

Pengenalan Wajah Menggunakan PCA dengan Memperhatikan Jumlah Data Latih dan Vektor Eigen

Rifki Kosasih

Program Studi Komputasi Matematika, Fakultas Teknik Informatika, Universitas Gunadarma
Jl. Margonda Raya No. 100, Depok, Indonesia 16424
e-mail : rifki_kosasih@staff.gunadarma.ac.id

Submitted Date: October 07th, 2020
Revised Date: January 31st, 2021

Reviewed Date: December 30th, 2020
Accepted Date: June 14th, 2021

Abstract

To find out if an employee is present, attendance is usually used. Attendance can be done in several ways, one of which is by filling in the attendance list that has been provided (manual attendance). However, this method is less effective because there is a possibility that employees who are not present will entrust attendance to employees who are present. Therefore, other ways are needed so that this does not happen. In this study, attendance was carried out using facial recognition. Face recognition is one of the fields used to recognize someone. A person's face usually has special characteristics that are easily recognized by people. These special characteristics are also called features. In this study, these features can be searched using the Principle Component Analysis (PCA) method. The PCA method is one of the methods used to produce features by reducing dimensions using eigenvectors from facial images (eigenface). The facial image used in this study consisted of 40 people with each person having 10 facial images with various expressions. Image data is divided into two parts, namely training data and test data. In this study, it is proposed to pay attention to the amount of training data and the number of eigenvectors used to get the best level of accuracy. From the research results, the highest level of accuracy occurs when the training data for each person is 7 and the test data for each person is 3 with an accuracy rate of 96.67%.

Keywords: Principle Component Analysis (PCA); feature; Manifold Learning; eigenface; eigenvector

Abstrak

Untuk mengetahui seorang karyawan perusahaan hadir biasanya digunakan absensi. Absensi kehadiran dapat dilakukan dengan beberapa cara, salah satunya adalah dengan mengisi daftar hadir yang sudah disediakan (absensi manual). Akan tetapi cara tersebut kurang efektif karena ada kemungkinan karyawan yang tidak hadir menitip absensi kepada karyawan yang hadir. Oleh karena itu, dibutuhkan cara lain agar hal tersebut tidak terjadi. Dalam penelitian ini absensi dilakukan dengan menggunakan pengenalan wajah. Pengenalan wajah merupakan salah satu bidang yang digunakan untuk mengenali seseorang. Wajah seseorang biasanya memiliki ciri khusus sehingga mudah dikenali oleh orang-orang. Ciri khusus tersebut disebut juga dengan fitur. Dalam penelitian ini fitur-fitur tersebut dapat dicari dengan menggunakan metode Principle Component Analysis (PCA). Metode PCA merupakan salah satu metode yang digunakan untuk menghasilkan fitur-fitur dengan cara mereduksi dimensi menggunakan vektor eigen dari citra wajah (eigenface). Citra wajah yang digunakan dalam penelitian ini terdiri dari 40 orang dengan tiap orang memiliki 10 citra wajah dengan berbagai ekspresi. Data citra ini dibagi menjadi dua bagian yaitu data latih dan data uji. Pada penelitian ini diusulkan untuk memperhatikan banyaknya data latih dan banyaknya vektor eigen yang digunakan untuk mendapatkan tingkat akurasi yang terbaik. Dari hasil penelitian diperoleh tingkat akurasi tertinggi terjadi ketika data latih tiap orang sebanyak 7 dan data uji tiap orang sebanyak 3 dengan tingkat akurasi sebesar 96,67%.

Kata Kunci : Principle Component Analysis (PCA); fitur; Manifold Learning; eigenface; vektor eigen

1. Pendahuluan

Dalam sebuah perusahaan, absensi karyawan merupakan suatu keharusan agar pihak perusahaan mengetahui apakah seseorang hadir atau tidak. Pada awalnya, sistem absensi yang digunakan adalah menggunakan absensi manual dengan cara menulis ke dalam buku absensi. Seiring perkembangan teknologi, cara tersebut dinilai kurang efisien, sehingga dibutuhkan suatu sistem absensi tanpa harus menggunakan absensi manual.

Banyak cara digunakan untuk melakukan absensi berbasis teknologi yaitu menggunakan fitur-fitur khusus pada manusia yang menjadi ciri khas yang membedakan manusia yang satu dengan yang lain seperti wajah, sidik jari dan lain-lain. Dalam penelitian ini digunakan fitur pengenalan wajah untuk absensi.

