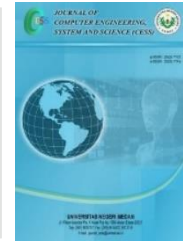


Contents list available at www.jurnal.unimed.ac.id

CESS (Journal of Computing Engineering, System and Science)

journal homepage: <https://jurnal.unimed.ac.id/2012/index.php/cess>



Implementasi Random Forest Pada Pengenalan Wajah Menggunakan Fitur Isomap

Implementation of Random Forest on Face Recognition Using Isomap Features

Rifki Kosasih^{1*}, Achmad Fahrurrozi², Desti Rimirasih³

^{1,2,3} Pusat Studi Komputasi Matematika, Teknik Informatika, Universitas Gunadarma
Jl. Margonda Raya No. 100, Depok, Indonesia, 16424

email: ¹rifki_kosasih@staff.gunadarma.ac.id, ²achmad_fahrurrozi@staff.gunadarma.ac.id,
³destimath@staff.gunadarma.ac.id

Submitted: 13 Mei 2022 | Revision: 07 Juni 2022 | Accepted: 29 Juli 2022

ABSTRAK

Sistem pengenalan wajah merupakan salah satu bidang yang digunakan untuk mengenali wajah seseorang. Dalam penelitian ini, data yang dikumpulkan merupakan data citra wajah yang terdiri dari 24 citra dengan komposisi 6 orang dan tiap orang memiliki 4 citra dengan berbagai ekspresi. Untuk mengenali wajah tersebut, dilakukan ekstraksi fitur wajah terlebih dahulu menggunakan metode isomap. Isomap merupakan metode reduksi dimensi yang dapat mereduksi dari dimensi tinggi menjadi fitur-fitur yang berdimensi rendah. Berdasarkan hasil ekstraksi diperoleh 4 fitur yang digunakan untuk mengklasifikasikan wajah. Untuk mengklasifikasikan wajah, digunakan algoritma random forest. Berdasarkan hasil penelitian diperoleh tingkat akurasi hasil klasifikasi sebesar 87,5%, nilai weighted average precision sebesar 81,25% dan nilai weighted average recall sebesar 87,5%.

Kata Kunci: *Pengenalan Wajah; Ekstraksi Fitur; Isomap; Random Forest*

ABSTRACT

Face recognition system is one of the fields used to recognize a person's face. In this study, the data collected is facial image data consisting of 24 images with a composition of 6 people and each person has 4 images with various expressions. To recognize the face, we perform facial feature extraction by using isomap method. Isomap is a dimensional reduction method that can reduce high dimensional to low dimensional features. Based on the extraction results obtained 4 features that are used to classify faces. To classify faces, we use a random forest algorithm. Based on the results of the study, the accuracy of the classification results was

*Penulis Korespondensi:
email: rifki_kosasih@staff.gunadarma.ac.id

87.5%, the weighted average precision value was 81.25% and the weighted average recall value was 87.5%.

Keywords: *Face Recognition; Feature Extraction; Isomap; Random Forest*

1. PENDAHULUAN

Kamera CCTV merupakan salah satu media yang digunakan untuk mengawasi lingkungan sekitar. Untuk mengenali seseorang yang memasuki lingkungan tersebut maka digunakan sistem pengenalan wajah. Wajah manusia merupakan salah satu fitur unik yang dapat digunakan untuk mengenali seseorang, karena tiap orang memiliki wajah yang berbeda-beda [1], [2].

Dalam melakukan pengenalan wajah, digunakan bantuan komputer untuk mengekstrak fitur wajah. Ekstraksi fitur wajah tersebut dilakukan untuk mengenali pola fitur wajah dan untuk mendapatkan fitur khusus pada wajah yang dapat dikenali oleh komputer. Salah satu metode yang digunakan pada penelitian terkait pengenalan wajah adalah metode manifold learning. Manifold learning merupakan metode reduksi dimensi yang digunakan untuk mengenali pola fitur wajah. Salah satu algoritma dalam manifold learning yang dapat digunakan adalah algoritma Isomap [3].

Isomap merupakan suatu prosedur reduksi dimensi dari dimensi tinggi ke dimensi yang lebih rendah yang diperkenalkan oleh Tenenbaum pada tahun 2000 [4]. Tenenbaum melakukan pengenalan pola pergerakan wajah dan pergerakan tangan dengan menggunakan kumpulan citra terkait [4].

