

## Klasifikasi Gender berbasis Wajah menggunakan Metode *Local Binary Pattern* dan *Random KNN*

Ruri Armandhani<sup>1</sup>, Randy Cahya Wihandika<sup>2</sup>, Muh. Arif Rahman<sup>3</sup>

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya  
Email: <sup>1</sup>ruriarmandhani@gmail.com, <sup>2</sup>rendicahya@ub.ac.id, <sup>3</sup>m\_arif@ub.ac.id

### Abstrak

Klasifikasi gender otomatis berbasis citra wajah merupakan salah satu topik penelitian yang menarik di dalam dunia visi komputer. Sistem klasifikasi gender otomatis memiliki peran penting dalam pengembangan aplikasi seperti sistem pengawasan otomatis dan sistem *monitoring*. Namun, komputer sulit untuk menemukan ciri khusus yang dapat membedakan gender seseorang sehingga dibutuhkan penerapan metode ekstraksi ciri terlebih dahulu. Selain itu, pemilihan metode klasifikasi juga berperan penting dalam keakuratan mengklasifikasikan gender. Tahap awal pada penelitian ini adalah melakukan deteksi wajah. Setelah itu, *pre-processing* dilakukan untuk mendapatkan citra bagian wajah yang ukurannya diubah menjadi berukuran 100x100 piksel. Kemudian, proses ekstraksi ciri tekstur dengan *Local Binary Pattern* (LBP) dilakukan pada citra hasil *pre-processing* tersebut. Citra tekstur yang dihasilkan oleh LBP dibagi menjadi beberapa bagian kecil yang disebut *region*. Lalu, nilai histogram 32-bin diambil dari setiap *region*. Setelah itu, seluruh histogram yang didapatkan dari beberapa *region* tersebut digabungkan menjadi satu vektor yang kemudian menjadi fitur histogram yang digunakan untuk mengklasifikasikan gender. Lalu, proses klasifikasi dilakukan dengan menggunakan metode *Random KNN*. Berdasarkan hasil pengujian pada penelitian ini, fitur terbaik yang dihasilkan dari proses ekstraksi ciri LBP dengan jumlah *region* 7x6. Rata-rata akurasi tertinggi yang dihasilkan pada fitur tersebut adalah 72,5% dengan menggunakan parameter nilai *k* dan nilai *r* yang paling optimal yaitu *k* = 11 dan *r* = 29.

**Kata kunci:** klasifikasi gender, ekstraksi ciri, *Local Binary Pattern*, *Random KNN*

### Abstract

*Automatic gender classification based on facial image is one of the interesting research topics in the world of computer vision. The automatic gender classification system has an important role in developing applications such as surveillance system and monitoring system. However, computers find it difficult to find a special characteristics that can distinguish someone's gender so that a feature extraction is needed. In addition, the selection of classification method is also important to get a better accuracy. The initial stage in this research is to do face detection. After that, pre-processing is done to get the face image only and the size of the image is normalized to 100x100 pixels. Then, the feature extraction process with Local Binary Pattern (LBP) method is done on the pre-processing image. Then, the texture image produced by LBP is divided into several small parts called region. The 32-bin histogram is extracted from each region. All of the histograms from each region are concatenated into a single vector which become the histogram feature used to classify gender. The classification was performed by Random KNN method. Based on the results of testing in this research, the best features produced from the LBP feature extraction which has 7x6 regions. The highest average accuracy produced by Random KNN is 72.5%. The optimal parameter value used for Random KNN in this research is *k* = 11 and *r* = 29.*

**Keywords:** gender classification, feature extraction, *Local Binary Pattern*, *Random KNN*

### 1. PENDAHULUAN

Saat ini, klasifikasi gender merupakan salah satu topik penelitian yang menarik di dalam

dunia visi komputer. Sistem klasifikasi gender otomatis memiliki peran penting dalam pengembangan aplikasi seperti sistem pengawasan otomatis dan sistem *monitoring*

(Tawhid & Dey, 2018), misalnya sistem untuk mengawasi suatu tempat yang hanya boleh dikunjungi oleh laki-laki atau perempuan saja. Selain itu, klasifikasi gender juga dapat bermanfaat di dalam dunia komersial. Salah satu contohnya adalah pada pengembangan suatu sistem yang dapat menghitung banyaknya laki-laki atau perempuan yang mengunjungi sebuah toko baju untuk keperluan evaluasi penjualan (Rani & Saepudin, 2013).

