel/ atribut	Deskripsi
A.1 (Lebar wajah)	Nilai lebar wajah dari perhitungan pixel pada gambar citra (titik 0 sampai titik 16)
A.2 (tinggi wajah)	Nilai tinggi wajah dari perhitungan pixel pada gambar citra (titik (17,27) sampai dengan titik 8)
A.3 (Jarak ₂)	Nilai jarak titik kanan-kiri rahang atas dari perhitungan pixel pada gambar citra (titik 2 sampai titik 14)
A.4 (Jarak ₃)	Nilai jarak titik kanan-kiri rahang bawah dari perhitungan pixel pada gambar citra (titik 5 sampai dengan titik 11)
A.5 (Indeks)	Nilai indeks morfologi wajah
A.6 (Jenis Kelamin)	Laki-laki (L) atau perempuan (P)

Tabel kelas kacamata baca ditunjukkan pada Tabel 4 dan Tabel 5. dan tabel kelas kacamata hitam ditunjukkan pada Tabel 4.

Tabel 4. Data kelas bingkai kacamata baca

Seri	Kode Kelas	Kacamata	Seri
Baca-00	0		Bulat
Baca-01	1		Kotak
Baca-02	2		Bulat
Baca-03	3		Kotak
Baca-04	4		Oval
Baca-05	5		Kotak

Tabel 5. Data kelas bingkai kacamata hitam

Seri	Kode Kelas	Kacamata	Kategori Bentuk
Hitam-00	0		Bulat
Hitam-01	1		Bulat
Hitam-02	2		Kotak
Hitam-03	3		Bulat
Hitam-04	4		Bulat
Hitam-05	5		Kotak
Hitam-06	6		Kotak

4. Hasil dan Pembahasan

Dalam penelitian ini menggunakan data yang berasal dari hasil survei optiker pada Optik Melawai. Untuk pemilihan kacamata ini ada dua macam yaitu bingkai kacamata hitam dan bingkai kacamata baca. Data menggunakan 109 gambar foto wajah pegawai Universitas Diponegoro yang diprediksi pilihan bingkai kacamataanya oleh optiker Melawai dengan 7 pilihan kacamata hitam dan 6 pilihan kacamata baca yang bentuknya bermacam-macam. Atribut dari data ini adalah lebar wajah, panjang wajah, jarak₂ yang merupakan jarak titik kanan-kiri rahang atas, jarak₃

yang merupakan jarak titik kanan-kiri rahang bawah, indeks wajah dan jenis kelamin. Jarak_2 dan jarak_3 digunakan untuk mengukur keruncingan rahang, sedangkan indeks wajah di dapat dari menghitung panjang wajah dibagi lebar wajah dikalikan 100.

Deteksi wajah Max Margin dilakukan untuk mengidentifikasi wajah dari citra gambar dan menampilkan 68 titik wajah yang kemudian akan dihitung jarak yang dibutuhkan untuk mendapatkan nilai dari masing-masing atribut.

Lebar wajah diambil dari jarak titik 0 sampai titik 16. Panjang wajah diambil dari titik tangan alis yang diambil dari koordinat titik (17,27) sampai dengan titik 8. Jarak_2 diambil dari jarak titik 2 sampai titik 14 dan jarak_3 diambil dari jarak titik 5 sampai dengan titik 11. Indeks wajah dihitung dari rumus :

$$I = \frac{\text{Panjang wajah}}{\text{Lebar wajah}} \times 100$$

$$I = \frac{125.06}{138.01} \times 100$$

$$I = 93.49$$

Untuk menampilkan *landmarks* titik wajah digunakan Dlib. Dlib adalah sebuah *library* yang dapat mendeteksi 68 titik wajah, perintah coding untuk menampilkan 68 titik wajah pada gambar wajah :

```
import cv2
import dlib
import numpy as np
import imutils
from imutils import face_utils
from videocaptureasync import
VideoCaptureAsync
PREDICTOR_PATH =
"shape_predictor_68_face_landmarks.dat"
predictor =
dlib.shape_predictor(PREDICTOR_PATH)
detector =
dlib.get_frontal_face_detector()
```