Beberapa penelitian telah dilakukan di bidang pengenalan wajah seperti Salamun menerapkan *Euclidean Distance* dalam pengklasifikasian wajah berdasarkan citra wajah yang direduksi menggunakan metode *Principle Component Analysis* (PCA). PCA mereduksi dimensi untuk mengekstrak fitur berdasarkan vektor eigen dari citra wajah. Pada penelitiannya citra wajah yang digunakan sebanyak 130 citra. Berdasarkan hasil penelitiannya diperoleh tingkat akurasi sebesar 82,27% [5].

Penelitian berikutnya dilakukan oleh Rosyani juga menerapkan metode *Principle Component Analysis* (PCA) tetapi menggunakan *Canberra Distance* untuk mencocokkan wajah yang ada pada data uji dengan wajah yang ada pada data latih. Dalam penelitiannya citra wajah yang digunakan sebanyak 40 buah citra wajah dan diperoleh tingkat akurasi sebesar 84,5% [6].

Peneliti berikutnya Fahrurozi menggunakan metode *Local Binary Pattern* (LBP) dan *Principle Component Analysis* (PCA) untuk melakukan pengenalan wajah dengan ekspresi dan tingkat pencahayaan yang bervariasi. Dalam penelitiannya dilakukan ekstraksi fitur dengan menggunakan metode PCA berdasarkan citra luaran dari operator LBP. Data yang digunakan adalah 40 data citra wajah yang terdiri dari 8 orang dengan tiap orang memiliki 5 citra wajah dengan ekspresi dan tingkat pencahayaan yang berbeda. Berdasarkan hasil penelitiannya, tingkat akurasi yang diperoleh sebesar 75% [7].

Pada penelitian selanjutnya, Kosasih melakukan pengelompokan wajah manusia dengan menggunakan metode isomap. Metode isomap akan mereduksi dimensi citra dari dimensi tinggi menjadi dimensi yang lebih rendah sehingga fitur hasil reduksi dapat divisualisasikan. Dalam penelitiannya, kosasih menggunakan citra yang terdiri dari 5 orang dengan tiap orang memiliki 4 variasi citra wajah. Berdasarkan hasil penelitiannya, citra-citra wajah yang berasal dari orang yang sama dengan variasi wajah yang berbeda-beda mengelompok ke dalam satu

wilayah akan tetapi dalam penelitiannya belum dilakukan pengenalan wajah dengan menggunakan metode klasifikasi wajah [8].

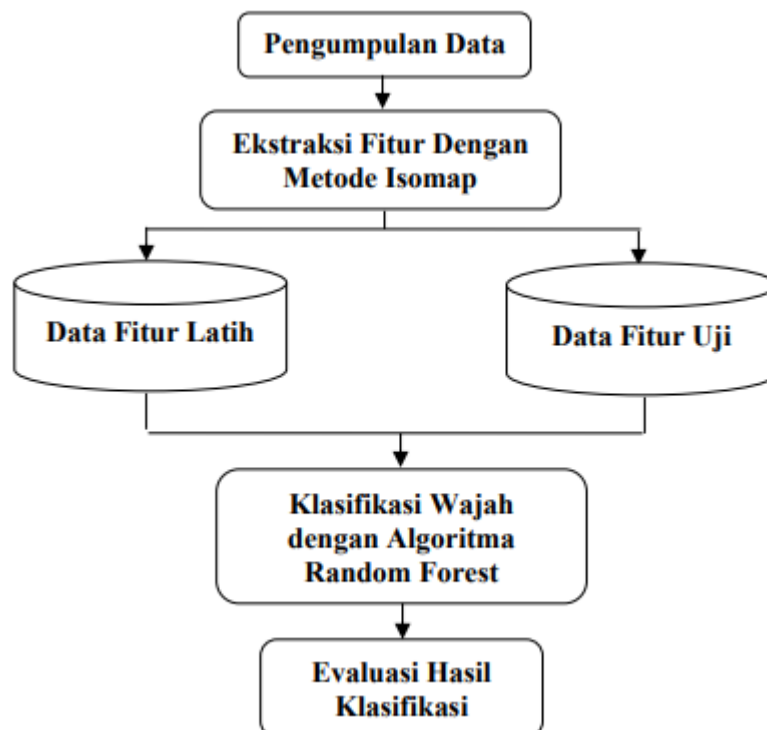
Kosasih melanjutkan penelitiannya, fitur-fitur wajah yang sudah dikelompokkan dengan metode isomap diklasifikasikan dengan menggunakan algoritma *K-Nearest Neighbor* (KNN) untuk mengenali wajah seseorang [9]. Dalam penelitiannya, citra wajah yang digunakan terdiri dari 6 orang dengan tiap orang memiliki 4 variasi citra wajah.