Dalam mengklasifikasikan gender, salah satu cara yang bisa dilakukan adalah dengan melihat wajah. Manusia dapat mengenali gender dengan akurasi mencapai lebih dari 90% hanya dengan melihat wajah seseorang meskipun tanpa gaya rambut dan terdapat perubahan lainnya pada wajah (Dey, Khan, & Ali, 2013). Hal tersebut dapat dilakukan dengan mudah oleh manusia tetapi sulit dilakukan oleh komputer. Komputer sulit untuk menemukan ciri khusus dari wajah yang dapat membedakan gender seseorang sehingga dibutuhkan penerapan metode ekstraksi ciri untuk menemukan karakteristik dari citra wajah dalam mengatasi kesulitan tersebut. Selain itu, pemilihan metode klasifikasi juga penting dalam keakuratan mengklasifikasikan gender.

Pada penelitian sebelumnya, Hatipoglu dan Köse (2016) telah berhasil melakukan penelitian terkait dengan pengenalan gender. Pada penelitian tersebut, metode *Local Binary Pattern* (LBP) digunakan untuk melakukan ekstraksi fitur tekstur pada data citra. Metode ini mempunyai kelebihan yaitu memiliki kemampuan diskriminatif dan proses komputasi yang sederhana. Penggunaan ekstraksi ciri tekstur LBP pada penelitian tersebut menghasilkan akurasi sebesar 88,13%.

Selain metode ekstraksi ciri, pemilihan metode klasifikasi juga tidak kalah penting dalam mengklasifikasikan gender. Pada penelitian sebelumnya yang telah dilakukan oleh TIN (2011), metode klasifikasi *K-Nearest Neighbor* (KNN) telah berhasil digunakan dalam melakukan estimasi gender dan umur seseorang. Metode ini diklaim memiliki hasil yang baik ketika diterapkan pada prototipe sistem pengenalan wajah secara *real-time*.

Selain KNN, terdapat beberapa metode klasifikasi lain yang juga dapat digunakan. Salah satunya adalah metode klasifikasi *Random KNN* yang pernah diterapkan pada penelitian yang dilakukan oleh Li dan Harner (2014) untuk mengklasifikasikan jenis kanker dengan tingkat akurasi tertinggi mencapai lebih dari 90%.

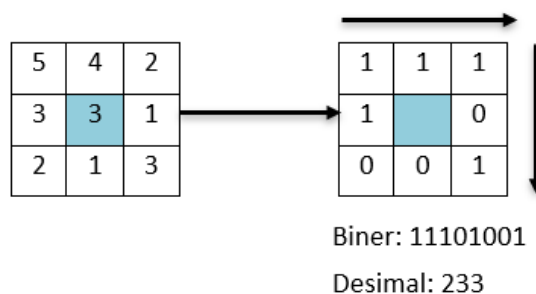
Metode *Random KNN* terbukti memiliki proses komputasi yang lebih cepat dibandingkan dengan KNN untuk penggunaan dataset yang memiliki dimensi tinggi. Metode *Random KNN* juga dapat diparalelkan pada mesin dengan beberapa unit komputasi.

Berdasarkan permasalahan yang telah dijelaskan pada paragraf sebelumnya, pada penelitian ini klasifikasi gender dilakukan dengan menggunakan metode LBP dan *Random KNN*. Pemilihan kedua metode ini diharapkan menghasilkan akurasi yang lebih baik daripada metode yang telah digunakan pada penelitian sebelumnya.

## 2. LANDASAN KEPUSTAKAAN

### 2.1. Local Binary Pattern

*Local Binary Pattern* (LBP) merupakan salah satu metode yang digunakan untuk melakukan ekstraksi fitur tekstur pada sebuah citra. Metode ini menggunakan bantuan operator matriks berukuran 3x3 yang berfungsi untuk mendapatkan nilai biner dari piksel tetangganya. Piksel tengah dari operator matriks tersebut digunakan sebagai *threshold*. Jika nilai piksel tetangga lebih besar atau sama dengan nilai piksel tengah maka piksel tetangga tersebut akan bernilai 1. Namun jika nilai piksel tetangga lebih kecil dari nilai piksel tengah maka piksel tetangga tersebut akan bernilai 0 (Hatipoglu & Köse, 2016). Urutan nilai biner didapatkan dengan cara melingkar seperti yang ditunjukkan pada Gambar 1.



Gambar 1. Operator LBP

Formulasi matematika dari operator LBP dapat dituliskan pada Persamaan (1) dan Persamaan (2).

$$LBP(x) = \sum_{i=1}^8 s(G(X_i) - G(X))2^{i-1} \quad (1)$$

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

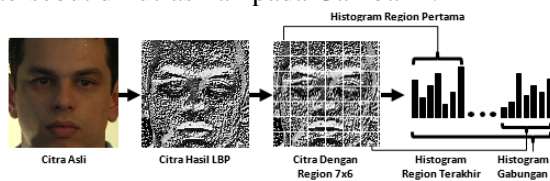
Keterangan :

$LBP(x)$  : nilai piksel baru yang dihasilkan oleh

operator LBP.

