

# ANALISA PERBANDINGAN METODE MAHALANOBIS DISTANCE DAN EUCLIDEAN DISTANCE PADA SISTEM PENGENALAN WAJAH

Arnold Nasir<sup>1)</sup>

<sup>1)</sup>Prodi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Atma Jaya Makassar  
Alamat e-mail: arnold\_nasir@outlook.com

## ABSTRACT

Dalam kurun waktu 20 tahun terakhir teknologi dibidang pengenalan citra mengalami peningkatan yang signifikan. Berbagai metode pencocokan kedekatan citra seperti Euclidean distance merupakan salah satu metode yang paling umum digunakan pada sebuah sistem pengenalan wajah. Akan tetapi metode tersebut memiliki beberapa kelemahan, seperti hasil pencocokan yang kurang akurat serta proses pencocokan data dengan dataset yang membutuhkan waktu yang lama. Guna mengatasi masalah tersebut maka penelitian ini mencoba untuk menerapkan metode Mahalanobis Distance terhadap sistem pengenalan wajah. Kegiatan penelitian ini berfokus untuk membandingkan tingkat akurasi dan lamanya waktu yang dibutuhkan oleh kedua metode tersebut dengan dataset pengujian yang bervariasi, seperti wajah yang tertutup sebagian maupun jarak antara pengguna dengan kamera. Hasil pengujian ini menunjukkan bahwa Mahalanobis Distance mampu meningkatkan keakuratan sistem untuk mengenali wajah seseorang sebesar 42%.

**Keywords:** Sistem pengenalan wajah, Eigenface, Mahalanobis Distance, Euclidean distance, Principal Component Analysis

## 1. PENDAHULUAN

Dalam kurun waktu 50 tahun terakhir, teknologi telah mengalami perkembangan pesat dan menjadi salah satu aspek penting dalam kehidupan manusia. Salah satu contoh bentuk perkembangan teknologi yang dapat diamati dengan seksama ialah perkembangan teknologi pengamanan, seperti penggunaan metode *Personal Identity Number* (PIN) yang umum dijumpai pada mesin-mesin ATM hingga beralih kepada metode yang memanfaatkan biometrik sebagai kontrol akses. Hal ini dikarenakan metode PIN telah berulang kali dibuktikan mudah untuk “dipecahkan” dengan menggunakan perangkat keras atau sejumlah aplikasi, sehingga teknologi kontrol akses berbasis biometrik semakin diminati. Biometrik adalah teknik otentikasi yang mengandalkan karakteristik fisik terukur yang dapat diperiksa secara otomatis[1]. Salah satu teknik biometrik yang saat ini digunakan adalah sistem pengenalan wajah.

Sistem pengenalan wajah dianggap sebagai salah satu teknologi biometrik yang banyak digunakan secara global dengan

kemampuan mengidentifikasi seseorang dari citra digital atau frame video dari sumber video. Pengenalan wajah juga dianggap sebagai salah satu sistem kontrol akses tercepat karena tidak perlu membutuhkan prosedur yang rumit untuk melakukan proses pengenalan wajah terhadap dataset yang tersimpan. Sistem pengenalan wajah juga berfungsi sebagai alat identifikasi pasif karena pengguna tidak harus melakukan aktivitas tertentu. Bandingkan dengan salah satu teknik biometrik seperti identifikasi sidik jari, pengguna harus meletakkan jarinya pada pembaca agar dikenali oleh sistem.

Secara umum pengenalan wajah dapat dibagi menjadi tiga proses utama, yaitu deteksi wajah, ekstraksi fitur, dan identifikasi wajah. Dari ketiga proses tersebut, penelitian ini berfokus pada proses identifikasi wajah dengan membandingkan penggunaan *distance classifier* yang berbeda, yaitu menggunakan metode *Euclidean distance* dan *Mahalanobis Distance*. Keduanya tetap menggunakan *Principal Component Analysis* (PCA) untuk

mengurangi data dan ekstraksi ciri dari pendekatan berbasis tampilan.