Citra-citra yang sudah diperoleh, selanjutnya direduksi dimensinya untuk mendapatkan fitur dengan menggunakan metode *isometric mapping* (isomap). Kemudian fitur-fitur tersebut diklasifikasikan dengan menggunakan algoritma KNN dengan jumlah tetangga terdekat $k=1$. Berdasarkan hasil klasifikasi diperoleh tingkat akurasi sebesar 83,33% [9].

Berdasarkan penelitian sebelumnya, Algoritma yang digunakan adalah KNN dan pemilihan data latih dan data uji tidak secara acak. Oleh karena itu, pada penelitian ini dilakukan pemilihan data uji dan data latih secara acak dan menggunakan metode pengklasifikasi lain seperti Random Forest. Selain itu, evaluasi yang dilakukan tidak hanya menghitung tingkat akurasi, tetapi juga menghitung recall dan presisi.

2. METODE PENELITIAN

Pada bagian ini, untuk melakukan pengenalan wajah dilakukan Langkah-langkah seperti pada Gambar 1.



Gambar 1. Tahapan Penelitian

Berdasarkan Gambar 1, tahapan pertama dalam penelitian ini adalah melakukan pengumpulan data citra wajah. Pada penelitian ini, citra wajah yang digunakan sebanyak 24 citra yang terdiri dari 6 orang dengan tiap orang memiliki 4 variasi wajah. Citra-citra tersebut

berukuran 280×315 dan diperoleh dari ORL *database*. Citra-citra tersebut dapat dilihat pada Gambar 2.

Setelah citra-citra tersebut diperoleh, selanjutnya dilakukan ekstraksi fitur citra wajah dengan menggunakan metode reduksi dimensi seperti Isomap. Isomap merupakan suatu metode yang dapat mereduksi dimensi sehingga diperoleh fitur-fitur dengan dimensi yang lebih rendah.



Gambar 2. Dataset Citra Wajah

2.1. Ekstraksi Fitur Dengan Metode Isomap

Misal diberikan video klip dari suatu CCTV yang memuat barisan frame $(f_1, f_2, f_3, \dots, f_k)$ berukuran $m \times n$ piksel, maka tiap frame tersebut berada pada ruang berdimensi $m.n$. Prinsip utama dari implementasi Isomap terhadap basis data berupa citra-citra yang merupakan frame video adalah dengan mentransformasi citra berukuran $m \times n$ tersebut menjadi vektor-vektor $v_1, v_2, v_3, \dots, v_k$ dalam ruang berdimensi d . Vektor-vektor $v_1, v_2, v_3, \dots, v_k$ berdimensi d disebut fitur manifold video [10]. Untuk keperluan visualisasi sebaran fitur, nilai d yang umumnya digunakan adalah lebih kecil dari 4. Tahap awal dalam implementasi Isomap terhadap basis data berupa citra adalah dengan menyusun vektor-vektor kolom atau baris dari citra tersebut menjadi vector berbentuk P -tupel (P pasangan terurut).

Pada penelitian ini, dilakukan penskalaan ulang pada 24 citra wajah yang digunakan, sedemikian sehingga memiliki ukuran yang sama, yaitu 280×315 piksel. Dengan demikian, tiap citra disusun ulang menjadi vektor berbentuk 88.200-tupel dalam ruang berdimensi 88.200, misal disimbolkan ruang X . Secara umum, terdapat 3 tahap dalam memperoleh fitur manifold video, yaitu:

- Bentuk graf $G(V, E)$, di mana V adalah himpunan vektor yang merepresentasikan citra wajah dalam basis data dalam penelitian ini. Himpunan E menyatakan himpunan busur dari tiap x_i dan x_j dalam ruang X , dinotasikan $d_X(i, j)$.

