



**SISTEM PENGENALAN WAJAH (FACE RECOGNITION)
MENGUNAKAN METODE PCA (PRINCIPAL COMPONENT
ANALYSIS)**

Oleh :

**NAMA : JONATHAN PURNAMA . H (A11.2017.10193)
VANNYA MAHESWARI (A11.2017.10407)
PROGRAM STUDI : TEKNIK INFORMATIKA – S1**

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO
SEMARANG
2020**

ABSTRAK

Identitas setiap orang dapat ditentukan berdasarkan mata, sidik jari dan juga wajah. Wajah merupakan salah satu yang dapat menentukan identitas seseorang karena wajah setiap orang mempunyai karakteristik yang berbeda dan unik. Identifikasi wajah sering digunakan dalam bidang akses untuk keamanan, absensi, dan juga digunakan untuk sistem pencarian seseorang. Sistem pengenalan wajah atau *face recognition* diciptakan menggunakan metode *Principal Component Analysis* (PCA) sebagai pengoptimalan dalam mereduksi dimensi. PCA merupakan algoritma reduksi dimensi yang mampu menghasilkan komponen wajah. Sistem pengenalan wajah (*face recognition*) ini melakukan identifikasi wajah seseorang berdasarkan gambar dari sebuah database yang kemudian diproses menjadi data latih. Setelah terbentuk data latih, sistem pengenalan wajah ini akan melakukan perbandingan antara data uji dengan data latih. Pada hasil akhir dari sistem pengenalan wajah ini akan menampilkan gambar data latih, gambar uji dan nama orang yang diuji. Jika data latih dengan data uji menghasilkan hasil yang berbeda, maka dinyatakan gagal. Begitupula sebaliknya jika hasil dari data latih dengan data uji sesuai, maka dinyatakan berhasil.

Kata Kunci : Wajah, *Face Recognition*, *Principal Component Analysis* (PCA)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Wajah merupakan sebuah identitas bagi semua orang. Hal ini dikarenakan wajah pada setiap orang umumnya berbeda satu dengan yang lainnya. Meskipun kembar identik sekalipun, jika diperhatikan dengan seksama akan ada perbedaan satu dengan yang lainnya. Hal ini dikarenakan wajah yang dimiliki oleh setiap orang mempunyai karakteristik masing – masing dan juga mempunyai keunikannya tersendiri. Keunikan itu sendiri berasal dari tipe bentuk wajah yang berbeda - beda seperti kotak, oval, bulat, dan lain-lain. Kemudian tata letak alis, jarak antara mata satu dan lainnya, bentuk hidung, bentuk bibir, hingga bentuk dagu sekalipun pasti mempunyai perbedaan dan keunikannya masing – masing.

Kemajuan teknologi tentunya tidak dapat kita hindari lagi di masa sekarang yang dimana semakin pesatnya ilmu pengetahuan yang dimiliki manusia dengan dibantu nya manusia dengan memperoleh informasi dan pengetahuan secara cepat akan semakin pesat juga suatu kemajuan teknologi. Kemajuan teknologi sendiri diciptakan untuk memberikan banyak kemudahan untuk kehidupan manusia [1], dimana kemajuan teknologi itu sendiri tercipta berdasarkan pikiran manusia yang berkembang selain itu juga berdasarkan permasalahan-permasalahan yang ada di kehidupan manusia. Masalah yang sering menyita perhatian adalah dalam pendeteksian atau pengidentifikasian. Sebagai contoh sistem identifikasi yang banyak dikembangkan saat ini adalah seperti sistem identifikasi *biometrik* atau menggunakan informasi biologis pada tubuh manusia seperti sidik jari, retina mata, suara maupun wajah manusia.[2] Salah satu topik yaitu pengenalan wajah merupakan topik yang banyak dikembangkan dan memiliki tingkat akurasi yang tinggi.[3]

Wajah manusia sendiri merupakan salah satu permasalahan yang menarik dan menjadi perhatian selama bertahun-tahun. Hal ini dikarenakan setiap wajah manusia mengandung banyak sekali informasi, topik yang cukup menarik perhatian, serta dipelajari secara intensif. Beberapa diantara penelitian yang

mempelajari informasi dari wajah manusia yaitu mengenai pengenalan wajah. Pengenalan wajah diterapkan dalam suatu sistem dibuat dengan meniru kinerja otak manusia yang mampu mencocokkan wajah seseorang dengan sebuah foto atau media yang lainnya dengan melakukan ekstraksi .

Sistem pengenalan wajah (*Face Recognition*) merupakan salah satu perkembangan teknologi yang sangat berkembang dengan pesat dan sangat dibutuhkan untuk berbagai kebutuhan yang mempermudah kegiatan. Pengenalan wajah sangat berperan penting yang dibutuhkan oleh berbagai aplikasi ataupun sistem berbasis komputer. Sistem pengenalan wajah itu sendiri banyak dimanfaatkan pada biometriks yang sering digunakan untuk mengidentifikasi seseorang dan dapat diimplementasikan untuk pencarian seseorang atau mengidentifikasi seseorang, dan juga sebagai mesin absensi. Dikarenakan Wajah memiliki karakteristik unik setiap masing - masing individu, sering kali Sistem pengenalan wajah ini juga diterapkan pada bidang keamanan (*security system*)[4] seperti izin akses mengontrol sesuatu atau akses masuk ke sebuah ruangan penting, pengawasan lokasi. Sistem identifikasi yang dikembangkan itu sendiri diambil berdasarkan perbedaan ciri wajah dari seseorang berbasis biometriks yang mempunyai keakurasian yang tergolong tinggi. Identifikasi tersebut mampu digunakan sebagai sistem pencarian wajah seseorang berdasarkan sebuah gambar yang berisikan wajah seseorang yang hendak dicari.

Sistem pengenalan wajah merupakan sistem pengenalan pada gambar yang tidak menggunakan bitmap pixel secara langsung namun tetapi bekerja dengan mengubah image wajah menjadi *domain feature*. Gambar dipresentasikan kedalam bentuk fitur yang telah di ekstraksi yang lebih *compact* dan kemudian digunakan untuk pengenalan, dan demikian hal tersebut dibutuhkan untuk menghemat komputasi.[4] Cara kerja dari sistem pendeteksi wajah ini membandingkan wajah yang ada pada database atau knowledgebase dengan data yang ada pada data latih. Jika hasil perbandingan antara kedua data tersebut cocok, maka indentifikasi wajah dinyatakan berhasil. Hal tersebut juga berlaku sebaliknya, jika hasil perbandingan antara kedua data tersebut tidak cocok, maka hasil identifikasi wajah tersebut dinyatakan gagal atau tidak teridentifikasi. [5]