## 2. TINJAUAN PUSTAKA

Adapun beberapa kajian penelitian sejenis yang digunakan sebagai acuan dalam kegiatan penelitian ini, diantaranya:

a. Mohit P. Gawande dan Prof. Dhiraj G. Agrawal (2014) [2]

Penelitian ini berjudul: "Face recognition using PCA and different distance classifiers". Penelitian ini membahas tentang penggunaan PCA dengan *distance classifier* yang berbeda untuk melihat perbedaan yang dihasilkan terhadap tingkat kesuksesan pencocokan dan ketahanan algoritma pengenalan wajah. Beberapa *distance classifier* yang diujikan pada penelitian ini antara lain *City-Block*, *Euclidean*, *Squared Euclidean*, dan *Square Chebyshev*. Hasil dari penelitian ini menunjukkan bahwa PCA mendapatkan hasil yang paling akurat dengan menggunakan *Euclidean distance* bila dibandingkan dengan metode *distance classifier* lainnya.

b. Mini Singh Ahuja dan Sumit Chhabra (2011)[3]

Penelitian ini berjudul: "Effect of Distance Measures in PCA Based Face Recognition". Penelitian ini membahas mengenai efek dari penggunaan berbagai *distance classifiers* terhadap tingkat keakuratan pengenalan wajah. Metode *distance classifiers* yang diujikan pada penelitian ini antara lain *City-Block*, *Squared Euclidean*, *Angle negative*, dan *Mahalanobis Distance*. Penelitian ini menggunakan dataset dengan jumlah 1,196 gambar dengan ketentuan bahwa setiap individu akan memiliki 5 gambar yang berbeda. Hasil dari kegiatan penelitian ini menunjukkan bahwa PCA dengan menggunakan *Mahalanobis Distance* unggul bila dibandingkan dengan *distance classifier* lainnya. Namun, tidak dijelaskan dengan detail faktor apa saja yang berperan dalam penentuan tingkat kesuksesan pengenalan wajah.

c. Lauren Barker (2011)[4]

Penelitian ini berjudul: "Face Recognition". Penelitian ini berfokus pada ekstraksi ciri yang bertujuan untuk

menentukan cara optimal dalam melakukan ekstraksi ciri diskriminasi dari satu set wajah. Penelitian ini dimulai dengan menggunakan *Euclidean distance* kemudian ditingkatkan dengan menerapkan konsep *Linear Discriminant Analysis* dan metode *Kernel*. Hal ini bertujuan untuk membangun solusi yang lebih kuat terhadap masalah ini.

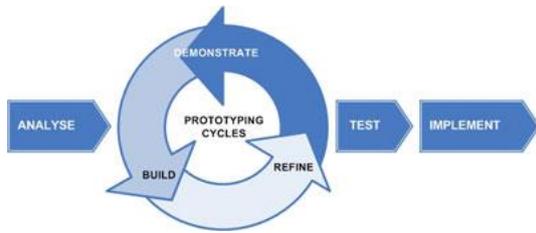
d. Yangfeng Ji, Tong Lin, Hongbin Zha[5]

Penelitian ini berjudul: "Mahalanobis Distance Based Non-Negative Sparse Representation for Face Recognition". Penelitian ini berfokus pada pengujian algoritma klasifikasi berbasis *sparse representation*. Hasil dari penelitian ini menunjukkan bahwa penggabungan *Mahalanobis Distance* dengan algoritma klasifikasi berbasis *sparse representation* memiliki hasil yang lebih akurat jika dibandingkan dengan menggunakan metode lainnya.