- b. Berikutnya didefinisikan jarak terpendek antar dua verteks dalam graf G, dinotasikan $d_G(i, j)$, dengan inisialisasi awal:
 $d_G(i, j) = d_X(i, j)$ jika x_j merupakan salah satu tetangga dari K tetangga dari x_i atau jarak antara x_i dan x_j kurang dari ϵ , untuk suatu ϵ bilangan riil positif, $d_G(i, j) = \infty$ jika tidak demikian. Nilai tiap busur diperbaharui dengan menerapkan pencarian jarak terpendek antar tiap verteks dalam graf G, secara matematis dirumuskan seperti persamaan (1):

$$d_G(i, j) = \min\{d_G(i, j), d_G(i, k) + d_G(k, j)\} \quad (1)$$

Untuk setiap x_i dan x_j . Pada penelitian ini, digunakan prinsip K-nearest neighbor dalam penentuan $d_X(i, j)$, dengan nilai K = 20. Selanjutnya bentuk matriks $D_G = [d_G(i, j)]$.

- c. Mengkonversi ruang X ke ruang Y berdimensi d, di mana Y berdimensi lebih rendah dibanding X. Hal ini dilakukan dengan membentuk matriks B_G yang diperoleh dari matriks D_G , dengan aturan seperti pada persamaan (2)

$$B_G = -\frac{1}{2} \left(I - \frac{1}{n} S \cdot S^T \right) D_G^2 \left(I - \frac{1}{n} S \cdot S^T \right) \quad (2)$$

dengan I menyatakan matriks Identitas berukuran 24 x 24, S menyatakan matriks yang entri-entri-nya adalah angka 1, dan D_G^2 menyatakan matriks yang entri-entri-nya diperoleh dengan mengkuadratkan tiap entri yang bersesuaian dalam D_G . Output dari isomap adalah matriks A yang diperoleh dari eigen dekomposisi, secara matematis dituliskan seperti pada persamaan (3)

$$Y = [\sqrt{\lambda_1}v_1 \quad \sqrt{\lambda_2}v_2 \quad \dots \quad \sqrt{\lambda_d}v_d]^T \quad (3)$$

dengan $\lambda_1, \lambda_2, \dots, \lambda_d$ adalah m buah nilai eigen dengan nilai tertinggi dari matriks B_G , dan v_1, v_2, \dots, v_d adalah vektor-vektor eigen yang berkorespondensi dengan nilai-nilai eigen tersebut. Pada penelitian ini, nilai d yang digunakan adalah d = 4. Dengan demikian, dalam penelitian ini diperoleh suatu himpunan fitur manifold video, berdasarkan nilai parameter K = 20 dan d = 4.

2.2. Random Forest

Sebuah pengklasifikasi Ensemble terdiri dari satu himpunan pengklasifikasi yang dilatih secara individual (seperti jaringan saraf atau pohon keputusan) yang prediksinya digabungkan untuk mengklasifikasikan sampel baru [11]. Penelitian sebelumnya telah menunjukkan bahwa Ensemble seringkali lebih akurat daripada pengklasifikasi tunggal mana pun dalam Ensemble [12], [13]. Metode yang populer untuk memproduksi Ensemble adalah Bagging [14] dan Boosting [15].

Random forest merupakan suatu teknik ensemble supervised machine learning dimana suatu ensemble yang dibangun berdasarkan pohon keputusan [1q]. Random forest didefinisikan sebagai pengklasifikasi yang terdiri dari kumpulan pengklasifikasi berstruktur pohon $\{h(x, \Theta_k), k=1,2,3 \dots\}$, dengan $\{\Theta_k\}$ adalah vektor acak yang independen dan berdistribusi identik dan setiap pohon membentuk satu unit memilih kelas yang paling populer pada data input x [16].

Setiap Pohon Keputusan dibuat dengan memilih data secara acak dari data yang tersedia. Misalnya Random forest untuk setiap Pohon Keputusan dapat dibangun dengan pengambilan sampel secara acak subset fitur, dan/atau dengan pengambilan sampel acak dari subset data pelatihan untuk setiap Pohon Keputusan (konsep Bagging). Pada random forest, fitur dipilih secara acak di setiap pemisahan keputusan [17]. Fitur acak dipilih dalam proses induksi. Prediksi dibuat dengan menggabungkan prediksi ensemble. Setiap pohon keputusan ditumbuhkan seperti dijelaskan pada [16] yaitu:

