BAB II
TINJAUAN PUSTAKA

2.1 Pengolahan Citra (*Image Processing*)

Citra adalah sinyal dua dimensi yang bersifat kontinu dan dapat diamati oleh sistem visual manusia. Secara matematis, citra adalah fungsi dua dimensi yang menyatakan intensitas cahaya pada bidang dua dimensi, yang disimbolkan dengan $f(x,y)$, di mana (x,y) adalah koordinat pada bidang dua dimensi dan $f(x,y)$ adalah intensitas cahaya (brightness) pada titik (x,y)[5]. Citra digital adalah representasi citra melalui pencuplikan (sampling) secara spasial dan temporal. Pencuplikan adalah proses untuk menentukan warna pada piksel tertentu pada citra dari sebuah gambar yang kontinu. Citra digital direpresentasikan sebagai matriks berukuran $M \times N$. Setiap elemen matriks menyatakan sebuah piksel (picture element), dan $M \times N$ menyatakan resolusi citra [5].

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1)\cdots & f(0,M) \\ f(1,0) & f(1,1)\cdots & f(1,M) \\ \vdots & \vdots & \vdots \\ f(N-1,0) & f(N-1,1) & f(N-1,M-1) \end{bmatrix} \tag{2.1}$$

2.2 *Principal Component Analysis* (PCA)

Principal Component Analysis (PCA) adalah pendekatan statistik yang digunakan untuk mengurangi jumlah variabel dalam pengenalan wajah. Dalam PCA, setiap gambar dalam set pelatihan direpresentasikan sebagai kombinasi linier dari vektor eigen berbobot yang disebut *Eigenfaces*. Vektor eigen ini diperoleh dari matriks kovarian dari set gambar pelatihan. Bobot ditemukan setelah memilih satu set *Eigenfaces* yang paling relevan. Pengenalan dilakukan dengan memproyeksikan gambar uji ke dalam subruang yang direntang oleh eigenfaces dan kemudian klasifikasi dilakukan dengan mengukur jarak Euclidean minimum [6].

Bayangkan sebuah gambar wajah I (x,y) dengan dimensi $N \times N$. Citra ini dapat digambarkan sebagai vektor dengan dimensi N^2 atau sebagai array nilai identitas dua dimensi $N \times N$. Citra dengan resolusi 256×256 akan digambarkan sebagai vektor dengan dimensi 65.536, yang setara dengan sebuah titik dalam ruang 65.536 dimensi. Untuk citra wajah, PCA berfokus menemukan nilai vektor terbaik untuk

menggambarkan distribusi data gambar wajah di seluruh ruang dimensi gambar. Berikut langkah-langkah untuk mengekstraksi fitur PCA:

1. Terdapat satu set gambar M dengan berukuran $N \times N$ dapat digambarkan dengan $I_1, I_2, I_3, ...I_M$ yang membentuk vektor berukuran N^2. Citra wajah direprentasikan sebagai $\Gamma_i = \{\Gamma_1, \Gamma_2, ..., \Gamma_M\}$, dimana setiap Γ_i adalah vektor dari dimensi N dan jumlah citra wajah dalam dataset pelatihan dilambangkan dengan M.

2. Reprentasikan setiap citra I_i menjadi sebuah vektor Γ_i sebagai berikut:

$$\begin{bmatrix} a_{1_1} & a_{1_2} & \cdots & a_{1_N} \\ a_{2_1} & a_{2_2} & \cdots & a_{2_N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N_1} & a_{2_2} & \cdots & a_{2_N} \end{bmatrix} \xrightarrow{\text{concatenate}} \begin{bmatrix} a_{1_1} \\ \vdots \\ a_{1_N} \\ \vdots \\ a_{2_N} \\ \vdots \\ a_{N_N} \end{bmatrix} = \Gamma_i \tag{2.2}$$