## 3. METODOLOGI PENELITIAN

Penelitian ini menggunakan Rapid Application Development (RAD) dalam mengembangkan aplikasi yang direncanakan. RAD merupakan metode pengembangan software yang lebih cepat dibandingkan metode pengembangan software tradisional karena RAD menerapkan penggunaan teknik yang berulang seperti yang dinyatakan oleh Murch[6]. Selain itu, James Martin berpendapat bahwa RAD tidak hanya membutuhkan waktu yang singkat dalam mengembangkan sebuah software tetapi juga memberikan kualitas end-product yang lebih baik bila dibandingkan metodologi tradisional seperti metode Waterfall[7].

Menurut Watkins[8], RAD memiliki kerangka waktu yang tetap dalam mengembangkan sebuah aplikasi ataupun sistem, mulai dari 30, 60, hingga 90 hari. Dengan demikian, apabila peneliti ingin mengembangkan suatu software yang ternyata waktu penyelesaiannya melebihi kerangka waktu telama yang dimiliki RAD maka proyek tersebut dipastikan tidak akan selesai dengan hasil yang diharapkan.



Gambar 1 Bagan Metode RAD

Gambar 1 menunjukkan fase-fase pengembangan yang terdapat dalam RAD, antara lain:

a) Analisa

Pada fase ini peneliti mengumpulkan data yang diperlukan dalam mengembangkan aplikasi yang dirancangkan. Observasi serta studi literatur merupakan instrumen penelitian yang dipilih untuk digunakan dalam mengumpulkan data. Dari kegiatan observasi diperoleh data yang diperlukan dalam membantu proses kegiatan analisa perbandingan kedua metode *distance classifier*. Setelah data-data telah dikumpulkan, kemudian dilakukan analisa terkait waktu yang dibutuhkan agar penelitian ini dapat diselesaikan sesuai dengan target waktu yang telah ditentukan.

b) Siklus Prototyping

Pada fase ini, peneliti mulai mengembangkan aplikasi secara bertahap berdasarkan hasil analisa terhadap data-data yang telah dikumpulkan. Proses pengembangan dimulai dengan pemodelan terhadap fitur yang didukung oleh aplikasi yang dirancangkan menggunakan diagram use case maupun diagram aktivitas. Selanjutnya, dengan menggunakan sistem yang telah ada sebagai pedoman dasar dalam merancang aplikasi yang diinginkan, peneliti kemudian mengubah atau menambahkan beberapa fitur baru ke dalam sistem yang akan menghasilkan sebuah prototipe. Dengan menggunakan prototipe sebagai acuan dalam pengembangan aplikasi proses pengembangan dapat berjalan lebih singkat. Hal ini dikarenakan peneliti hanya melakukan perubahan maupun perbaikan pada prototipe sehingga tidak perlu merubah rancangan desain aplikasi secara keseluruhan. Setelah peneliti selesai membangun dan menyempurnakan prototipe, maka selanjutnya prototipe tersebut akan didemonstrasikan. Pada tahapan ini, prototipe akan diuji oleh beberapa pengguna, kemudian pengguna

memberikan evaluasi terhadap prototipe, dengan memberikan beberapa masukan berupa perbaikan-perbaikan kecil maupun besar. Namun, jika pengguna berpendapat bahwa prototipe tidak memenuhi ekspektasi calon pengguna, maka peneliti harus mengulang proses mengembangkan prototipe.

c) Pengujian

Setelah prototipe telah memenuhi kriteria yang telah ditetapkan, maka prototipe akan masuk dalam fase pengujian. Pada fase ini, prototipe akan diuji dengan serangkaian uji kelayakan yang telah ditentukan oleh peneliti untuk mengetahui apakah masih terdapat kecacatan yang mungkin tidak nampak pada saat proses pengembangan. Hasil dari setiap uji kelayakan akan dicatat dan dirangkum sehingga dapat digunakan untuk pengembangan prototipe yang serupa dikemudian hari.

d) Implementasi

Fase implementasi merupakan fase terakhir dari proses pengembangan sebuah aplikasi menggunakan metode RAD. Setelah prototipe berhasil melewati fase pengujian maka prototipe telah siap sepenuhnya untuk digunakan atau dioperasikan.