- a. Dengan Sampling N secara acak dari data asli, Jika jumlah kasus dalam set pelatihan adalah N tetapi dengan penggantian. Sampel ini akan digunakan sebagai set pelatihan untuk menumbuhkan pohon. Pada penelitian ini digunakan N = 200.
- b. Untuk jumlah M variabel input, variabel m dipilih sedemikian rupa sehingga $m < \text{node}$, m variabel dipilih secara acak dari M dan pemisahan terbaik pada m ini digunakan untuk memisahkan node. Selama pertumbuhan hutan, nilai m dipertahankan konstan. Dalam penelitian ini digunakan nilai M = 4.
- c. Setiap pohon ditanam sejauh mungkin. Tidak ada pemangkasan yang digunakan.

2.3. Evaluasi Model

Pada penelitian ini, dilakukan evaluasi model melalui pengukuran performa hasil klasifikasi berdasarkan parameter Akurasi, Presisi, dan *Recall*. Perhitungan Akurasi, Presisi, dan *Recall* didasarkan pada *confusion matrix* hasil klasifikasi. *Confusion matrix* atau sering dikenal sebagai Tabel Kontingensi [18], merupakan tabel yang terdiri dari aktual dan data prediksi dari masing-masing kelas yang diamati. Data aktual terletak pada bagian kolom sementara data prediksi terletak pada bagian baris, seperti disajikan oleh Tabel 1 [19].

Tabel 1. *Confusion Matrix*

	Aktual Positif	Aktual Negatif
Prediksi Positif	True Positive (TP)	False Negative (FN)
Prediksi Negatif	False Positive (FP)	True Negative (TN)

Berdasarkan Tabel 1, dalam suatu penelitian klasifikasi terdapat empat kemungkinan kejadian (instance). Jika instance positif dikelompokkan sebagai positif maka disebut sebagai *true positive* (TP) dan jika instance positif dikelompokkan sebagai negatif disebut *false positive* (FP). Jika instance negatif dinyatakan sebagai negatif maka disebut sebagai *true negative* (TN) dan jika instance negatif diprediksi positif maka disebut *false negative* (FN) [20]. Dengan demikian, secara matematis tingkat Akurasi dihitung dengan formula pada persamaan (4).

$$Akurasi = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \quad (4)$$

Di sisi lain, Presisi menyatakan perbandingan antara banyaknya kelas positif yang diprediksi dengan benar dan banyaknya data kelas positif, sementara *Recall* atau biasa disebut Sensitivitas, menyatakan perbandingan antara jumlah kelas positif yang diprediksi dengan benar dan semua data yang diprediksi sebagai kelas positif. Secara matematis, Presisi dan *Recall* berturut-turut dihitung menggunakan persamaan (5) dan persamaan (6) [21].

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

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

Jika kasus yang dihadapi adalah klasifikasi multi kelas atau klasifikasi yang memuat lebih dari dua kelas, misal N kelas, maka instance positif yang dimaksud adalah kelas yang diamati, sedemikian sehingga terdapat total N buah nilai Presisi dan Recall, satu untuk setiap kelas yang diamati. Untuk pengamatan performa suatu model klasifikasi multi kelas secara menyeluruh, umumnya Presisi dan Recall dihitung secara rata-rata, diantaranya dengan *Weighted Average Precision* (WAP) dan *Weighted Average Recall* (WAR) seperti pada persamaan (7) dan (8)[22].

$$WAP = \frac{\sum_{i=1}^n p_i d_i}{\sum_{i=1}^n d_i} \quad (7)$$

$$WAR = \frac{\sum_{i=1}^n r_i d_i}{\sum_{i=1}^n d_i} \quad (8)$$

Dengan p_i merupakan nilai presisi untuk kelas ke- i , d_i merupakan jumlah data aktual pada kelas ke- i dan r_i merupakan nilai recall untuk kelas ke- i .

3. HASIL DAN PEMBAHASAN

Pada penelitian ini, data yang dikumpulkan sebanyak 24 citra wajah yang terdiri dari 6 orang dengan tiap orang memiliki 4 citra wajah dengan ekspresi dan tingkat pencahayaan yang berbeda. Setelah citra wajah diperoleh, selanjutnya dilakukan ekstraksi fitur wajah dengan menggunakan metode *isometric mapping* (isomap). Metode ini digunakan untuk mereduksi citra yang berdimensi tinggi menjadi fitur-fitur yang berdimensi lebih rendah. Hasil ekstraksi fitur dengan menggunakan metode isomap ditunjukkan pada Tabel 2.