sedangkan perintah untuk menghitung indeks wajah adalah sebagai berikut :

```
landmarks, rects = get_landmarks(image)
(bX, bY, bW, bH) =
face_utils.rect_to_bb(rects[0])
lebarWajah = np.linalg.norm(landmarks[0] -
landmarks[16])
jarak2 = np.linalg.norm(landmarks[2] -
landmarks[14])
jarak3 = np.linalg.norm(landmarks[5] -
landmarks[11])
tinggiWajah =
np.linalg.norm(np.matrix((landmarks[27,0],
landmarks[17,1])) - landmarks[8])
indeksMorfo = tinggiWajah / lebarWajah *
100
print("lebar: {:.2f}, tinggi: {:.2f},
indeks: {:.2f}".format(lebarWajah,
tinggiWajah, indeksMorfo))
print(landmarks.shape)
```

Dari citra wajah Gambar 4.1 didapatkan nilai atribut x1; x2; x3; x4; x5 berturut-turut adalah 138,01; 129,03; 130,03; 90,05; 93,49. Dari hasil perhitungan indeks wajah tersebut maka dapat

diidentifikasi wajah tersebut adalah oval seperti pada Gambar 5 :



Karakter Wajah: Leptoprosop (Oval)

Gambar 5. Hasil identifikasi wajah

4.1. Hasil aplikasi

Prediksi kacamata digunakan untuk memprediksi bingkai kacamata. Dalam aplikasi ini terdapat jendela foto untuk mengambil citra langsung ataupun unggah berkas citra, dan jendela untuk hasil prediksi kacamata. Kacamata hasil pilihan sistem langsung disematkan pada citra wajah agar pengguna langsung dapat mengetahui gambaran bentuk kacamata tanpa menggunakan produk aslinya dan fasilitas tombol suka/tidak suka agar pengguna dapat memberikan penilaian untuk kacamata yang dipilih oleh sistem sebagai referensi informasi dalam tren maupun stok produk untuk perusahaan. Contoh hasil prediksi kacamata baca dapat dilihat seperti pada Gambar 6 dan contoh hasil untuk kacamata hitam dapat dilihat pada Gambar 7 :



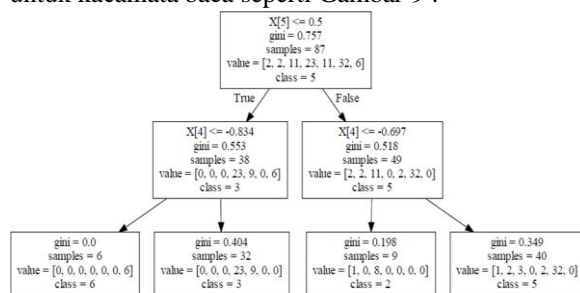
Gambar 6. Contoh hasil input dan output pada kacamata baca



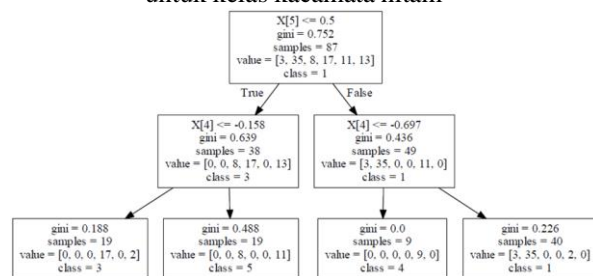
Gambar 7. Contoh hasil input dan output pada kacamata hitam

4.2. Pembahasan Algoritma CART

Salah satu parameter dalam sebuah pohon keputusan adalah kedalaman pohon yang dalam pemrograman disebut *Max depth*. *Max depth* digunakan untuk mengontrol kedalaman maksimum pohon yang dibuat. *Max depth* dapat digambarkan sebagai panjang jalan terpanjang dari akar pohon ke daun. Simpul *root* dianggap memiliki kedalaman 0. Untuk mendapatkan akurasi yang baik perlu dicoba beberapa nilai kedalaman pohon. Aturan nilai kedalaman pohon adalah minimal 2 dan maksimal adalah jumlah atribut yang dipakai. Dalam kasus ini atribut yang dipakai adalah 6 atribut maka nilai kedalaman pohon tidak boleh lebih dari 6. Hasil dari bentuk pohon untuk output kacamata hitam dengan kedalaman pohon 2 adalah seperti Gambar 8 dan untuk kacamata baca seperti Gambar 9 :