- $G(X_i)$  : nilai piksel tetangga ke-i.
- $G(X)$  : nilai piksel tengah.
- $i$  : urutan piksel tetangga.
- $s(t)$  : nilai biner.
- $t$  : selisih nilai piksel tengah dan piksel tetangga.

Kemudian citra hasil LBP operator tersebut dibagi menjadi beberapa *region*. Pada setiap *region* dilakukan proses ekstraksi histogram. Selanjutnya, histogram yang dihasilkan pada tiap-tiap *region* tersebut digabungkan menjadi satu (Shan, 2012). Ilustrasi langkah-langkah tersebut diilustrasikan pada Gambar 2.



Gambar 2. Ilustrasi Ekstraksi Histogram Pada Tiap Region

### 2.2. K-Nearest Neighbor

*K-Nearest Neighbor* (KNN) merupakan salah satu metode klasifikasi yang data ujinya diklasifikasikan berdasarkan jarak terhadap data latih yang menjadi tetangga terdekat data uji tersebut. Perhitungan jarak antara data uji dan data latih dilakukan menggunakan jarak *Euclidean*. Langkah-langkah yang perlu dilakukan pada metode KNN adalah sebagai berikut (Liantoni, 2015):

1. Menentukan jumlah tetangga sebanyak  $k$ .
2. Menghitung jarak antara data uji dan data latih menggunakan *Euclidean Distance*.
3. Urutkan jarak yang didapatkan berdasarkan nilai yang terkecil.
4. Tentukan kelompok data berdasarkan kelas mayoritas pada nilai  $k$ .

Secara umum, perhitungan jarak *Euclidean* dapat ditulis pada Persamaan (3) (Hastuti, Widodo, & Dewi, 2018).

$$d_i(P, Q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \tag{3}$$

Keterangan :

$d_i(P, Q)$  : nilai jarak *Euclidean* dari data latih P ke data uji Q.

- $p_i$  : nilai data latih pada fitur ke-i.
- $q_i$  : nilai data uji pada fitur ke-i.
- $n$  : jumlah fitur.

### 2.3. Random KNN

*Random KNN* merupakan metode klasifikasi yang mengadopsi teknik klasifikasi *Random Forest* dan *K-Nearest Neighbor* (KNN). *Random KNN* menggunakan KNN sebagai pengklasifikasi dasar yang mudah diimplementasikan dan juga stabil (Li & Harner, 2014). Pada metode ini pengklasifikasi KNN digunakan beberapa kali dengan menggunakan jumlah fitur yang berbeda-beda. Jumlah fitur yang digunakan pada setiap KNN didapatkan secara acak. Kemudian hasil klasifikasi setiap KNN tersebut digabungkan dan dilakukan *voting*. Secara umum, langkah-langkah yang dilakukan dalam metode *Random KNN* adalah sebagai berikut:

1. Menentukan banyaknya KNN yang digunakan yaitu sebanyak  $r$  KNN.
2. Menentukan jumlah fitur sebanyak  $m$  yang digunakan untuk setiap KNN. Nilai  $m$  dapat ditentukan dengan Persamaan (4).

$$m = \sqrt{p} \tag{4}$$

3. Mengambil fitur sebanyak  $m$  pada koleksi fitur keseluruhan secara acak sehingga didapatkan koleksi fitur baru  $F^{(m)}$ . Secara matematis dapat dituliskan pada Persamaan (5) dan Persamaan (6).

$$F = \{f_1, f_2, \dots, f_p\} \tag{5}$$

$$F^{(m)} = \{f_{j1}, f_{j2}, \dots, f_{jm} | f_{jl} \in F, 1 \leq l \leq m\} \tag{6}$$

Keterangan :

- $F$  : koleksi seluruh fitur yang digunakan.
- $p$  : jumlah seluruh fitur.

$F^{(m)}$  : koleksi fitur acak sebanyak  $m$ .

$m$  : jumlah fitur acak pada tiap KNN.

$f_{jl}$  : nilai fitur acak yang diambil dari

koleksi seluruh fitur.

4. Menentukan nilai  $k$  dan melakukan pengklasifikasian KNN biasa sebanyak  $r$  kali menggunakan koleksi fitur  $F^{(m)}$ .
5. Menggabungkan seluruh hasil klasifikasi dari setiap KNN.
6. Melakukan *voting* terhadap hasil klasifikasi yang telah digabungkan.

### 2.4. Confusion Matrix

*Confusion matrix* merupakan metode evaluasi yang sering digunakan untuk masalah klasifikasi. Pada metode evaluasi ini pengukuran kinerja dilakukan dengan menghitung akurasi. Selain itu, terdapat dua pengukuran tambahan

yaitu dengan menghitung *sensitivity* dan *specificity*. Kedua pengukuran tersebut dibutuhkan pada kasus data dengan kelas tidak seimbang sehingga pengukuran kinerja dari dari suatu metode klasifikasi menjadi lebih detail. Pada penelitian ini, pengukuran *sensitivity* menunjukkan seberapa baik model yang digunakan dalam memprediksi citra wajah seseorang sebagai laki-laki. Sedangkan pengukuran *specificity* menunjukkan seberapa baik model yang digunakan dalam memprediksi citra wajah seseorang sebagai perempuan. Model confusion matrix pada penelitian ini ditunjukkan pada Tabel 1.