Tabel 2. Fitur Hasil Isomap (18 dari 24 Data Citra)

No	fitur1	fitur2	fitur3	fitur4	Kelas
1.	1865,715	539,9132	-732,798	130,2959	1
2.	2591,17	-913,433	-676,285	1808,866	1
3.	3353,991	895,3262	-1905,87	-547,668	1
4.	-1669,34	5691,88	4499,27	3580,42	2
5.	-537,349	8090,167	4172,715	906,2219	2
6.	-1386,63	7475,373	4079,722	387,3025	2
7.	-5018,16	3229,181	-3885,48	-3687,17	3
8.	-3337,56	4378,413	-3404,94	-1299,32	3
9.	-3752,17	1644,93	-2640,35	-4241,87	3
10.	-8374,26	-2811,57	730,7217	1752,626	4
11.	-9293,96	-3889,42	392,4702	2361,835	4
12.	-8536,55	-2661,72	882,6216	-3641,11	4
13.	2871,136	-2708,39	-1427,35	-999,04	5
14.	4279,461	-13795,8	-2087,92	1299,07	5
15.	10899,96	-5659,91	11283,26	-5583,11	5
16.	2191,061	-553,263	-972,016	-44,2337	6

17.	7471,734	1367,492	-3430,37	8072,981	6
18.	692,9532	-5742,52	573,569	2388,598	6

Pada Tabel 2 diperoleh 24 data dengan tiap data memiliki 4 buah fitur yang digunakan sebagai masukan dalam pengenalan wajah. Data-data yang sudah diperoleh selanjutnya secara acak dibagi menjadi 2 bagian yaitu 16 data sebagai data latih dan 8 data sebagai data uji.

Tahapan selanjutnya adalah melakukan klasifikasi wajah dengan menggunakan algoritma random forest. Pada penelitian ini, banyaknya sampling N yang digunakan adalah 200. Hasil klasifikasi dengan menggunakan algoritma random forest dapat dilihat pada Tabel 3.

Tabel 3. Hasil Klasifikasi dengan Algoritma Random Forest

No	Aktual	Prediksi
1	6	1
2	5	5
3	2	2
4	4	4
5	3	3
6	4	4
7	3	3
8	1	1

Pada Tabel 3, terdapat 8 data uji yang digunakan. Kolom 2 merupakan data sebenarnya sedangkan pada kolom 3 merupakan data hasil prediksi dengan menggunakan algoritma random forest. Setelah itu, dilakukan pengujian model dengan menggunakan matriks confusion seperti pada Tabel 4.

Tabel 4. Matriks Confusion dari Algoritma Random Forest

		Prediksi					
		1	2	3	4	5	6
Aktual	1	1	0	0	0	0	0
	2	0	1	0	0	0	0
	3	0	0	2	0	0	0
	4	0	0	0	2	0	0
	5	0	0	0	0	1	0
	6	1	0	0	0	0	0

Berdasarkan Tabel 4, terdapat 1 orang yang merupakan wajah orang ke-1 pada keadaan sebenarnya diprediksi sebagai orang ke-1. Wajah orang ke-2 pada keadaan sebenarnya diprediksi sebagai orang ke-2 sebanyak satu orang. Wajah orang ke-3 pada keadaan sebenarnya diprediksi sebagai orang ke-3 sebanyak dua orang. Wajah orang ke-4 pada keadaan sebenarnya diprediksi sebagai orang ke-4 sebanyak dua orang. Wajah orang ke-5 pada keadaan sebenarnya diprediksi sebagai orang ke-5 sebanyak satu orang sedangkan wajah orang ke-6 pada keadaan sebenarnya tetapi diprediksi sebagai orang ke-1 sebanyak satu orang. Setelah itu, dilakukan evaluasi model dengan menghitung tingkat akurasi menggunakan persamaan (4)

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{TN}+\text{FN}} \times 100\% \\
 &= \frac{7+0}{7+1+0+0} \times 100\% \\
 &= \frac{7}{8} \times 100\% = 87,5\%
 \end{aligned}$$