## 4. HASIL DAN PEMBAHASAN

### 4.1 Pendekatan Deteksi Wajah

Beberapa metode pendekatan deteksi wajah yang diajukan dalam penelitian ini antara lain:

1. *Appearance-Based method*: umumnya dalam metode pencocokan template, template tersebut telah ditentukan sebelumnya oleh para ahli. Namun, pada *Appearance-Based method* template dipelajari dari contoh gambar menggunakan analisis statistik atau pembelajaran mesin untuk mendapatkan karakteristik yang relevan dari sebuah wajah.
2. *Feature-Based method*: fitur invariant dari sebuah wajah digunakan untuk mendeteksi tekstur dan warna kulit. Namun fitur wajah tersebut dapat rusak akibat penerangan yang berlebih, noise, dan oklusi.
3. *Knowledge-Based method*: metode ini berasal dari pengetahuan terhadap fitur

wajah yang telah diketahui. Akan tetapi masalah yang dihadapi dengan metode ini adalah kesulitan dalam menerjemahkan pengetahuan manusia terhadap fitur-fitur wajah kedalam sebuah aturan yang tepat dan mudah dimengerti.

Dari ketiga pendekatan metode deteksi wajah tersebut, peneliti memutuskan untuk menggunakan *Appearance-Based method*.

#### 4.2 Perhitungan terhadap Eigenfaces

Dalam penelitian ini, terdapat beberapa tahapan yang diperlukan dalam mempersiapkan dataset yang akan dipakai dalam pengujian, yaitu:

- Menyiapkan data citra wajah 2D yang dapat digambarkan sebagai vektor 1D dengan menggabungkan setiap baris menjadi vektor yang panjang.
- Mengurangi nilai rata-rata; Matrix rata-rata harus dihitung, kemudian dikurangkan dari wajah asli, dan hasilnya disimpan pada variabel  $\Phi_i$

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n$$

$$\Phi_i = \Gamma_i - \Psi$$

- Menghitung matriks co-variant

$$A = \Phi^T \Phi$$

- Menghitung vektor eigen dan nilai eigen dari matriks co-variant.
- Menghitung *eigenfaces*

$$[\Phi]X_i = f_i$$

- Mengklasifikasikan wajah

$$\Omega_k = \Omega_k^T (\Gamma_k - \Psi)$$



Gambar 2 sample eigenfaces dari pengguna

#### 4.3 Euclidean distance

*Euclidean distance* merupakan *distance classifier* yang paling umum digunakan untuk menghitung jarak kedekatan antara citra pembelajaran dengan

objek yang akan diklasifikasikan. Dengan demikian, dapat disimpulkan bahwa *Euclidean distance* menghitung jarak minimum antara citra yang akan diuji dengan citra yang dikenali dari database. Jika jaraknya kecil, maka dapat disimpulkan bahwa citra yang dibandingkan serupa sehingga dapat ditentukan mana gambar yang paling mirip dengan database.

Pendekatan Perhitungan *distance classifiers* dengan menggunakan metode *Euclidean distance* dapat dirumuskan sebagai berikut:

$$d(x, y) = \|x - y\|^2 = \sum_{i=1}^k (x_i - y_i)^2$$

Secara formal bentuk diatas merupakan klasifikasi *nearest neighbor*, karena gambar diklasifikasikan dengan menempatkannya ke label titik terdekat dalam rangkaian pelatihan yang mengukur seluruh jarak pada ruang ciri.

Wajah kemudian ditentukan menjadi milik kelas  $k$  jika jarak kedekatannya kurang dari ambang batas yang telah ditentukan sebelumnya, yaitu:

$$\epsilon_k = \min_{i=1, \dots, g} \{\epsilon_i | \epsilon_i < \theta_i\}$$

Penulis perlu menentukan ambang batas  $\theta_i$  untuk masing-masing  $i = 1, \dots, g$  individu. Hal ini nantinya memberikan jarak maksimum yang diijinkan gambar yang dimaksud.