Pengenalan wajah merupakan salah satu topik di computer vision yang digunakan untuk sistem keamanan, absensi (Kardile, 2017) dan lain-lain. Dalam melakukan pengenalan wajah, terdapat tiga proses (Budi et al., 2016). Proses pertama adalah proses pendeteksian wajah. Setelah itu dilakukan proses ekstraksi fitur. Proses ekstraksi fitur dilakukan untuk mendapatkan ciri khusus pada wajah. Proses terakhir adalah melakukan pengenalan wajah berdasarkan hasil ekstraksi fitur. Pada proses ini digunakan metode jarak antara dua buah fitur untuk pengenalan wajah.

Salah satu metode yang digunakan pada pengenalan wajah adalah Principle Component Analysis (PCA). PCA merupakan salah satu teknik dari Manifold Learning yaitu suatu teknik mereduksi dimensi yang digunakan untuk mengekstrak informasi (fitur) yang optimal dari citra dengan cara memilih vektor eigen yang maksimum (Kosasih et al., 2015)

Metode Principle Component Analysis juga digunakan untuk ekstraksi fitur dan merepresentasikan data (Zein, 2018). Metode ini mentransformasikan citra-citra wajah menjadi himpunan yang terdiri dari karakteristik citra fitur yang disebut eigenface (Muliawan et al., 2015).

Dalam proses pengenalan, citra wajah diproyeksikan dari ruang dimensi yang tinggi ke ruang dimensi yang lebih rendah tanpa menghilangkan banyak informasi dan di klasifikasikan dengan menggunakan metode jarak terdekat seperti jarak euclid (Mulyadi & Zaini, 2020).

Beberapa penelitian telah dilakukan dibidang pengenalan wajah seperti kosasih,

melakukan pengklasteran wajah manusia dengan menggunakan metode isomap. Dari hasil penelitiannya terlihat bahwa citra-citra wajah yang berasal dari orang yang sama dengan ekspresi wajah yang berbeda-beda mengelompok ke dalam satu area akan tetapi dalam penelitiannya belum dilakukan pengklasifikasian wajah (Kosasih & Fahrurrozi, 2017).

Peneliti berikutnya (Kosasih, 2020) menggunakan metode isomap dan KNN untuk melakukan pengenalan wajah. Pada penelitiannya, data yang digunakan adalah 24 citra yang terdiri dari 6 orang dengan tiap orang memiliki 4 citra dengan berbagai ekspresi. Metode isomap digunakan untuk mengekstrak fitur citra wajah. Setelah itu dilakukan klasifikasi dengan menggunakan metode KNN. Berdasarkan hasil penelitiannya diperoleh tingkat akurasi sebesar 83,33%.

Selanjutnya Ardiansiah melakukan pengenalan wajah dengan menggunakan data pelatihan dan pengujian sistem menggunakan *software development kit Luxand FaceSDK* (Ardiansiah et al., 2016). Berikutnya (Fandiansyah et al., 2017; Putra & Harjoko, 2018) menggunakan metode LDA untuk pengenalan wajah.

Peneliti selanjutnya (Fahrurrozi & Kosasih, 2017) menggunakan metode PCA dan *Local Binary Pattern* (LBP) untuk melakukan pengenalan wajah. Dalam penelitiannya Fahrurrozi et al menggunakan 40 data citra wajah yang terdiri dari 8 orang dengan tiap orang memiliki 5 citra wajah dengan berbagai ekspresi. Berdasarkan penelitiannya, tingkat akurasi yang diperoleh untuk mengenali wajah seseorang adalah sebesar 75%. Untuk meningkatkan akurasi maka Oleh pada penelitian ini digunakan sampel citra wajah yang lebih banyak dari penelitian sebelumnya.

Selanjutnya (Salamun & Wazir, 2016) dan (Rosyani, 2017) menggunakan metode PCA untuk melakukan pengenalan wajah. (Salamun & Wazir, 2016) dalam penelitiannya menggunakan dataset sebanyak 130 citra wajah dan diperoleh tingkat akurasi dengan metode tersebut sebesar 82,27%. (Rosyani, 2017) menggunakan Canberra Distance pada fitur hasil PCA. Pada penelitiannya ata yang digunakan sebanyak 40 buah citra wajah dan diperoleh tingkat akurasi sebesar 84,5%. Akan tetapi pada penelitian (Rosyani, 2017; Salamun & Wazir, 2016) tidak memperhatikan banyaknya data latih yang digunakan dan berapa banyak eigenvector (eigenface) yang digunakan agar mendapatkan tingkat akurasi yang terbaik.