Sistem pengenalan wajah memiliki banyak metode dan algoritma yang cukup banyak, di antaranya yang cukup populer adalah algoritma PCA atau biasa kita sebut *Principal Component Analysis*. PCA merupakan suatu teknik linear yang berfungsi untuk memproyeksikan data berupa vektor yang mempunyai dimensi tinggi ke vektor yang mempunyai dimensi lebih rendah, jadi PCA bekerja dengan cara mengidentifikasi pola – pola pada data lalu mengekspresikan data tersebut ke bentuk – bentuk yang lain agar dapat disimpulkan kesamaan atau perbedaan antar data ataupun pola. [2]

Tujuan penggunaan algoritma PCA (*Principal Component Analysis*) sendiri adalah mereduksi dimensi yang besar pada ruang data menjadi suatu dimesi yang lebih kecil atau bisa kita permudah mendeskripsikan data lebih sederhana. Dimensi pada gambar berupa matriks $M \times N$ dan mempunyai tiga layer RGB. Pose wajah yang ada dengan citra RGB kemudian di convert ke citra grayscale yang kemudian matriks tersebut akan di reduksi.[2] Setelah itu penghitungan tingkat kesamaan (*similarity degree*) antara citra tes dengan citra uji menggunakan metode jarak. PCA itu sendiri adalah salah satu algoritma yang digunakan untuk pengenalan berdasarkan *appearance based*. PCA juga merupakan salah satu algoritma reduksi dimensi yang dapat menghasilkan komponen wajah yaitu *eigenface*. *Eigenface* merupakan salah satu metode pengenalan wajah yang cukup mudah dan sederhana untuk diimplementasikan. *Eigenface* ini merupakan salah satu algoritma pengenalan wajah yang berdasarkan kepada algoritma PCA (*Principal Component Analysis*). Proses pendeteksian wajah dilakukan mulai dari pengambilan gambar kemudian melakukan pendeteksian kulit dan pendeteksian wajah untuk selanjutnya melakukan *face recognition*. Sistem ini melakukan pendekatan pada warna kulit menggunakan warna YcbCr untuk menghasilkan pendeteksian wajah yang lebih akurat.[3] Pemrosesan gambar wajah menggunakan kombinasi operator morfologi dan *elliptical shape* dari sebuah wajah untuk selanjutnya proses segmentasi wajah. Kemudian seleksi atau pemilihan *eigenvector* dilakukan untuk mengetahui *eigenvector* mana yang sesuai dengan kandungan informasi yang lebih tinggi.

1.2 Rumusan Masalah

1. Apakah metode *Principal Component Analysis* (PCA) mampu mengidentifikasi wajah seseorang dari suatu gambar dan membandingkannya dengan gambar yang menjadi data latih pada database?
2. Seberapa akuratkah metode *Principal Component Analysis* (PCA) untuk mendeteksi wajah seseorang?

1.3 Batasan Masalah

1. Program dibuat menggunakan Matlab
2. Data gambar atau citra yang digunakan untuk diuji harus menggunakan citra gambar RGB atau hitam putih
3. Citra yang digunakan mempunyai format .tiff
4. Citra wajah pada database, data latih dan data latih mempunyai ukuran yang sama
5. Data uji langsung di bandingkan dengan data latih pada suatu database
6. Proses training menggunakan metode *Principal Component Analysis* (PCA)
7. Program dibuat menggunakan Matlab

1.4 Tujuan

Adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Untuk membuat software yang dapat mengidentifikasi wajah dan menguji metode *Principal Component Analysis* (PCA) mampu mengidentifikasi wajah seseorang
2. Untuk mengetahui seberapa akurat metode *Principal Component Analysis* (PCA) dalam mengidentifikasi wajah seseorang

1.5 Manfaat Penelitian

1.5.1. Manfaat Bagi Universitas Dian Nuswantoro

1. Menambah kajian pustaka mengenai penelitian yang terkait.
2. Menambah jurnal online di kampus yang telah ada.

1.5.2. Manfaat Bagi Pembaca

1. Dapat menambah wawasan dan pengetahuan pembaca mengenai identifikasi wajah seseorang menggunakan metode *Principal Component Analysis* (PCA).
2. Dapat dijadikan sebagai referensi dan dapat diimplementasikan dalam kehidupan nyata misalnya sebagai sistem pencarian seseorang.

1.5.3. Manfaat Bagi Penulis

1. Menambah pengetahuan mengenai sistem pengenalan wajah menggunakan metode *Principal Component Analysis* (PCA)
2. Dapat menambah pengalaman berdasarkan penelitian tersebut.
3. Dapat menguji metode *Principal Component Analysis* (PCA) yang mampu mengidentifikasi wajah seseorang.

BAB II

TINJAUAN PUSTAKA

2.1. Tinjauan Studi

Menurut hasil pencarian di google scholar, topik face recognition merupakan topik yang hangat dan sangat berkembang pesat di masyarakat. Hal ini ditandai dengan banyaknya topik pengenalan wajah yang diangkat dan dipublikasikan. Adapun beberapa topik pengenalan wajah yang menggunakan metode PCA (*Principal Component Analysis*), dan metode – metode lainnya. Berikut adalah beberapa journal yang akan di review:

Perani Rosyani (2017) melakukan penerapan metode PCA (*Principal Component Analysis*) untuk membuat sistem pengenalan wajah dengan bantuan *canberra distance*. Penelitian dilakukan dengan melakukan pengambilan 10 pose wajah yang berbeda lalu dilakukan penerapan metode PCA untuk pengoptimalan dalam mereduksi dimensi. Setelah citra berhasil di ekstraksi maka akan di hitung tingkat kemiripan atau tingkat kesamaan (*similarity degree*) antara data uji dengan data latih menggunakan metode jarak yaitu *Canberra distance*. Hasil penelitian yang didapatkan dalam rata-rata nilai untuk *Canberra distance* sebesar 77,59.

Dian Esti Pratiwi, Agus Harjoko (2016), mereka menerapkan metode PCA (*Principal Component Analysis*) yang dianggap cukup mudah diimplementasikan pada sistem pengenalan wajah. Citra diolah untuk mencari nilai *eigenvector*, *eigenvalue* dan *average image* yang diproyeksikan ke dalam subruang PCA. Lalu akan dicari perbandingan terkecil proyeksi PCA antara file database dengan data inputan dengan menggunakan metode *nearest neighbor*. Pada penelitian ini didapatkan hasil tingkat keberhasilan 82,81% dengan melakukan pengujian menggunakan ekspresi senyum dan citra tanpa ekspresi pada 8 orang dan 16 buah citra wajah. Faktor yang memengaruhi keberhasilan pengenalan adalah pencahayaan pada wajah, jarak wajah terhadap webcam dan banyaknya gambar wajah yang tersimpan pada database.

Abdu Rakhman Syakhala, Diyah Puspitaningrum, Endina Putri Purwandari (2017) pada penelitian mereka melakukan perbandingan antara metode PCA

(*Principal Component Analysis*) dengan metode HMM (*Hidden Markov Model*) dalam pengenalan identitas seseorang melalui pengenalan wajah. Tujuan dari penelitian yang dilakukan adalah mengetahui perbedaan, kekurangan, dan kelebihan. Setelah dilakukan pengujian terhadap kedua metode disimpulkan bahwa metode PCA unggul lebih akurat yaitu mendapatkan akurasi 86,6% dibandingkan metode HMM yang hanya mendapat akurasi 77,7% dengan maksimal iterasi 2000 dan toleransi 0.1%.