Tabel 1. Model Confusion Matrix Dua Kelas

Kelas Asli	Kelas Hasil Klasifikasi	
	Laki-laki	Perempuan
Laki-laki	TP	FN
Perempuan	FP	TN

Pengukuran akurasi, *sensitivity* dan *specificity* masing-masing dilakukan dengan menggunakan Persamaan (7), Persamaan (8) dan Persamaan (9) (Handayani, Jamal, & Septiandri, 2017).

$$Akurasi = \frac{TP+TN}{TP+FP+FN+TN} \tag{7}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{8}$$

$$Specificity = \frac{TN}{FP+TN} \tag{9}$$

Keterangan :

- TP : Data dengan kelas asli laki-laki dan kelas hasil klasifikasinya laki-laki.
- FN : Data dengan kelas asli laki-laki dan kelas hasil klasifikasinya perempuan.
- TN : Data dengan kelas asli perempuan dan kelas hasil klasifikasinya perempuan.
- FP : Data dengan kelas asli perempuan dan kelas hasil klasifikasinya laki-laki.

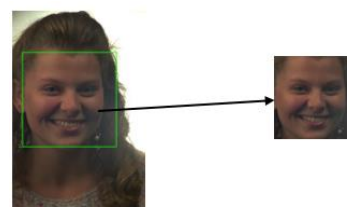
### 3. METODOLOGI

#### 3.1. Teknik Pengumpulan Data

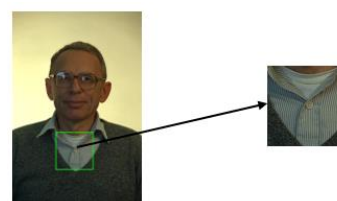
Data yang digunakan pada penelitian ini adalah data sekunder yang bersumber dari basis data Facial Recognition Technology (FERET). Data ini memiliki dua jenis kondisi yaitu data wajah tanpa aksesoris dan data wajah dengan aksesoris. Aksesoris yang dimaksudkan adalah kacamata. Data wajah tanpa aksesoris berjumlah 500 (250 perempuan dan 250 laki-laki). Sedangkan data wajah dengan aksesoris berjumlah 133 (29 perempuan dan 104 laki-laki).

Sebelum data digunakan untuk klasifikasi,

data terlebih dahulu dilakukan proses *cleaning* data. Proses *cleaning* data adalah proses menghapus data-data yang buruk. Data yang buruk adalah data wajah yang tidak terdeteksi sebagai wajah saat dilakukan proses deteksi wajah. Contoh data wajah yang baik dan contoh data wajah yang buruk dapat dilihat pada Gambar 3 dan Gambar 4.



Gambar 3. Data Wajah yang Baik



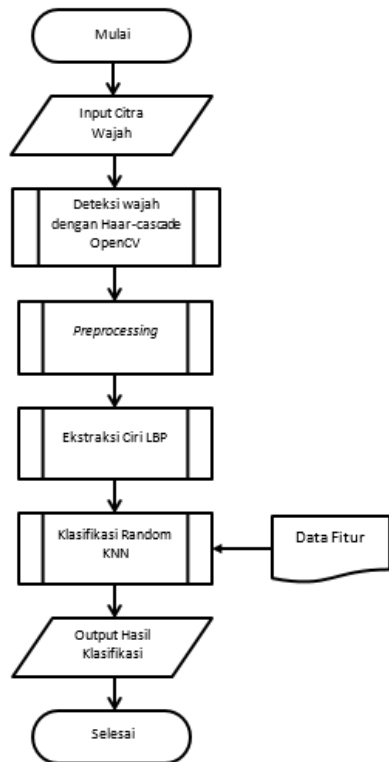
Gambar 4. Data Wajah yang Buruk

#### 3.2. Diagram Alir Sistem

Proses awal yang dilakukan adalah memasukkan data citra wajah yang akan dilakukan ekstraksi ciri tekstur. Sebelum dilakukan ekstraksi ciri tekstur, citra wajah terlebih dahulu dilakukan proses deteksi wajah untuk mengetahui posisi wajah pada citra tersebut. Deteksi wajah dilakukan dengan menggunakan library *Haar-cascade Detection* OpenCV. Setelah itu, citra yang sudah terdeteksi wajahnya akan dilakukan *pre-processing*. Pada tahap *pre-processing*, proses awal yang dilakukan yaitu melakukan *cropping* untuk mendapatkan bagian wajahnya saja. Kemudian citra hasil *cropping* tersebut diubah ukurannya menjadi 100x100 piksel. Lalu citra wajah yang masih dalam bentuk RGB dikonversi menjadi *grayscale*. Selanjutnya, citra wajah *grayscale* akan diperbaiki kontrasnya dengan menerapkan proses ekualisasi histogram. Setelah *pre-processing*, tahap selanjutnya adalah melakukan ekstraksi ciri tekstur menggunakan metode LBP. Pada tahap ini, operator LBP yang digunakan adalah operator matriks berukuran 3x3. Citra hasil proses LBP akan dibagi menjadi beberapa bagian. Setelah itu, nilai histogram pada setiap bagian akan diambil dan digabungkan menjadi satu sehingga menjadi fitur akhir yang