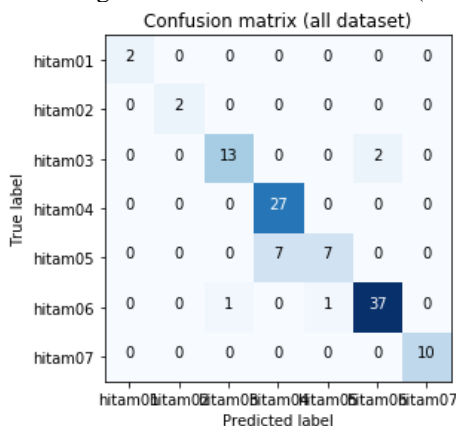


Gambar 8. Pohon keputusan dengan max_depth = 2 untuk kelas kacamata hitam



Gambar 9. Pohon keputusan dengan max_depth = 2 untuk kelas kacamata baca

Dari penelitian ini model pohon yang paling baik adalah yang menggunakan max_dept 6. Dari hasil percobaan untuk kacamata hitam, nilai akurasi terbaik diperoleh pada max_depth = 6 yaitu sebesar 90% dengan matrix Confusion Matrix (all dataset) :



Nilai akurasi pada max_depth = 6 diperoleh dari persamaan :

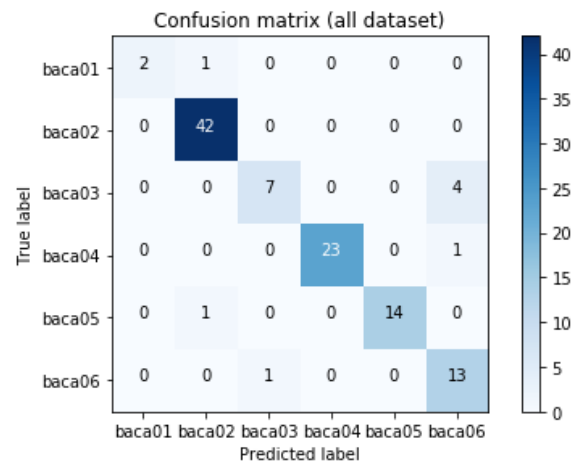
$$Akurasi = \frac{\sum_{i=1}^l TP_i + TN_i}{\sum_{i=1}^l TP_i + TN_i + FP_i + FN_i}$$

$$Akurasi = \frac{2 + 2 + 13 + 27 + 7 + 37 + 10}{2 + 2 + 13 + 27 + 7 + 37 + 10 + 2 + 7 + 1 + 1}$$

$$Akurasi = \frac{98}{109}$$

$$Akurasi = 0.90 = 90\%$$

Sedangkan untuk kacamata baca nilai akurasi terbaik diperoleh pada max_depth = 6 yaitu sebesar 93% dengan matrix Confusion Matrix (all dataset) :



Nilai akurasi pada max_depth = 6 diperoleh dari persamaan :

$$Akurasi = \frac{\sum_{i=1}^l TP_i + TN_i}{\sum_{i=1}^l TP_i + TN_i + FP_i + FN_i}$$

$$Akurasi = \frac{2 + 42 + 7 + 23 + 14 + 13}{2 + 42 + 7 + 23 + 14 + 13 + 1 + 4 + 1 + 1 + 1}$$

$$Akurasi = \frac{101}{109}$$

$$Akurasi = 0.93 = 93\%$$

4.3. Validasi hasil penelitian

Perbandingan Evaluasi antara hasil train test data, confusion matriks dan 10 fold cross validation dapat dilihat pada Tabel 6.