Salamun, Firman Wazir (2016) dalam penelitian yang dilakukan mereka memilih metode PCA atau (*Principal Component Analysis*) sebagai metode untuk penelitian dalam pembuatan rancang bangun sistem pengenalan wajah. PCA menghitung rata – rata tiap *flatvector pixel* dari gambar – gambar yang telah tersimpan di database lalu akan didapatkan nilai *eigenface* setiap gambar. Dari setiap nilai *eigenface* inilah yang akan dicari nilai terdekatnya untuk menentukan gambar citra wajah yang akan dikenali. Hasil pengujian menunjukkan tingkat keberhasilan yang cukup tinggi yaitu 82,27% dari jumlah data wajah sebanyak 130 gambar.

Dodit Suprianto, Rini Nur Hasanah, Purnomo Budi Santosa(2016) pada penelitian ini, mereka membuat sebuah sistem pengenalan wajah yang berbasis *real-time* dengan menggunakan metode – metode seperti *Adaboost*, *Eigenface* PCA & MySQL. Dimana setiap metode diatas memiliki fungsi tersendiri masing – masing, Pendeteksian wajah menggunakan metode *Adaboost* sedangkan metode *eigenface* PCA melakukan pengenalan tiap wajah yang telah terdektesi dan tahap akhir adalah pencocokkan wajah yang telah tersimpan di database MySQL. Menurut penelitian yang dilakukan didapatkan tingkat keberhasilan sebesar 80% dalam mengidentifikasi wajah.

Tabel 2.1. State of The Art

No	Peneliti	Tahun	Judul	Masalah	Metode	Hasil
2	Perani Rosyani	2017	Pengenalan Wajah Menggunakan Metode <i>Principal Component Analysis</i> (PCA) dan <i>Canberra Distance</i>	Membuat sistem pengenalan wajah dengan berdasarkan analogi 'memori otak manusia' yang mampu mengenali seseorang melalui memori gambaran wajah orang yang telah dilihat sebelumnya dan diaplikasikan menggunakan metode <i>Principal Component Analysis</i> (PCA)	1. Metode <i>Principal Component Analysis</i> (PCA). 2. Metode jarak (<i>Canberra Distance</i>)	1. Mengekstraksi gambar menggunakan PCA cukup rumit karena harus mentranspose matriks ukuran $m \times n$ menjadi $1 \times n$. 2. Tingkat kesamaan yang dihasilkan setelah proses reduksi dan ekstraksi menggunakan PCA didapatkan rata-rata nilai untuk <i>Canberra distance</i> adalah 77,59
3	1. Dian Esti Pratiwi 2. Agus Harjoko	2016	Implementasi Pengenalan Wajah Menggunakan PCA (<i>Principal Component Analysis</i>)	Mengimplementasi kan sistem identifikasi wajah yang merupakan salah satu informasi biologis dan memiliki tingkat keakuratan tinggi sehingga dapat dijadikan sebagai sistem pencarian seseorang di dalam sebuah gambar berisi wajah	<i>Principal Component Analysis</i> (PCA)	1. Hasil pengujian menunjukkan dari 16 sampel yang diujikan sebanyak 64 kali pengenalan menghasilkan 53 pengenalan benar dan 11 pengenalan salah. 2. Wajah sejumlah 160 wajah dari delapan orang memerlukan waktu kurang lebih dua menit pada setiap satu kali proses pengenalan. 3. Prosentase keberhasilan pengenalan wajah pada sampel yang dilakukan adalah 82,81%. 4. Faktor yang mempengaruhi keberhasilan pengenalan yaitu kesesuaian ekspresi antara data latih dengan data uji, faktor pencahayaan (semakin terang, semakin baik), dan jarak wajah dengan <i>webcam</i> adalah 25 cm.

No	Peneliti	Tahun	Judul	Masalah	Metode	Hasil
8	1. Abdu Rakhman Syakhaka 2. Diyah Puspitaningrum 3. Endina Putri Purwandani	2017	Perbandingan Metode <i>Principal Component Analysis</i> (PCA) dengan metode <i>Hidden Markov Model</i> (HMM) dalam Pengenalan Identitas Seseorang Melalui Wajah	Mengetahui perbedaan, kekurangan, kelebihan, dan perbandingan dari metode PCA dan metode HMM sebagai pengenalan identitas melalui wajah	1. <i>Principal Component Analysis</i> (PCA) 2. <i>Hidden Markov Model</i> (HMM)	1. Metode <i>Principal Component Analysis</i> (PCA) dan Metode <i>Hidden Markov Model</i> (HMM) berhasil melakukan pengenalan wajah seseorang dengan hasil PCA secara umum lebih baik daripada HMM. 2. Pengenalan wajah terbaik diperoleh dari citra wajah manusia tanpa background dengan tingkat akurasi tertinggi sebesar 86,6% pada PCA sedangkan HMM sebesar 77,7% dengan maksimum iterasinya 2000 dan toleransi 0,1 3. Metode PCA lebih cepat dalam pengenalan wajah dari 1-1,5 detik sedangkan pada metode HMM 2-7,5 detik

No	Peneliti	Tahun	Judul	Masalah	Metode	Hasil
9	1. Salamun 2. Firman Wazir	2016	Rancang Bangun Sistem Pengenalan Wajah dengan Metode <i>Principal Component Analysis</i>	<p>1. Bagaimana dapat mengenali suatu pola, khususnya pola wajah</p> <p>2. Mengukur tingkat keberhasilan pengenalan wajah dengan menggunakan algoritma <i>Principal Component Analysis</i> (PCA) sesuai dengan batasan yang telah ditentukan</p> <p>3. Menghasilkan aplikasi pengenalan wajah menggunakan metode <i>Principal Component Analysis</i></p> <p>4. Apakah pengenalan wajah dengan PCA dapat mencapai 50%?</p> <p>5. Bagaimana melakukan proses pengenalan citra wajah menggunakan metode PCA?</p>	Principal Component Analysis (PCA)	<p>1. Pengenalan pola wajah dapat dikenali menggunakan nilai <i>pixel</i> dari data wajah</p> <p>2. Proses pengenalan wajah menggunakan pendekatan metode <i>Principal Component Analysis</i> sensitif terhadap perubahan cahaya, jarak, ekspresi wajah, sudut pandang wajah dan perubahan wajah yang terlalu ekstrim. Jika citra yang digunakan sebagai <i>training set</i> maupun sebagai citra input memiliki intensitas cahaya yang berbeda dan tidak berada pada posisi yang sama dengan citra <i>training set</i> maka proses tersebut tidak dapat memberikan hasil yang akurat.</p> <p>3. Metode <i>Principal Component Analysis</i> dapat diimplementasikan untuk pengenalan wajah dengan tingkat akurasi 81%.</p> <p>4. Wajah dapat dikenali apabila nilai jarak antara <i>image</i> dengan <i>image test</i> diatas nilai <i>threshold</i>.</p>