digunakan untuk proses klasifikasi. Proses klasifikasi pada penelitian ini menggunakan metode *Random KNN*. Alur keseluruhan proses yang dilakukan sistem pada penelitian ini ditunjukkan pada Gambar 5.



Gambar 5. Diagram Alir Sistem

#### 4. PENGUJIAN DAN ANALISIS

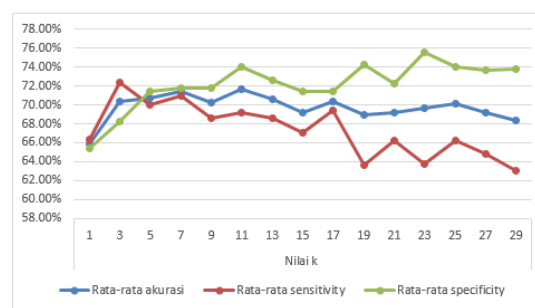
##### 4.1. Pengujian Nilai *k*

Pengujian ini dilakukan untuk mengetahui pengaruh dari nilai *k* terhadap hasil akurasi, *sensitivity* dan *specificity*. Pengujian ini juga berguna untuk mencari nilai *k* yang paling optimal untuk digunakan pada pengujian selanjutnya. Nilai parameter lainnya yang digunakan adalah *r* = 11 dan jumlah *region* 7x6. Pengujian ini diujikan pada data wajah tanpa aksesoris sebanyak sepuluh kali percobaan. Hasil pengujian dari masing-masing nilai *k* dapat dilihat pada Tabel 2 dan Gambar 6.

Tabel 2. Hasil Pengujian Nilai *k*

Nilai <i>k</i>	Rata-rata akurasi	Rata-rata <i>sensitivity</i>	Rata-rata <i>specificity</i>
1	65,9%	66,4%	65,4%
3	70,3%	72,4%	68,2%
5	70,7%	70%	71,4%
7	71,4%	71%	71,8%

9	70,2%	68,6%	71,8%
11	71,6%	69,2%	74%
13	70,6%	68,6%	72,6%
15	69,2%	67%	71,4%
17	70,4%	69,4%	71,4%
19	68,9%	63,6%	74,2%
21	69,2%	66,2%	72,2%
23	69,7%	59,2%	69,2%
25	69,7%	63,8%	75,6%
27	70,1%	66,2%	74%
29	69,2%	64,8%	73,6%



Gambar 6. Grafik Hasil Pengujian Nilai *k*

Berdasarkan pada Tabel 2 hasil rata-rata akurasi tertinggi dihasilkan pada *k* = 11 yaitu sebesar 71,6%. Sedangkan rata-rata hasil akurasi terendah dihasilkan pada *k* = 1 sebesar 65,9%. Kemudian, pada Gambar 6 menunjukkan bahwa nilai *k* berpengaruh terhadap rata-rata hasil akurasi. Rata-rata hasil akurasi cenderung menurun ketika nilai *k* < 11 dan *k* > 11. Selain itu, pada Gambar 6 juga dapat dilihat bahwa nilai *k* yang semakin besar memiliki rata-rata *sensitivity* yang cenderung menurun. Hal tersebut menunjukkan bahwa ketika nilai *k* semakin besar ketepatan sistem dalam mengklasifikasikan data gender laki-laki semakin buruk. Namun sebaliknya pada *specificity*, nilai *k* yang semakin besar memiliki rata-rata *specificity* yang semakin tinggi. Hal tersebut menunjukkan bahwa ketika nilai *k* semakin besar ketepatan sistem dalam mengklasifikasikan data gender perempuan semakin baik.