#### 4.4 Mahalanobis distance

*Mahalanobis distance* didefinisikan sebagai ruang dimana varians sampel di sepanjang dimensi masing-masing adalah satu. Adapun Barker[4] berpendapat bahwa *Mahalanobis distance* merupakan jarak antara dua vektor yang memperhitungkan korelasi kumpulan data. Relasi antara kedua vektor dapat didefinisikan sebagai berikut:

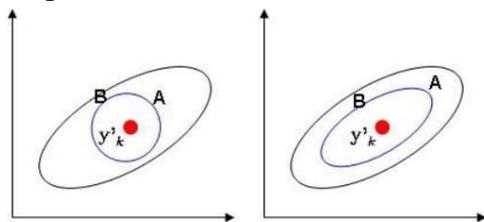
$$m_i = \frac{x_i}{\sigma_i} \quad n_i = \frac{y_i}{\sigma_i}$$

$$d(x, y) = \sqrt{\sum_{i=1}^k (m_i - n_i)^2}$$

#### 4.5 Perbandingan Euclidean dan Mahalanobis distance

Dalam menugaskan gambar yang tidak diketahui ke salah satu dari  $g$  individu yang dikenal, penulis menggunakan pendekatan *nearest neighbor* karena penulis berasumsi semakin dekat vektor ciri  $y_j$  dengan vektor ciri rata-rata kelas  $y'_k$ , semakin mirip gambar yang tidak diketahui dari individu  $k$ th. Namun, hal tersebut juga berguna untuk mengetahui apakah himpunan vektor ciri yang termasuk dalam kelas tersebut dan rentang besar atau kecil. Hal inilah yang memperlihatkan keunggulan *Mahalanobis distance* di atas *Euclidean*, karena memperhitungkan variabilitas kumpulan data.

Jarak Mahalanobis memberikan persamaan sebuah ellipsoid dimana semua titik pada ellipsoid ini memiliki *Mahalanobis distance* yang sama dari pusat, yang merupakan kelas rata-rata  $y'_k$ . Sumbu ellipsoid memperhitungkan korelasi di ruang fitur. Arah di mana ellipsoid memiliki poros kecil berarti bahwa vektor fitur dalam arah ini harus lebih dekat ke pusat untuk diklasifikasikan sebagai kelas ini, sedangkan arah di mana ellipsoid memiliki sumbu besar berarti vektor ciri diperbolehkan untuk jauh dari pusat. Jadi, kemungkinan vektor ciri yang termasuk kelas bergantung pada jarak dan pusat. Jika dibandingkan dengan jarak *Euclidean* yang membentuk bola, di mana semua vektor fitur pada bidang ini memiliki jarak *Euclidean* yang sama dari pusat. Jadi ini tidak memperhitungkan korelasi dari ruang data.



Gambar 3 Diagram sebelah kiri menunjukkan pola *Euclidean distance*, sementara diagram sebelah kanan menunjukkan pola *Mahalanobis distance*

#### 4.6 Implementasi Euclidean dan Mahalanobis distance

Proses pengujian terhadap kedua metode tersebut adalah dengan mengimplementasikan kedalam sebuah

sistem pengenalan wajah dari penelitian sebelumnya.

Jumlah training set yang dipakai dalam penelitian ini berjumlah 2,808 citra wajah. Setiap citra memiliki ukuran 180 x 220 dan telah mengalami proses perubahan dari citra berwarna menjadi citra *greyscale*. Setiap citra akan diubah dari sebuah array menjadi sebuah vector dengan panjang  $N = 180 \times 220$  dengan menggabungkan setiap baris dari array gambar.