No	Peneliti	Tahun	Judul	Masalah	Metode	Hasil
10	1. Dodit Suprianto 2. Rini Nur Hasanah 3. Purnomo Budi Santosa	2016	Sistem Pengenalan Waja Secara Real-Time dengan <i>Adaboost</i> , <i>Eigenfface</i> PCA & MySQL	1. Mengimple mentasikan metode <i>Adaboost</i> dan Metode <i>Eigenface</i> PCA ke dalam sebuah sistem pengenalan wajah secara <i>real-time</i> . 2. Pengenalan wajah yang diintegrasikan dengan data profil bermanfaat di berbagai sektor, misalnya bidang keamanan, pengawasan, kontrol akses, robotika, intelejen, militer, presensi dan lain-lain	1. Metode <i>Adaboost</i> 2. <i>Eigenface</i> PCA	1. Perancangan dan implementasi pengenalan wajah dengan metode <i>Adaboost</i> dan <i>Eigenface</i> PCA telah berhasil dilakukan dalam penelitian. 2. Rata-rata tingkat keberhasilan pengenalan wajah dengan kedua metode tersebut mencapai 80% pada berbagai kondisi berbeda (jarak objek dengan sensor, pencahayaan, posisi, atribut, dan mimik muka) 3. <i>Adaboost</i> dan <i>Eigenface</i> PCA memiliki kelebihan pada proses kecepatan mengambil keputusan untuk mengenali wajah di kondisi <i>real-time</i> .

2.2. Tinjauan Pustaka

2.2.1. Pengolahan Citra Digital

Pengolahan citra digital merupakan salah satu subjek atau bidang ilmu kejuruan dari teknologi informasi yang sangat terkenal dan menarik di era ini. Proses pengolahan citra digital bertujuan untuk memanipulasi dan menganalisis citra dengan bantuan komputer. Pengolahan citra digital (Digital Image Processing) sendiri adalah sebuah disiplin ilmu yang mempelajari tentang teknik-teknik mengolah citra. Dari kalimat diatas dapat kita artikan bahwa citra yang dimaksud yaitu adalah gambar diam (foto) maupun gambar bergerak (yang berasal dari webcam maupun berupa

video yang di input). Sedangkan digital disini mempunyai maksud bahwa pengolahan citra/gambar dilakukan secara digital menggunakan komputer.[6]

Pada Pengolahan Citra Digital, citra yang bersifat data kontinyu yang merupakan hasil kumpulan intensitas cahaya pada bidang dua dimensi maka harus diubah menjadi data diskrit yang bersifat numerik dengan melakukan perhitungan matematis agar dapat dilakukan pengolahan [7]. Sebuah citra dapat diwakili dengan sebuah matriks 2 dimensi yang terdiri dari baris dan kolom. Dengan matriks inilah perhitungan dan teknik – teknik pengolahan citra dapat dilakukan.

2.2.2. Pengenalan Wajah

Wajah adalah salah satu ukuran fisiologis atau bagian tubuh yang paling mudah dan sering digunakan untuk membedakan identitas tiap individu antar satu dengan yang lainnya. Wajah juga memegang peranan penting selain menunjukkan identitas seperti menunjukkan emosi, kemampuan manusia dalam mengenali wajah sangatlah hebat. Manusia dapat membedakan atau mengenali banyak wajah bahkan dengan terjadinya perubahan terhadap wajah seseorang seperti karena bertambahnya usia ataupun diberikan aksesoris seperti kacamata dan perubahan gaya rambut manusia tetap dapat mengenalinya [8]. Hal demikian membuat sebuah sistem atau aplikasi pengenalan wajah menjadi hal yang sungguh menarik dalam bidang ilmu teknik informasi karena hal ini dapat di aplikasikan di berbagai bidang lainnya seperti keamanan atau security ataupun untuk hiburan seperti *filter* wajah di berbagai aplikasi sosial media. Sistem pengenalan wajah adalah suatu sistem yang membuat sebuah mesin dapat mengenali wajah seseorang sesuai dengan gambar wajah yang telah dilatih dan disimpan di dalam mesin tersebut. [9] Dalam proses pengenalan wajah, komputer menggunakan informasi mentah dari pixel citra yang dihasilkan melalui camera ataupun melalui gambar yang diinputkan kemudian direpresentasikan dalam metode – metode yang akan digunakan untuk mengolah seperti algoritma PCA (*Principal Component Analysis*), *FaceNet*, dan lain sebagainya.

2.2.3. Principal Component Analysis

Secara sederhana PCA adalah transformasi linear untuk menentukan sistem koordinat yang baru dari sebuah dataset. Teknik PCA ini mereduksi atau mengurangi informasi data yang besar dari sebuah citra wajah tanpa menghilangkan informasi yang

ada pada sebuah citra wajah. Algoritma PCA melakukan penguraian citra wajah kedalam kumpulan fitur karakteristik yang disebut sebagai “Eigenface”. Hal tersebut yang kemudian dikatakan sebagai Principal Component dalam sebuah data trainingset, fitur utama dari sebuah algoritma PCA adalah merekonstruksi dari beberapa citra asli dari trainingset dengan mengkombinasikan eigenface [8].

2.2.3.1. EigenFace

Metode eigenface adalah bagaimana cara meguraikan informasi yang relevan dari sebuah citra wajah, kemudian mengubahnya ke dalam satu set kode yang paling efisien dan membandingkan kode wajah tersebut dengan database berisi beragam wajah yang telah dikodekan secara serupa. *Eigenfaces* PCA digunakan untuk mereduksi dimensi sekumpulan atau ruang gambar sehingga basis atau sistem koordinat yang baru dapat menggambarkan model yang khas dari kumpulan tersebut dengan lebih baik. Model yang diinginkan merupakan sekumpulan wajah yang dilatihkan. Fitur yang baru tersebut akan dibentuk melalui kombinasi linear. Komponen fitur ruang karakter ini tidak akan saling berkolerasi dan akan memaksimalkan perbedaan yang ada pada variabel aslinya [10]. Sekumpulan wajah yang dilatih direpresentasikan dalam sebuah *vektor flat* (gabungan vektor) dan digabungkan menjadi matriks tunggal. Vektor eigen diekstraksi dan disimpan ke dalam database. Data wajah yang dilatih kemudian diproyeksi dalam *feature space* bernama *face space* yang ditentukan oleh vektor eigen. Vektor eigen (v_i) dan nilai eigen (λ_i) dari matrik kovarian adalah

$$Cv_i = \lambda_1 v_1$$

$$Cv_i = \lambda_i I v_i$$

Solusi menentukan nilai eigen

$$\det(C - \lambda_i I) = 0$$

Solusi menentukan vektor eigen

$$(C - \lambda_i I)v_i = 0$$

Menentukan eigenface

$$\mu_i = \sum_{k=1}^M v_{ik} \Phi_k$$

Wajah diklasifikasikan dengan memproyeksikan wajah Γ ke ruang wajah sebagai berikut:

$$\omega k = \mu k^T (\Gamma k - \Psi)$$

Bobot membentuk vektor $\Omega k^T = [\omega_1, \omega_2, \dots, \omega_M]$, yang berisi proyeksi ke setiap vektor eigen. Klasifikasi dilakukan dengan menghitung jarak Ωk dari Ω , di mana Ω merupakan vektor bobot yang mendefinisikan beberapa kelas.