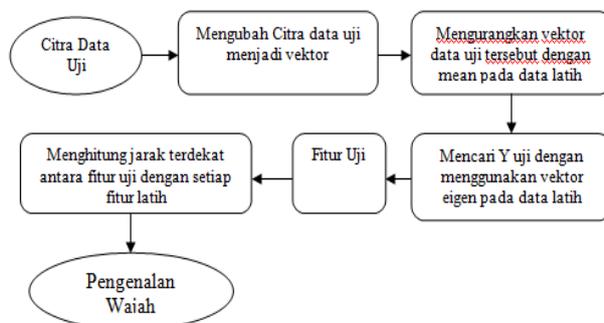
Oleh karena itu, dalam penelitian ini diusulkan pengenalan wajah menggunakan metode Principle Component Analysis (PCA) dengan memperhatikan banyaknya vektor eigen dan banyaknya data latih yang digunakan untuk mendapatkan tingkat akurasi yang terbaik.

2. Metode Penelitian

Pada penelitian ini, dilakukan beberapa tahapan penelitian untuk mengenali wajah seseorang yang dapat dilihat pada Gambar 1 dan Gambar 2.



Gambar 1. Tahapan Penelitian Pada Data Latih



Gambar 2. Tahapan Penelitian Pada Data Uji

Tahapan pertama dalam penelitian ini adalah melakukan pengumpulan data citra wajah. Pada penelitian ini data yang digunakan adalah data citra wajah yang terdiri dari 40 orang. Tiap orang memiliki 10 citra dengan berbagai ekspresi wajah sehingga total data citra yang digunakan adalah 400 citra yang berukuran sama.

Data citra tersebut dibagi menjadi dua yaitu data citra latih (Gambar 1) dan data citra uji (Gambar 2). Pada penelitian ini dibuat variasi pemilihan data latih dan data uji yang dapat dilihat pada Tabel 1.

Sampel citra yang digunakan pada

penelitian ini diperoleh dari ORL database seperti pada Gambar 3.



Gambar 3. Sampel Data Citra Wajah (12 dari 400 Data Citra)

Tahapan selanjutnya adalah menggunakan metode PCA untuk mendapatkan fitur latih.

2.1. Principle Component Analysis (PCA)

Principle component analysis merupakan salah satu teknik untuk mereduksi dimensi (Nasution, 2020). Tahapan metode PCA diterapkan pada data latih yang sudah dikumpulkan. Tahapan tersebut dapat dilihat pada Gambar 1 yaitu pertama setiap sampel data citra latih diubah menjadi vektor-vektor, kemudian digabung menjadi matriks X.

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1N} \\ x_{21} & \ddots & & x_{2N} \\ \vdots & & \ddots & \\ x_{M1} & x_{M2} & \dots & x_{MN} \end{bmatrix}_{M \times N}$$

Tahapan kedua adalah mencari *mean* vektor (m) dari matriks X berdasarkan rumus (1)

$$m = \frac{1}{N} \begin{bmatrix} \sum_{k=1}^N x_{1k} \\ \sum_{k=1}^N x_{2k} \\ \vdots \\ \sum_{k=1}^N x_{Mk} \end{bmatrix} \quad (1)$$

Di mana:

M = banyaknya baris dari matriks X,
 N = banyaknya kolom dari matriks X
 m = vektor rata-rata dari matriks X

Setelah itu membentuk matriks \bar{m} dengan cara menggandakan vektor m sebanyak N kali seperti rumus (2)

$$\bar{m} = [m \ m \ \dots \ m]_{M \times N} \quad (2)$$

Tahapan selanjutnya adalah menghitung Matriks kovariansi C dengan rumus (3)

$$C = (X - \bar{m})(X - \bar{m})^T \quad (3)$$

Setelah matriks kovariansi diperoleh, tahapan berikutnya adalah mencari vektor eigen yang bersesuaian dengan *eigen value* dari matriks C. Setelah itu dilakukan pengurutan vektor eigen dengan memperhatikan nilai eigen yang tertinggi kemudian dipilih beberapa vektor eigen yang tertinggi (maksimum) yang disebut dengan matriks P. Dalam penelitian ini diusulkan untuk memperhatikan berapa banyak vektor eigen yang harus dipilih.

Tahapan berikutnya adalah melakukan transformasi matriks X berukuran MxN menjadi matriks Y yang berdimensi lebih rendah. Y adalah matriks principle component yang berhubungan dengan transformasi linear P direpresentasikan dengan persamaan (4):

$$Y = P (X - \bar{m}) \quad (4)$$

Matriks Y yang diperoleh merupakan matriks yang mengandung fitur-fitur yang digunakan untuk pengenalan wajah.