Setelah persiapan tersebut dilakukan maka selanjutnya peneliti melakukan 5 tahapan proses sebagai berikut:

1. Menghitung dan mempertimbangkan perhitungan citra rata-rata wajah dari set pelatihan.
2. Menghitung vector eigen dan nilai eigen dari matriks kovariansi.
3. Menentukan jumlah eigenface yang saat ini dimiliki.
4. Memproyeksikan setiap gambar dalam pelatihan yang ditetapkan ke dalam ruang wajah.
5. Memperhitungkan satu gambar dalam pengujian yang ditetapkan pada saat itu.

Setelah melakukan kelima tahapan tersebut, maka diperoleh hasil pengujian metode *Euclidean* dan *Mahalanobis distance* terhadap sistem pengenalan wajah sebagai berikut:

Tabel 1 Data akurasi wajah dengan *Euclidean distance*

No.	Subjek	Image	Konfiden	Akurasi
1	Pria1		0.75	76.34
2	Wanita1		0.78	81.1
3	Pria2		0.64	73.5
4	Pria3		0.73	79.7

Tabel 2 Data akurasi wajah dengan *Mahalanobis distance*

No.	Subjek	Image	Konfiden	Akurasi
1	Pria1		0.88	94.13
2	Wanita1		0.89	95.17
3	Pria2		0.91	91.22
4	Pria3		0.85	96.94

## 5. KESIMPULAN

Dari hasil penelitian yang telah dilakukan, penulis menyimpulkan sebagai berikut:

1. *Mahalanobis distance* lebih akurat dalam memberikan hasil pengenalan wajah dibandingkan *Euclidean distance*.
2. Tingkat akurasi pengenalan wajah dengan metode *Mahalanobis distance* rata-rata mencapai 94.35%.
3. Para pengguna menyatakan bahwa aplikasi tergolong mudah untuk dioperasikan.
4. Aplikasi dirasa cukup akurat dan cepat dalam mengenali wajah pengguna.

## 6. DAFTAR PUSTAKA

- [1] Webopedia, 2017. *Definition of biometric template* [online]. Dapat diakses pada [http://www.webopedia.com/TERM/B/biometric\\_template.html](http://www.webopedia.com/TERM/B/biometric_template.html)
- [2] Gawande, Mohit P. dan Dhiraj G. Agrawal. 2014. *Face recognition using PCA and different distance classifiers* [online]. Dapat diakses pada <http://www.iosrjournals.org/iosr-jece/papers/Vol9-Issue1/Version-6/A09160105.pdf>
- [3] Ahuja, Mini S. dan Sumit Chhabra. 2011. *Effect of Distance Measures in PCA Based Face Recognition* [online]. Dapat diakses pada <http://www.ijecbs.com/July2011/35.pdf>
- [4] Barker, L. 2011. *Face Recognition* [online]. Dapat diakses pada [http://www.maths.dur.ac.uk/users/kasper.peeters/pdf/face\\_recognition/reports/Barker.pdf](http://www.maths.dur.ac.uk/users/kasper.peeters/pdf/face_recognition/reports/Barker.pdf)
- [5] Ji, Yangfeng, Tong Lin, Hongbin Zha. no date. *Mahalanobis Distance Based Non-Negative Sparse Representation for Face Recognition* [online]. Dapat diakses pada [http://www.cis.pku.edu.cn/faculty/vision/lintong/papers/ICMLA09\\_Ji.pdf](http://www.cis.pku.edu.cn/faculty/vision/lintong/papers/ICMLA09_Ji.pdf)
- [6] Murch, R. 2012, *Project Management: Best Practices for IT Professionals*, Ed.1, Prentice Hall, New Jersey.
- [7] Martin, J. 1991, *Rapid Application Development*, Macmillan Col Div, New York.
- [8] Watkins, J. 2009, *Agile Testing: How To Succeed in an Extreme Testing Environment*, Cambirdge University Press, Cambridge