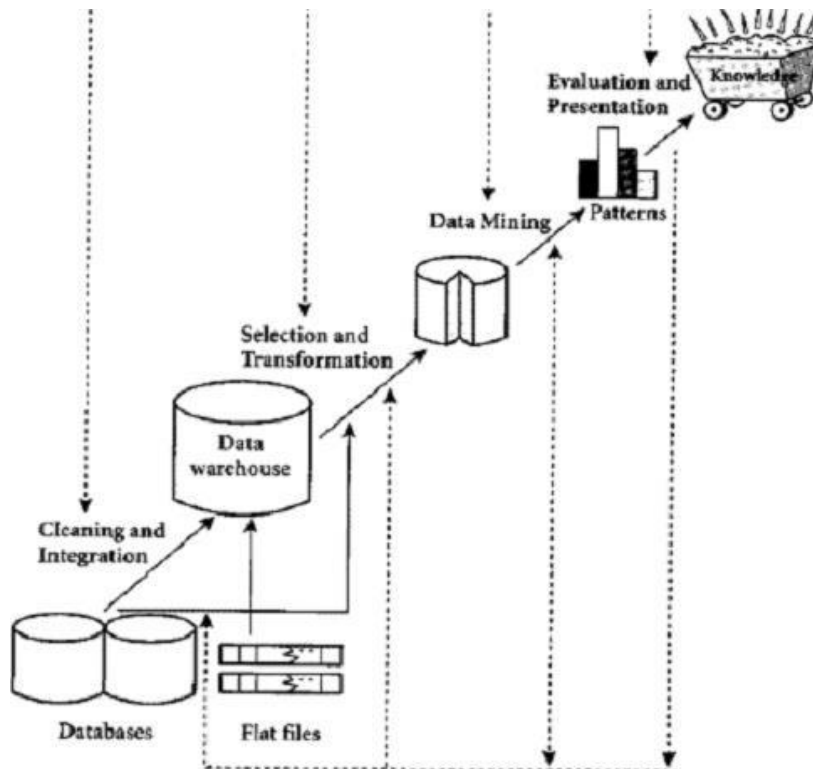
2.2.4. MATLAB

Matlab atau nama lain dari *Matrix Laboratory* merupakan bahasa pemrograman yang hadir dengan fungsi dan karakteristik yang berbeda dengan bahasa pemrograman lain yang sudah ada lebih dahulu seperti Delphi, Basic maupun C++. Matlab merupakan bahasa pemrograman level tinggi yang dikhususkan untuk kebutuhan komputasi teknis, visualisasi dan pemrograman seperti komputasi matematik, analisis data, pengembangan algoritma, simulasi dan pemodelan dan grafik-grafik perhitungan [11]. MATLAB telah berkembang menjadi sebuah environment pemrograman yang canggih yang berisi fungsi-fungsi built-in untuk melakukan tugas pengolahan sinyal, aljabar linier, dan kalkulasi matematis lainnya. MATLAB juga berisi toolbox yang berisi fungsi-fungsi tambahan untuk aplikasi khusus. MATLAB bersifat extensible, dalam arti bahwa seorang pengguna dapat menulis fungsi baru untuk ditambahkan pada library ketika fungsi-fungsi built-in yang tersedia tidak dapat melakukan tugas tertentu. [12]

2.2.5. Data Mining

Data mining merupakan sebuah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi, maka dari itu pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses Knowledge Discovery in Database (KDD) secara keseluruhan. Istilah Knowledge Discovery in

Database (KDD) dan data mining seringkali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain, dan salah satu tahapan dalam keseluruhan proses KDD adalah data mining. Sebagai suatu rangkaian proses, data mining dibagi menjadi beberapa tahap. [13] Tahap – tahap tersebut mempunyai sifat interaktif yang mempunyai arti pemakai terlibat langsung atau melalui perantaraan *knowledge base*. Tahap-tahapannya yaitu sebagai berikut :



1. Pembersihan Data

Pada tahapan ini merupakan tahap pembersihan data yang mencakup membuang duplikasi data, memeriksa inkonsisten, dan memperbaiki kesalahan pada data. Proses ini membuang semua data yang tidak digunakan agar proses data mining lebih mudah dan data yang dihasilkan lebih efektif.

2. Integrasi Data

Pada tahap integrasi ini yaitu menggabungkan data – data yang akan digunakan untuk melalui proses mining dari berbagai database ke dalam sebuah database baru atau yang dinamakan warehouse.

3. Transformasi Data

Pada beberapa teknik data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan seperti beberapa teknik standar seperti analisis, asosiasi, dan klastering dapat menerima input data kategorikal. Proses binning diperlukan untuk membagi – bagi data yang berupa angka numerik menjadi beberapa interval. Di sini dilakukan pula pemilihan data yang diperlukan oleh teknik data mining. Transformasi dan pemilihan data menentukan kualitas dari hasil data mining karena beberapa karakteristik dari teknik data mining tertentu yang tergantung pada tahapan ini.

4. Pengaplikasian Teknik Data Mining

Aplikasi Teknik data mining hanya merupakan salah satu bagian dari proses data mining. Pada tahap ini merupakan proses yang penting untuk menemukan pengetahuan berharga yang tersembunyi dari data dengan menggali informasi dari data tersebut.

5. Evaluasi Pola yang Ditemukan

Tahap evaluasi ini hasil teknik data mining berupa pola-pola atau model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Jika hasilnya tidak sesuai dengan hipotesa, terdapat beberapa alternatif yang dapat diambil seperti menjadikannya sebuah umpan balik untuk memperbaiki proses data mining atau mencoba teknik data mining yang lain atau menerima hasil ini sebagai suatu hasil yang diluar dugaan dan bermanfaat.