Semua tahapan yang berada pada metode PCA digunakan untuk memperoleh fitur latih dan fitur uji seperti pada Gambar 2.

3. Hasil dan Pembahasan

Pada penelitian ini data yang digunakan adalah 400 data citra wajah yang terdiri dari 40 orang dan setiap orang memiliki 10 citra dengan berbagai variasi wajah. Selanjutnya, data citra yang sudah diperoleh dibagi menjadi dua yaitu data latih dan data uji. Jumlah data latih yang digunakan adalah sebanyak 5, 6, 7, 8, 9 data latih perorang.

Setelah itu digunakan metode Principle Component Analysis (PCA) untuk mendapatkan fitur latih dan fitur uji. Selanjutnya dicari fitur latih yang terdekat dengan fitur uji dengan menggunakan metode perhitungan jarak euclid.

Jika fitur uji sesuai dengan fitur latih yang dicari maka metode PCA berhasil dalam mengenali wajah. Hasil pengenalan wajah tersebut dapat dilihat pada Tabel 1.

Tabel 1. Hasil Pengenalan Wajah

Jumlah data latih Per orang	Jumlah data uji per orang	Jumlah vektor eigen	Akurasi
5	5	10%	90,5%
5	5	20%	91,5%
5	5	30%	91,5%
5	5	40%	91,5%
6	4	10%	95,625%
6	4	20%	95,625%
6	4	30%	95,625%
6	4	40%	95,625%
7	3	10%	96,67%
7	3	20%	96,67%
7	3	30%	96,67%
7	3	40%	96,67%
8	2	10%	96,25%
8	2	20%	96,25%
8	2	30%	96,25%
8	2	40%	96,25%
9	1	10%	95%
9	1	20%	95%
9	1	30%	95%
9	1	40%	95%

Pada penelitian ini jumlah data latih dan data uji tiap orang diambil dengan beberapa variasi pemilihan seperti pada Tabel 1. Jumlah data laith tiap orang yang dipakai adalah 5,6,7,8 dan 9 citra sehingga jumlah data uji yang dipakai adalah sisa dari data latih.

Pada baris kedua di Tabel 1, jika jumlah data latih tiap orang adalah 5 dan ada 40 orang maka sampel data latih yang digunakan ada 200 citra. Begitu juga dengan data uji, jika jumlah data uji perorang adalah 5 maka sampel data uji ada 200 citra.

Pada penelitian ini diperhatikan juga banyaknya vektor eigen yang digunakan untuk mendapatkan fitur latih dan fitur uji. Dalam penetiitan ini total vektor eigen yang diperoleh ada 1584 vektor eigen. Karena kita hanya memilih beberapa vektor eigen maksimum maka dalam penelitian ini dilakukan pemilihan vektor eigen sebesar 10%, 20%, 30% dan 40% dari total vektor eigen yang ada seperti pada kolom 3 pada Tabel 1.

Jika dilihat pada Tabel 1 terjadi perubahan persentase tingkat akurasi pada jumlah data latih tiap orang 5 ketika jumlah vektor eigen diperbanyak dari 10% menjadi 20%. Perubahan persentase tingkat akurasi juga terjadi ketika jumlah data latih diperbanyak dari 5 menjadi 7 akan tetapi tingkat akurasi menurun jika jumlah data latih per orang adalah 8 dan 9.

Dalam penelitian ini diperoleh bahwa persentase akurasi tertinggi didapat ketika jumlah data latih per orang 7 yaitu sebesar 96,67%.

4. Kesimpulan

Pengenalan wajah merupakan sesuatu hal yang penting yang digunakan untuk mengenali seseorang. Banyak bidang yang memanfaatkan pengenalan wajah ini, salah satunya untuk sistem absensi menggunakan wajah. Wajah seseorang biasanya memiliki ciri khusus sehingga mudah dikenali oleh orang-orang. Dalam penelitian ini digunakan metode Principle Component Analysis (PCA) untuk mendapatkan fitur-fitur yang dapat membedakan wajah seseorang dengan orang yang lainnya. Metode PCA menghasilkan fitur-fitur dengan cara mereduksi dimensi dengan menggunakan vektor eigen (eigenface) dari citra wajah. Citra wajah yang digunakan dalam penelitian ini adalah 400 citra wajah dengan komposisi 10 orang dan tiap orang memiliki 10 citra wajah dengan berbagai variasi wajah.