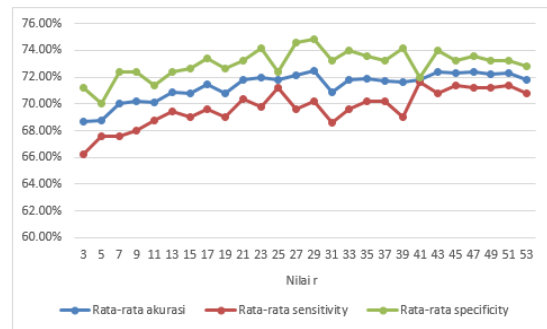
##### 4.2. Pengujian Nilai *r*

Pengujian ini dilakukan untuk mengetahui pengaruh dari nilai *r* terhadap hasil akurasi, *sensitivity* dan *specificity*. Pengujian ini juga digunakan untuk mencari nilai *r* yang paling optimal. Pemilihan nilai *r* yang merupakan

bilangan ganjil berfungsi untuk menghindari kebingungan dalam penentuan kelas saat terjadi kesamaan jumlah *voting* pada kelas laki-laki dan perempuan. Sedangkan nilai parameter lainnya yang digunakan pada pengujian ini adalah nilai *k* dengan rata-rata akurasi terbaik yaitu  $k = 11$  dan jumlah *region* 7x6. Pada pengujian ini data yang digunakan adalah data wajah tanpa aksesoris. Pengujian ini dilakukan sebanyak sepuluh kali percobaan untuk setiap nilai *r*. Hasil pengujian dari masing-masing nilai *r* ditunjukkan pada Tabel 3 dan Gambar 7.

Tabel 3. Hasil Pengujian Nilai *r*

Nilai <i>r</i>	Rata-rata akurasi	Rata-rata <i>sensitivity</i>	Rata-rata <i>specificity</i>
3	68,7%	66,2%	71,2%
5	68,8%	67,6%	70%
7	70%	67,6%	72,4%
9	70,2%	68%	72,4%
11	70,1%	68,8%	71,4%
13	70,9%	69,4%	72,4%
15	70,8%	69%	72,6%
17	71,5%	69,6%	73,4%
19	70,8%	69%	72,6%
21	71,8%	70,4%	73,2%
23	72,%	69,8%	74,2%
25	71,8%	71,2%	72,4%
27	72,1%	69,6%	74,6%
29	72,5%	70,2%	74,8%
31	70,9%	68,6%	73,2%
33	71,8%	69,6%	74%
35	71,9%	70,2%	73,6%
37	71,7%	70,2%	73,2%
39	71,6%	69%	74,2%
41	71,8%	71,6%	72%
43	72,4%	70,8%	74%
45	72,3%	71,4%	73,2%
47	72,4%	71,2%	73,6%
49	72,2%	71,2%	73,2%
51	72,3%	71,4%	73,2%
53	71,8%	70,8%	72,8%



Gambar 7. Grafik Hasil Pengujian Nilai *r*

Berdasarkan pada Tabel 3 rata-rata akurasi terbaik dihasilkan pada  $r = 29$  yaitu sebesar 72,50%. Kemudian, pada Gambar 7 menunjukkan bahwa nilai *r* dapat mempengaruhi rata-rata hasil akurasi. Ketika nilai  $r < 29$  rata-rata hasil akurasi yang dihasilkan cenderung menurun. Namun ketika  $r > 29$  rata-rata hasil akurasi mulai konvergen yaitu sebesar  $\pm 72\%$ . Tidak hanya itu, nilai *r* juga berpengaruh terhadap *sensitivity* dan *specificity* seperti yang ditunjukkan pada Gambar 7. Nilai *r* yang semakin besar membuat nilai *sensitivity* konvergen pada  $\pm 71\%$  dan nilai *specificity* konvergen pada  $\pm 73\%$ . Beberapa hal tersebut dapat terjadi karena semakin besar nilai *r* maka semakin banyak fitur serupa yang digunakan pada proses klasifikasi KNN di dalam random KNN, sehingga hasil *voting* pada proses akhir random KNN semakin baik.

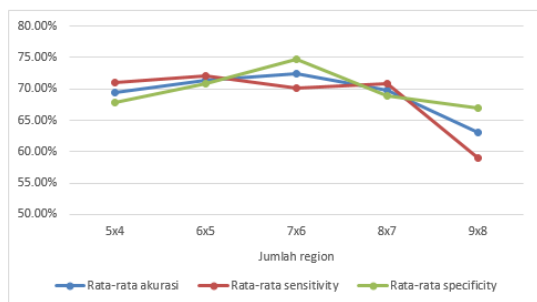
### 4.3. Pengujian Jumlah *Region*

Pengujian ini dilakukan untuk mengetahui pengaruh jumlah *region* terhadap hasil akurasi, *sensitivity* dan *specificity*. Jumlah *region* yang digunakan untuk pengujian adalah 5x4, 6x5, 7x6 8x7, dan 9x8. Nilai parameter *k* dan *r* yang digunakan dalam pengujian ini adalah yang memiliki akurasi terbaik yaitu  $k = 11$  dan  $r = 29$ . Pengujian ini juga diujikan pada data wajah tanpa aksesoris sebanyak sepuluh kali percobaan. Hasil pengujian dari masing-masing jumlah *region* ditunjukkan pada Tabel 4 dan Gambar 8.