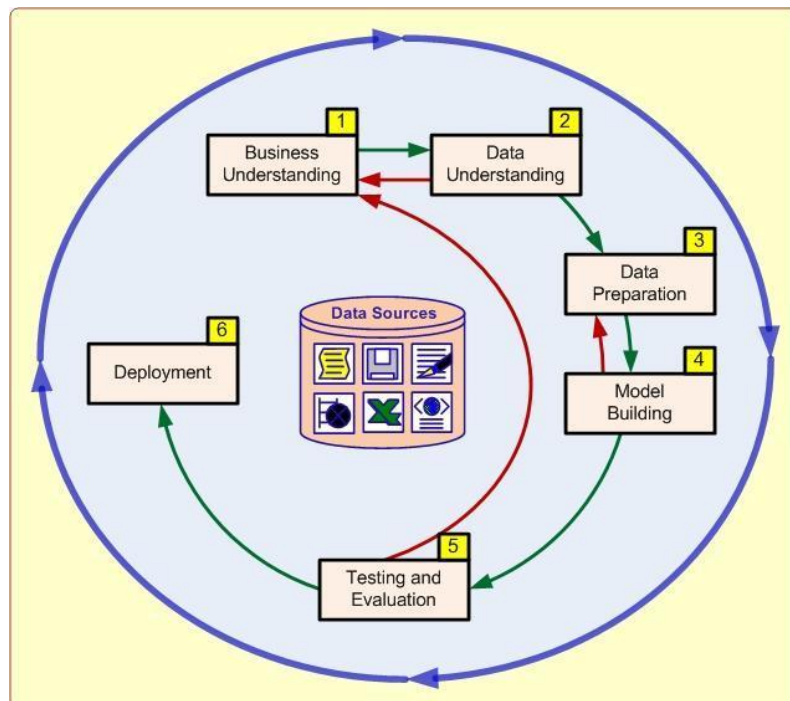
6. Presentasi Pola yang Ditemukan untuk Menghasilkan Aksi

Tahap terakhir adalah bagaimana memformulasikan keputusan atau aksi dari hasil yang didapat. Pada tahap ini terkadang melibatkan orang-orang yang tidak memahami data mining. Hal ini dikarenakan presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang merupakan satu tahapan yang diperlukan dalam proses data mining. Visualisasi dapat membantu menyampaikan hasil data mining.

2.2.6. Cross-Industry Standard Process Model for Data Mining (CRISP-DM)

CRISP-DM merupakan salah satu proses data mining yang menggambarkan urutan langkah pada penelitian Data Mining untuk mengatasi masalah.

1. Pemahaman Bisnis (*Business Understanding*)
2. Pemahaman Data (*Data Understanding*)
3. Pengolahan Data (*Data Preparation*)
4. Permodelan (*Modelling*)
5. Evaluasi (*Evaluation*)
6. Penyebaran (*Deployment*)



Business Understanding adalah tahap pertama dalam melakukan *crispdm* dimana tahap ini adalah memahami tujuan dan kebutuhan dari sudut pandang bisnis, kemudian diterjemahkan ke dalam pendefinisian masalah pada data mining. Lalu selanjutnya menentukan rencana dan strategi agar dapat tercapai tujuan tersebut.

Data Understanding tahap ini adalah tahap pengumpulan data yang akan dilanjutkan dengan proses untuk mendapatkan pemahaman yang mendalam tentang data yang didapat. Pada tahap ini kita mengidentifikasi masalah kualitas data dan mendeteksi jika terdapat bagian menarik dari data yang akan digunakan untuk melakukan hipotesa

Data Preparation tahapan ini meliputi semua kegiatan untuk membangun data yang akan diproses pada tahap pemodelan yang dibuat dari data mentah. Pada tahap ini juga dapat diulang beberapa kali, selain membangun data tahap ini juga mencakup pemilihan tabel, record, dan atribut data lainnya sekaligus melakukan proses pembersihan data atau melakukan transformasi data yang akan dimasukkan dalam tahap pemodelan

Modeling pada tahap ini yang akan dilakukan adalah pemilihan dan penerapan teknik – teknik pemodelan dan beberapa parameternya akan disesuaikan untuk mendapatkan nilai yang optimal. Secara khusus, dapat melakukan beberapa teknik yang berbeda dalam data mining yang sama.

Evaluation tahap ini model telah terbentuk dan diharapkan mempunyai kualitas yang baik jika dilihat dari sudut pandang Analisa data. Dan ditahap ini akan dilakukan evaluasi terhadap keefektifan dan kualitas suatu model sebelum digunakan. Di tahap ini akan dinilai apakah model dapat mencapai tujuan yang ditetapkan pada fase awal yaitu business understanding atau tidak.

Deployment tahapan terakhir dalam crisp-dm adalah melakukan deployment dimana tahapan ini knowledge atau informasi yang telah diperoleh akan diatur dan dipresentasikan dalam bentuk khusus sehingga dapat digunakan oleh pengguna. Tahapan ini dapat berupa membuat laporan sederhana atau meimplementasikan proses data mining yang berulang dalam perusahaan.

BAB III

METODE PENELITIAN

3.1 Metode Pengumpulan Data

Data pada penelitian ini menggunakan data yang diperoleh dengan cara observasi. Pengumpulan data dengan cara observasi berarti data dikumpulkan dengan cara mendatangi beberapa orang untuk dimintai foto *selfie* atau foto diri dengan berbagai pose foto dan sudut foto yang berbeda pada setiap fotonya. Jumlah sampel foto yang dibutuhkan sebanyak 20 foto per orangnya. Selain itu juga menggunakan beberapa Studi Pustaka (*Library Research Method*) yang berhubungan dengan pengolahan citra, metode *eigenface*, *Principal Component Analysis* (PCA), serta mencari informasi, artikel dan literatur dari internet.

3.2 Teknik Analisa Data

Data yang digunakan yaitu berupa foto berasal dari 20 orang dan terdiri dari 10 laki-laki dan 10 perempuan. Masing-masing orang mempunyai 20 foto yang diambil dengan pose dan sudut yang berbeda sehingga total data keseluruhan adalah 400 foto dan dengan 10 atribut yang berpengaruh terhadap pencarian wajah seseorang. Terdapat beberapa tahapan yang dilakukan untuk mendapatkan data yang berkualitas yaitu sebagai berikut :

1. Pembersihan Data (*Data Cleaning*)

Tahap pertama yang dilakukan adalah melakukan pemberisihan data yang tidak terpakai atau data yang kembar. Data dapat dinyatakan tidak terpakai apabila contohnya foto tidak tampak wajah sama sekali maka foto tersebut akan digantikan oleh foto lain dari orang tersebut. Jika tidak ada foto lain dari orang tersebut dan jika foto yang dibuang terlalu banyak, maka akan diputuskan untuk meminta foto ulang.

2. Integrasi dan Transformasi Data

Setelah melalui proses data cleaning, maka tahap selanjutnya adalah tahap integrasi dan transformasi data. Tahap integrasi digunakan dalam menganalisis data korelasi, atribut yang redundan atau duplikat data. Kemudian melakukan tahapan transformation untuk meningkatkan akurasi dan mengefisiensi fungsi algoritma.