Dari hasil penelitian terjadi kenaikan persentase ketika jumlah data latih perorang bertambah dari 5 menjadi 7, akan tetapi persentase berkurang jika data latih ditambah menjadi 8 dan 9 sehingga diperoleh tingkat akurasi tertinggi terjadi ketika data latih tiap orang sebanyak 7 dan data uji tiap orang sebanyak 3 dengan tingkat akurasi sebesar 96,67%.

Untuk penelitian selanjutnya sebagai perbandingan dilakukan juga pengenalan wajah menggunakan metode lainnya seperti LDA, ANN dan yang lainnya untuk mengetahui metode yang terbaik dalam melakukan pengenalan wajah.

Referensi

- Ardiansyah, Setiawan, W., & Linawati. (2016). Sistem Pengenalan Wajah Dengan Metode Face Features. *E-Journal SPEKTRUM*, 3(2), 21–25.
- Budi, A., Suma'inna, & Maulana, H. (2016). Pengenalan Citra Wajah Sebagai Identifier Menggunakan Metode Principal Component Analysis (PCA). *Jurnal Teknik Informatika*, 9(2), 166–175. <https://doi.org/10.15408/jti.v9i2.5608>
- Fahrurrozi, A., & Kosasih, R. (2017). Face Recognition Using Local Binary Pattern Combined With PCA For Images Under Various Expression and Illumination. *Proceeding on International Workshop on Academic Collaboration 2017, May*, 1–7.
- Fandiansyah, Sari, J. Y., & Ningrum, I. P. (2017). Pengenalan Wajah Menggunakan Metode Linear Discriminant Analysis dan k Nearest Neighbor. *Jurnal ULTIMATICS*, 9(1), 1–9. <https://doi.org/10.31937/ti.v9i1.557>
- Kardile, H. A. (2017). Face Recognition Using Pca and Eigen Face Approach. *International Research Journal of Engineering and Technology (IRJET)*, 4(12), 460–463.
- Kosasih, R. (2020). Kombinasi Metode Isomap dan KNN Ppada Image Processing untuk Pengenalan Wajah. *CESS (Journal of Computer Engineering System and Science)*, 5(2), 166–170.
- Kosasih, R., & Fahrurrozi, A. (2017). Clustering of Face Images by Using Isomap method. *Proceeding on International Workshop on Academic Collaboration 2017, May*, 52–56.
- Kosasih, R., Madenda, S., Karyati, C. M., & Etp, L. (2015). Determination the Optimal Position from T1 and T2 Weighted MR Imaging of the Abdominal Aortic Aneurysm. *Advanced Science, Engineering and Medicine*, 7(10), 915–919. <https://doi.org/10.1166/ asem.2015.1778>
- Muliawan, M. R., Irawan, B., & Brianorman, Y. (2015). Implementasi Pengenalan Wajah Dengan Metode Eigenface Pada Sistem Absensi. *Jurnal Coding*, 03(1), 41–50.
- Mulyadi, R., & Zaini. (2020). Identifikasi Pengenalan Wajah Perokok Menggunakan Metode Principal Component Analysis. *Jurnal Resti*, 4(5), 892–898.
- Nasution, M. Z. (2020). Face Recognition based Feature Extraction using Principal Component Analysis (PCA). *Journal of Informatics and Telecommunication Engineering*, 3(2), 182–191.
- Putra, I. N. T. A., & Harjoko, A. (2018). Pengenalan Wajah Berbasis Mobile Menggunakan Fisherface dan Distance Classifier. *JST (Jurnal Sains Dan Teknologi)*, 7(1), 135–145. <https://doi.org/10.23887/jst-undiksha.v7i1.13267>
- Rosyani, P. (2017). Pengenalan Wajah Menggunakan Metode Principal Component Analysis (PCA) dan Canberra Distance. *Jurnal Informatika Universitas Pamulang*,

- 2(2), 118.
<https://doi.org/10.32493/informatika.v2i2.1515>
- Salamun, & Wazir, F. (2016). Rancang Bangun Sistem Pengenalan Wajah Dengan Metode Principal Component Analysis. *Jurnal Teknologi Dan Sistem Informasi UNIVRAB*, 1(2), 59–75.
- <https://doi.org/10.36341/rabit.v1i2.25>
- Zein, A. (2018). Pendeteksian Multi Wajah dan Recognition Secara Real Time Menggunakan Metoda Principle Component Analysis (PCA) dan Eigenface. *Jurnal Teknologi Informasi ESIT*, 12(01), 1–7.