Sedangkan nilai presisi untuk tiap kelas dan nilai *weighted average precision* adalah:

$$p_1 = \frac{\text{TP}_1}{\text{TP}_1+\text{FP}_1} \times 100\% = \frac{1}{1+1} \times 100\% = 50\%$$

$$p_2 = \frac{\text{TP}_2}{\text{TP}_2+\text{FP}_2} \times 100\% = \frac{1}{1+0} \times 100\% = 100\%$$

$$p_3 = \frac{\text{TP}_3}{\text{TP}_3+\text{FP}_3} \times 100\% = \frac{2}{2+0} \times 100\% = 100\%$$

$$p_4 = \frac{\text{TP}_4}{\text{TP}_4+\text{FP}_4} \times 100\% = \frac{2}{2+0} \times 100\% = 100\%$$

$$p_5 = \frac{\text{TP}_5}{\text{TP}_5+\text{FP}_5} \times 100\% = \frac{1}{1+0} \times 100\% = 100\%$$

$$p_6 = \frac{\text{TP}_6}{\text{TP}_6+\text{FP}_6} \times 100\% = \frac{0}{0+1} \times 100\% = 0\%$$

$$\text{WAP} = \frac{1 \times 50\% + 1 \times 100\% + 2 \times 100\% + 2 \times 100\% + 1 \times 100\% + 0 \times 100\%}{1+1+2+2+1+0} = 81,25\%$$

Selain itu, dihitung juga recall dan weighted average recall dari hasil matriks confusion menggunakan persamaan (6) dan (8) yang dapat dilihat pada Tabel 5.

Tabel 5. Hasil Evaluasi Dari Algoritma Random Forest

Kategori	Random Forest		Jumlah Data Uji
	Presisi	Recall	
1	50 %	100%	1
2	100%	100 %	1
3	100%	100 %	2
4	100 %	100 %	2
5	100%	100%	1
6	0%	0%	1
Weighted Average	81,25%	87,5%	

Berdasarkan hasil penelitian pada algoritma random forest diperoleh tingkat akurasi sebesar 87,5%, nilai weighted average precision (WAP) sebesar 81,25% dan nilai weighted average recall (WAR) sebesar 87,5%.

4. KESIMPULAN

Untuk membantu pihak keamanan dalam mengawasi lingkungan sekitar dengan menggunakan kamera CCTV, dibutuhkan suatu sistem yang dapat mengenali seseorang. Salah satu cara untuk mengenali seseorang adalah dengan menggunakan fitur wajah. Wajah merupakan fitur unik yang dapat membedakan seseorang dengan orang yang lain.

Pada penelitian ini, data yang dipakai merupakan citra wajah yang terdiri dari 6 orang dengan tiap orang memiliki 4 citra yang memiliki ekspresi dan pencahayaan yang berbeda sehingga total citra sebanyak 24 citra. Setelah itu dilakukan ekstraksi citra dengan menggunakan metode isomap. Isomap merupakan suatu metode reduksi dimensi yang digunakan untuk mencari fitur wajah yang dapat membedakan seseorang dengan orang yang lain.

Tahapan selanjutnya adalah melakukan klasifikasi dengan menggunakan algoritma random forest. Berdasarkan hasil penelitian diperoleh tingkat akurasi sebesar 87,5%, nilai weighted average precision sebesar 81,25% dan nilai weighted average recall sebesar 87,5%.

Untuk penelitian yang akan datang, data yang digunakan diperbanyak. Selain itu, sebagai pembandingan digunakan juga metode lain untuk melakukan pengenalan wajah seperti metode deep learning.