3. Reduksi Data

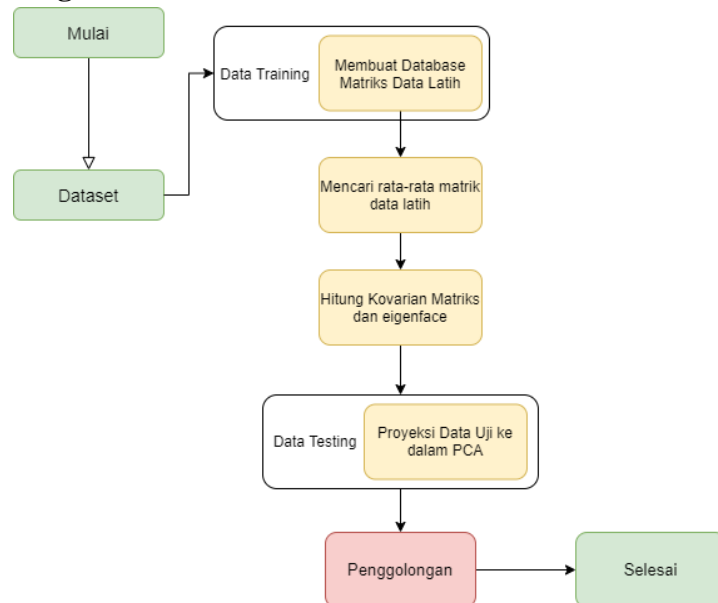
Tahapan reduksi data yaitu mengurangi jumlah atribut dan record agar data yang diolah menjadi lebih efektif. Berikut merupakan atribut data sebelum dan sesudah dilakukannya reduksi data :

No.	Atribut data penelitian sebelum direduksi
1	Nama
2	Gender
3	Rambut
4	Golongan
5	Sampel Data
6	Tgl Pengambilan
7	Data Latih
8	Data Uji
9	Proses Benar
10	Proses Salah

No.	Atribut data penelitian yang digunakan
1	Nama
2	Gender
3	Golongan
4	Data Latih
5	Data Uji
6	Proses Benar

3.3 Model yang Diusulkan

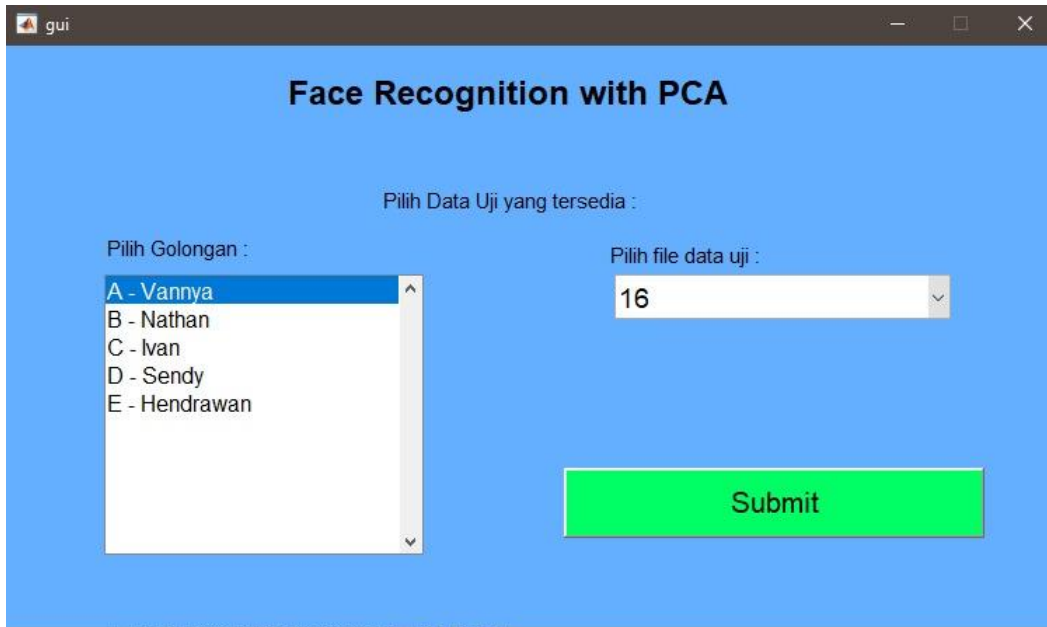
3.3.1. Modeling



Gambar 3.1

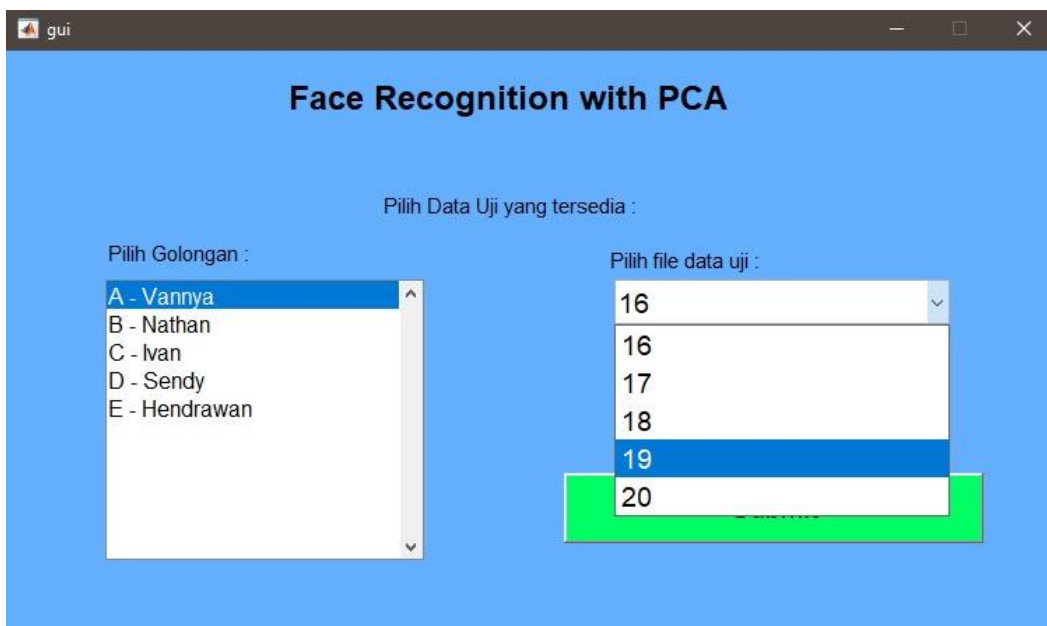
3.3.2. Model Design Interface

Pada pilih golongan merupakan *List Box* yang merupakan list yang berisi golongan apa saja yang terdapat dalam database data uji