Tabel 4. Hasil Pengujian Jumlah *Region*

Jumlah <i>Region</i>	Panjang Fitur	Rata-rata akurasi	Rata-rata <i>sensitivity</i>	Rata-rata <i>specificity</i>
5x4	640	69,4%	71%	67,8%
6x5	960	71,4%	72%	70,8%
7x6	1344	72,5%	70,2%	74,8%

8x7	1792	69,8%	70,8%	68,8%
9x8	2304	63%	59%	67%



Gambar 8. Grafik Hasil Pengujian Jumlah Region

Berdasarkan hasil pengujian yang ditunjukkan pada Tabel 4 nilai rata-rata akurasi tertinggi dihasilkan pada jumlah region 7x6 yaitu sebesar 72,5%. Sedangkan rata-rata hasil akurasi terendah dihasilkan pada jumlah region 9x8 yaitu sebesar 63%. Grafik pada Gambar 8 menunjukkan bahwa banyaknya jumlah region berpengaruh terhadap rata-rata hasil akurasi. Semakin banyak jumlah region yang digunakan maka rata-rata akurasi yang dihasilkan akan semakin rendah. Hal tersebut juga terjadi pada tingkat sensitivity dan specificity. Penurunan tingkat akurasi tersebut terjadi karena saat jumlah region semakin banyak maka fitur yang dihasilkan juga menjadi lebih banyak tetapi representasi ciri dari suatu citra tersebut menjadi kurang baik. Contoh fitur histogram yang kurang baik yang dihasilkan oleh LBP dengan region 9x8 ditunjukkan pada Tabel 5.

Tabel 5. Contoh Fitur Histogram yang Kurang Baik pada Region 9x8

No. Citra	Kel- as	Fitur Histogram
00224_940128	F	82 0 0 0 0 0 0 0 0 2 ...
00237_940128	M	75 0 0 0 0 0 0 0 1 0 ...
00302_940128	F	88 0 0 0 0 0 0 0 0 1 ...

Berdasarkan Tabel 5 dapat dilihat bahwa ketika region semakin besar maka nilai fitur histogram menjadi tidak merata dan banyak nilai histogram yang bernilai 0. Jika fitur yang bernilai 0 tersebut banyak terambil dan digunakan pada proses klasifikasi KNN yang ada di dalam Random KNN maka laki-laki dan perempuan akan sulit dibedakan karena

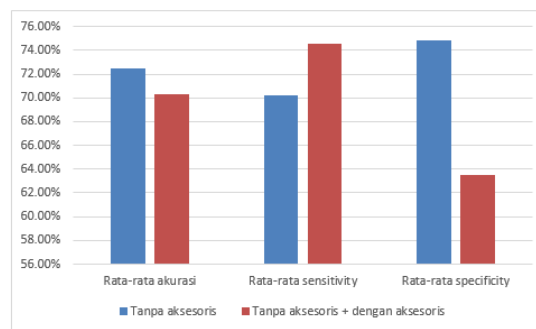
perbedaan pada jarak Euclidean hanya sedikit atau bahkan dapat menjadi sama persis. Hal tersebut juga kemudian berpengaruh terhadap hasil akhir voting yang kemudian menjadi semakin kurang baik.

#### 4.4. Pengujian Kondisi Data Wajah

Pengujian ini dilakukan untuk mengetahui pengaruh kondisi data wajah terhadap hasil akurasi. Kondisi data wajah yang dimaksud adalah data wajah tidak menggunakan aksesoris dan data wajah menggunakan aksesoris. Aksesoris yang digunakan pada data wajah tersebut adalah kacamata. Pengujian ini diujikan dua kali pada jumlah data yang berbeda yaitu 500 data tanpa aksesoris (250 laki-laki dan 250 perempuan) dan 633 data campuran antara tanpa aksesoris dan dengan aksesoris (354 laki dan 279 perempuan) dengan sepuluh kali percobaan pada masing-masing jumlah data. Pengujian ini dilakukan dengan menggunakan nilai k, nilai r dan jumlah region dengan akurasi terbaik yaitu k = 11, r = 29, dan jumlah region 7x6. Hasil dari pengujian ini dapat dilihat pada Tabel 6 dan Gambar 9.

Tabel 6. Hasil Pengujian Kondisi Data Wajah

Kondisi Data	Rata-rata akurasi	Rata-rata sensitivity	Rata-rata specificity
Tanpa Aksesoris	72,5%	70,2%	74,8%
Tanpa aksesoris + dengan aksesoris	70,28%	74,56%	63,49%



Gambar 9. Grafik Hasil Pengujian Kondisi Data

Berdasarkan pada Tabel 6 dan Gambar 9 dapat dilihat bahwa penambahan data wajah yang menggunakan aksesoris dapat mempengaruhi hasil klasifikasi. Hal tersebut bisa dari rata-rata hasil akurasi yang dihasilkan menjadi turun sekitar ±2% dari 72,5% menjadi 70,28%. Selain itu, apabila dilihat dari rata-rata sensitivity dan specificity, pengklasifikasian

gender laki-laki lebih baik dibandingkan dengan gender perempuan. Indikasi tersebut dapat dilihat berdasarkan kenaikan rata-rata *sensitivity* sebesar  $\pm 4\%$  dan penurunan rata-rata *specificity* yang cukup signifikan yaitu sebesar  $\pm 11\%$ .