Tabel 6. Perbandingan hasil akurasi

Max_depth	Akurasi Output Kacamata Baca			Akurasi Output Kacamata Hitam		
	Train Test Data	CM	K-Fold	Train Test data	CM	K-Fold
2	86%	83%	83%	86%	81%	81%
3	73%	86%	86%	86%	82%	82%
4	73%	88%	87%	86%	85%	84%
5	64%	91%	90%	91%	86%	87%
6	64%	93%	93%	91%	90%	91%

Dari tabel 6 dapat dilihat bahwa hasil akurasi untuk kacamata baca dan kacamata hitam memang berbeda dikarenakan masing-masing mempunyai data yang berbeda dan kelas yang berbeda. Parameter max_depth 6 dipilih untuk kedua kelas

kacamata baik hitam dan baca karena pada pengujian CM dan *K-fold* mendapatkan nilai akurasi tertinggi, walaupun terlihat pada kelas kacamata baca saat train test data mempunyai nilai yang lebih rendah dari nilai *max_depth* lainnya. Dari hasil pengukuran kinerja sistem identifikasi bentuk wajah berdasarkan model morfologi indeks wajah untuk pemilihan bingkai kacamata menggunakan algoritma CART mampu memprediksi dengan sangat baik pada pemilihan kacamata baca dan kacamata hitam dengan hasil sebagai berikut: nilai akurasi prediksi untuk bingkai kacamata baca 93% dan 91% untuk kacamata hitam.

5. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, dapat diambil kesimpulan, bahwa penerapan model morfologi indeks wajah untuk identifikasi bentuk wajah dapat mengidentifikasi bentuk wajah oval, kotak dan bulat. Indeks wajah tersebut dapat digunakan sebagai atribut pemilihan bingkai kacamata menggunakan algoritma CART menghasilkan prediksi yang sangat baik yaitu dengan akurasi prediksi untuk output kacamata baca adalah 93%, dan output kacamata hitam adalah 91%. Berdasarkan hal tersebut, maka tujuan sistem yang dirancang dapat memberikan satu referensi bingkai kacamata terpenuhi dan dapat digunakan sebagai acuan untuk melakukan prediksi terkait pemilihan bingkai kacamata.

Daftar Pustaka

- Berry, M. J. and Linoff, G. S., 2004. Data Mining Techniques for Marketing, Sales, Customer Relationship Management, Second ed., Indianapolis, Indiana: Wiley Publishing, Inc.
- Clifford, W. B., and Irvin, M. B., 2007. *System for Ophthalmic Dispensing Third Edition*, Butterworth-Heinemann an imprint of Elsevier Inc, Philadelphia USA.
- Chuan, N. K., Sivaji, A., Shahimin, M. M., and Saad, N., 2013. Kansei Engineering for e-commerce Sunglasses Selection in Malaysia. *Procedia - Social and Behavioral Sciences*, 97, 707–714.
- Glinka, J., Myrtati, D.A., dan Toetik, K., 2008. *Metode Pengukuran Manusia*, Airlangga University Press, Surabaya.
- Huang, S.H., Yang, Y.I., and Chu, C.H., 2012. Human-centric design personalization of 3D glasses frame in markerless augmented reality. *Advanced Engineering Informatics*, 26(1), 35–45.
- King, D.E., 2015. *Max-margin object detection*, *CoRR*, arXiv:1502.00046v1
- Kourou, K., Exarchos, T.P., Exarchos, K.P., Karamouzis, M. V. and Fotiadis, D.I., 2015. Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8–17.
- Mane, D.R., Kale, A.D., Bhai, M.B. and Hallikerimath, S., 2010. Anthropometric and anthroposcopic analysis of different shapes of faces in group of Indian population: A pilot study. *Journal of Forensic and Legal Medicine*, 17(8), 421–425.
- Rutkowski, L., Jaworski, M., Pietruczuk, L., and Duda, P., 2014. The CART decision tree for mining data streams. *Information Sciences*, 266, 1–15.
- Singh, N., Daniel, A.K. and Chaturvedi, P., 2017. Template matching for detection & recognition of frontal view of human face through Matlab. *2017 International Conference on Information Communication and Embedded Systems, ICICES 2017*, (Icices).
- Trabelsi, A., Elouedi, Z. and Lefevre, E., 2019. Decision tree classifiers for evidential attribute values and class labels. *Fuzzy Sets and Systems*, 366, 46–62.
- Yesmin, T., Thwin, S.S., Afrin Urmi, S., Wai, M. M., Zaini, P. F., dan Azwan, K., 2014. A Study of facial index among Malay population. *Journal of Anthropology*, 2014, 1–4.
- Yuan, M., Khan, I. R., Farbiz, F., Niswar, A. and Huang, Z., 2011. A mixed reality system for virtual glasses try-on. *Proceedings of VRCAI 2011: ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications to Industry*, 363–366.