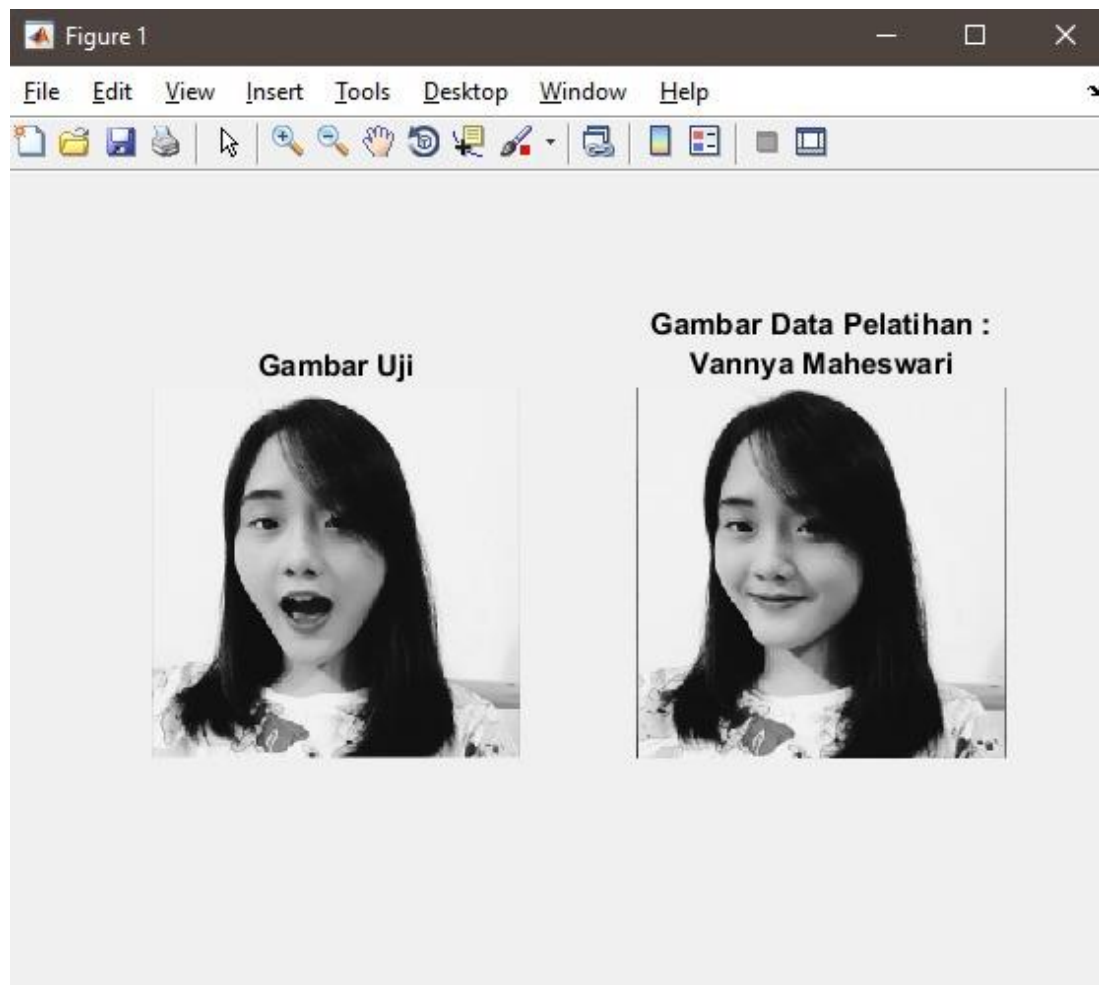
Gambar 3.2

Pada pilihan file data uji merupakan tombol *combo box* yang akan menampilkan angka 16-20 yaitu nama file dari tiap golongan



Gambar 3.2

Setelah user memilih golongan dan nomor file uji, dapat menekan tombol submit yang akan menjalankan program dan fungsi matlab yang tersedia.



Gambar 3.3

Gambar diatas merupakan gambar hasil akhir tampilan setelah program dijalankan.

BAB IV

JADWAL PENELITIAN DATA MINING

4.1. Jadwal Penyusunan Penelitian Data mining

Dalam melakukan penelitian data mining ini, kami membuat tahapan dengan melakukan jadwal penyusunan seperti berikut :

Tabel 4.1 Jadwal Penyusunan Penelitian Data Mining

No	Kegiatan	Februari		Maret				April
		1	2	3	4	5	6	7
1	Pengumpulan Dataset							
2	Membuat BAB I							
3	Membuat BAB II							
4	Membuat BAB III							
5	Mencari Sumber Pustaka							
6	Mengolah Data							
6	Membuat Program							
7	Membuat Bab IV							

Daftar Pustaka

- [1] M. Ngafifi, “Kemajuan Teknologi Dan Pola Hidup Manusia Dalam Perspektif Sosial Budaya,” *J. Pembang. Pendidik. Fondasi dan Apl.*, vol. 2, no. 1, pp. 33–47, 2014, doi: 10.21831/jppfa.v2i1.2616.
- [2] 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.
- [3] A. H. Dian Esti Pratiwi, “Implementasi Pengenalan Wajah Menggunakan PCA (Principal Component Analysis),” *Justus Liebig's Ann. Chem.*, vol. 502, no. 1, pp. 74–85, 2016, doi: 10.1002/jlac.19335020105.
- [4] D. Parikesit, “Face Recognition Menggunakan Metode PCA,” 2012.
- [5] Derisma, “Faktor-Faktor yang Mempengaruhi Sistem Pengenalan Wajah Menggunakan Metode Eigenface pada Perangkat Mobile Berbasis Android,” *J. Politek. Caltex Riau*, vol. 2, no. 2, pp. 127–136, 2016.
- [6] R. D. Kusumanto and A. N. Tompunu, “PENGOLAHAN CITRA DIGITAL UNTUK MENDETEKSI OBYEK MENGGUNAKAN PENGOLAHAN WARNA MODEL NORMALISASI RGB,” vol. 2011, no. Semantik, 2011.
- [7] N. Ahmad and A. Hadinegoro, “Metode Histogram Equalization untuk Perbaikan Citra Digital,” *Semin. Nas. Teknol. Inf. Komun. Terap.*, vol. 2012, no. Semantik, pp. 439–445, 2012, [Online]. Available: <http://publikasi.dinus.ac.id/index.php/semantik/article/view/185>.
- [8] J. Rekursif, A. R. Syakhala, D. Puspitaningrum, and E. P. Purwandari, “PERBANDINGAN METODE PRINCIPAL COMPONENT ANALYSIS (PCA) DENGAN METODE HIDDEN MARKOV MODEL (HMM) DALAM PENGENALAN IDENTITAS,” vol. 3, no. 2, pp. 68–81, 2017.
- [9] S. M. Kom and U. Abdurrah, “RANCANG BANGUN SISTEM PENGENALAN WAJAH DENGAN METODE,” no. August 2016, 2017, doi: 10.36341/rabit.v1i2.25.
- [10] E. Pca, “Sistem Pengenalan Wajah Secara Real-Time,” vol. 7, no. 2, pp. 179–

184, 2016.

- [11] A. Firmansyah, “D a s a r - d a s a r P e m r o g r a m a n M a t l a b,” pp. 1–10, 2007.
- [12] B. Cahyono, “Penggunaan Software Matrix Laboratory (Matlab) Dalam Pembelajaran Aljabar Linier,” vol. 1, pp. 45–62.
- [13] Y. Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4.5,” *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017.