## 5. KESIMPULAN

Berdasarkan hasil dari pengujian yang telah dilakukan pada penelitian ini, dapat disimpulkan bahwa fitur terbaik yang dihasilkan oleh ekstraksi ciri LBP adalah yang menggunakan jumlah *region*  $7 \times 6$  dan 32-bin histogram. Panjang fitur histogram yang dihasilkan adalah sebesar 1344. Semakin besar jumlah *region* maka panjang fitur akan semakin bertambah dan nilai fitur histogram semakin tidak merata serta banyak menghasilkan fitur histogram yang bernilai 0. Hal tersebut mengakibatkan hasil *voting* dari Random KNN menjadi kurang baik. Di samping itu, hasil klasifikasi dari metode *Random KNN* dengan menggunakan fitur tersebut dapat menghasilkan rata-rata akurasi tertinggi sebesar 72,5% dengan ketentuan parameter nilai  $k = 11$  dan  $r = 29$ . Selain itu, penggunaan aksesoris pada wajah dapat membuat rata-rata akurasi menjadi menurun sekitar  $\pm 2\%$ .

Adapun saran yang didapatkan pada penelitian ini untuk penelitian selanjutnya yaitu mengembangkan suatu metode yang dapat mengatasi masalah penggunaan aksesoris pada wajah.

## 6. DAFTAR PUSTAKA

- Dey, E. K., Khan, M., & Ali, M. H. (2013). Computer Vision-Based Gender Detection from Facial image. *International Journal of Advanced Computer Science*, 3(8), 428-433.
- Handayani, A., Jamal, A., & Septiandri, A. A. (2017). Evaluasi Tiga Jenis Algoritme Berbasis Pembelajaran Mesin untuk Klasifikasi Jenis Tumor Payudara. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 394-403.
- Hastuti, M. T., Widodo, A. W., & Dewi, C. (2018). Identifikasi Kondisi Kesehatan Ayam Petelur Berdasarkan Ciri Warna HSV Dan Gray Level Cooccurrence Matrix (GLCM) Pada Citra Jengger Dengan Klasifikasi K-Nearest Neighbour. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1054-1062.
- Hatipoglu, B., & Köse, C. (2016). *Gender Recognition from Face Images Using PCA and LBP*. [online] Tersedia di: <[http://www.eleco.org.tr/openconf\\_2015/modules/request.php?module=oc\\_proceedings&action=view.php&id=159&a=Accept+as+Poste](http://www.eleco.org.tr/openconf_2015/modules/request.php?module=oc_proceedings&action=view.php&id=159&a=Accept+as+Poste)> [Diakses 11 Januari 2019]
- Li, S., & Harner, E. J. (2014). Random KNN. *IEEE International Conference on Data Mining Workshop*.
- Liantoni, F. (2015). *Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor*. [Online] Tersedia di: <[https://www.researchgate.net/profile/Febri\\_Liantoni/publication/295547632\\_Klasifikasi\\_Daun\\_Dengan\\_Perbaikan\\_Fitur\\_Citra\\_Menggunakan\\_Metode\\_K-Nearest\\_Neighbor/links/56cb229308ae1106370b7377.pdf](https://www.researchgate.net/profile/Febri_Liantoni/publication/295547632_Klasifikasi_Daun_Dengan_Perbaikan_Fitur_Citra_Menggunakan_Metode_K-Nearest_Neighbor/links/56cb229308ae1106370b7377.pdf)> [Diakses 11 Januari 2019]
- Philips, P. J., Wechsler, H., Huang, J., & Rauss, P. (1998). The FERET database and evaluation procedure for face recognition algorithms. *Image and Vision Computing*, 16(5), 295-306.
- Rani, S., & Saepudin, D. (2013). Klasifikasi Jenis Kelamin Berdasarkan Citra Wajah Menggunakan Algoritma ADABOOST-SVM. *SEMNAS TEKNOLOGI ONLINE*, 1(1).
- Shan, C. (2012). Learning local binary patterns for gender classification on real-world face images. *Pattern Recognition Letters*, 431-437.
- Tawhid, M. N., & Dey, E. K. (2018). A Gender Recognition System from Facial Image. *International Journal of Computer Applications*, 5-14.
- TIN, H. H. (2011). Gender And Age Estimation Based On Facial Images. Yangon, Myanmar.