REFERENSI

- [1] R. Kosasih, "Penggunaan Metode Linear Discriminant Analysis Untuk Pengenalan Wajah Dengan Membandingkan Banyaknya Data Latih," *J. Ilm. Teknol. dan Rekayasa*, vol. 26, no. 1, pp. 25–34, 2021.
- [2] R. Kosasih and C. Daomara, "Pengenalan Wajah dengan Menggunakan Metode Local Binary Patterns Histograms (LBPH)," *Media Inform. Budidarma*, vol. 5, pp. 1258–1264, 2021, doi: 10.30865/mib.v5i4.3171.
- [3] L. Van der Maaten, E. Postma, and J. Van den Herik, "Dimensionality Reduction : A Comparative Review," Tilburg, Netherlands, 2009.
- [4] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science (80-.)*, vol. 290, no. December, pp. 2319–2323, 2000.
- [5] Salamun and F. Wazir, "Rancang Bangun Sistem Pengenalan Wajah Dengan Metode Principal Component Analysis," *J. Teknol. dan Sist. Inf. UNIVRAB*, vol. 1, no. 2, pp. 59–75, 2016, doi: 10.36341/rabit.v1i2.25.
- [6] P. Rosyani, "Pengenalan Wajah Menggunakan Metode Principal Component Analysis (PCA) dan Canberra Distance," *J. Inform. Univ. Pamulang*, vol. 2, no. 2, p. 118, 2017, doi: 10.32493/informatika.v2i2.1515.
- [7] A. Fahrurozi and R. Kosasih, "Face Recognition Using Local Binary Pattern Combined With PCA For Images Under Various Expression and Illumination," in *Proceeding on International Workshop on Academic Collaboration 2017*, 2017, no. May, pp. 1–7.
- [8] R. Kosasih and A. Fahrurozi, "Clustering of Face Images by Using Isomap method," in *Proceeding on International Workshop on Academic Collaboration 2017*, 2017, no. May, pp. 52–56.

- [9] R. Kosasih, "Kombinasi Metode Isomap dan KNN Pada Image Processing Untuk Pengenalan Wajah," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 5, no. 2, pp. 166–170, 2020.
- [10] S. Menaria and D. Mukherjee, "Video Manifold Feature Extraction Based on ISOMAP," *Int. J. Eng. Sci. Invent.*, vol. 4, no. 4, pp. 64–67, 2015, [Online]. Available: www.ijesi.org.
- [11] V. Y. Kullarni and P. K. Sinha, "Random Forest Classifier: A Survey and Future Research Directions," *Int. J. Adv. Comput.*, vol. 36, no. 1, pp. 1144–1156, 2013.
- [12] A. Krogh and J. Vedelsby, "Neural Network Ensembles, Cross Validation, and Active Learning Anders," pp. 6–7, 2010.
- [13] D. Opitz and R. Maclin, "Popular ensemble learning: an empirical study," *J. Artif. Intell. Res.*, vol. 11, pp. 169–198, 1999.
- [14] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996, doi: 10.1007/bf00058655.
- [15] R. E. Schapire, "The Boosting Approach to Machine Learning: An Overview BT - Nonlinear Estimation and Classification," *Nonlinear Estim. Classif.*, vol. 171, no. Chapter 9, pp. 149–171, 2003, [Online]. Available: http://link.springer.com/10.1007/978-0-387-21579-2_9
- [16] M. Reza, S. Miri, and R. Javidan, "A Hybrid Data Mining Approach for Intrusion Detection on Imbalanced NSL-KDD Dataset," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 6, pp. 1–33, 2016, doi: 10.14569/ijacsa.2016.070603.
- [17] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, "Random forests and decision trees," *IJCSI Int. J. Comput. Sci. Issues*, vol. 9, no. 5, pp. 272–278, 2012.
- [18] I. Amalia, "Pengenalan Citra Tanda Tangan Menggunakan Gray Level Co-Occurrence Matrix (GLCM) Dan Probabilistic Neural Network (PNN)," *E-Jurnal Politek. negeri Lhokseumawe*, vol. 14, pp. 29–34, 2014.
- [19] L. A. Septiandi, E. M. Yuniarno, and A. Zaini, "Deteksi Kedipan dengan Metode CNN dan Percentage of Eyelid Closure (PERCLOS)," *Tek. Its*, vol. 10, no. 1, pp. A56–A57, 2021.
- [20] T. Yulianti, M. Telaumbanua, H. D. Septama, and H. Fitriawan, "Pengaruh Seleksi Fitur Citra Terhadap Klasifikasi Tingkat the Effect of Image Feature Selection on the Local Beef," vol. 10, no. 1, pp. 85–95, 2021.
- [21] D. P. Lestari, R. Kosasih, T. Handhika, Murni, I. Sari, and A. Fahrurozi, "Fire Hotspots Detection System on CCTV Videos Using You only Look Once (YOLO) Method and Tiny YOLO Model for High Buildings Evacuation," in *2nd International Conference of Computer and Informatics Engineering IC2IE*, 2019, pp. 87–92.
- [22] M. Sokolova and G. Lapalme, "A Systematic Analysis of Performance Measures for Classification Tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.