3. Rata-rata set pelatihan ditentukan oleh

$$\Psi = \frac{1}{M} \sum_{i=1}^{M} \Gamma_i \tag{2.3}$$

4. Mencari rata-rata dari beberapa wajah yang berbeda dengan vektor $\{\Phi_1, \Phi_2, \Phi_3, \ldots, \Phi_M\}$

$$\Phi_i = \Gamma_i - \Psi \tag{2.4}$$

5. Mencari matriks kovarian dengan persamaan sebagai berikut:

$$C = A \times A^T, \quad \text{dimana} \quad A = [\Phi_1, \Phi_2, \Phi_3, \ldots, \Phi_M] \tag{2.5}$$

   Dalam kasus ini, matriks kovarian C menggambarkan hubungan antara dua matriks dan varians yang ada dalam dataset. Sementara itu, A^T adalah transpos dari matriks A, di mana matriks A menyimpan perbedaan antara setiap citra latih dan citra wajah rata-rata. Keunikan dari matriks kovarian ini adalah kemampuannya untuk menghasilkan vektor eigen dan nilai eigen yang diperlukan.

6. Matriks berukuran $w \times h$ dari citra wajah, yang diperlakukan sebagai vektor

berdimensi N, menghasilkan matriks kovarian C dengan dimensi N^2. Perhitungan matriks kovarian yang besar ini dapat dihindari dengan menggunakan persamaan berikut:

$$C \times u_i = \lambda_i \times u_i \tag{2.6}$$

Vektor eigen u_i dan nilai eigen λ_i merupakan nilai eigen yang berhubungan dengan citra wajah tersebut.

7. Kita dapat menentukan vektor eigen dan nilai eigen dari matriks kovarian (C). Misalkan $L = A^T \times A$. L adalah matriks berdimensi $M \times M$, di mana M lebih kecil dari N. Ini membuat perhitungan vektor eigen dan nilai eigen menjadi lebih sederhana dari matriks tersebut dengan menyelesaikan persamaan berikut:

$$\begin{aligned} L \times V_i &= d_i \times V_i \\ A^T \times A \times V_i &= d_i \times V_i \\ A \times A^T \times A \times V_i &= d_i \times A \times V_i \\ C \times A \times V_i &= d_i \times A \times V_i \end{aligned}$$

Dari persamaan di atas, dapat disimpulkan bahwa $u_i = A \times V_i$ dan $\lambda_i = d_i$ merupakan vektor eigen dan nilai eigen dari matriks kovarian. Dengan menggunakan matriks L, kita dapat memperoleh vektor eigen dan nilai eigen sebanyak M.

8. Langkah selanjutnya adalah secara heuristik memilih K vektor eigen yang terbalik. Setelah itu, sebelum melakukan perhitungan bobot, Φ_i direpresentasikan sebagai kombinasi linear dari vektor eigen U_i dengan cara berikut:

$$\Phi_i = \sum_{J=1}^{K} W_j u_j, \quad \text{dimana} \quad u_i \quad \text{adalah eigenface.} \tag{2.7}$$

Sehingga nilai bobot dapat dihitung dengan rumus:

$$W_j = U_j^T \Phi_i \tag{2.8}$$

Langkah selanjutnya citra training yang sudah dinormalisasi, diubah menjadi

sebuah vektor berikut:

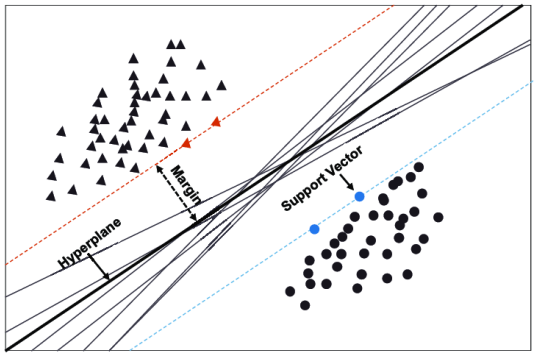
$$\Omega = \begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ \vdots \\ W_K \end{bmatrix} \tag{2.9}$$

Di mana $i = 1, 2, 3, \ldots, M$. Pada akhirnya, vektor yang telah dihitung untuk setiap gambar dalam *training set* akan disimpan sebagai template. Setelah bobot vektor diketahui, langkah berikutnya adalah tahap pengenalan dengan menerapkan proses klasifikasi.

2.3 *Support Vector Machine* (SVM)

Support Vector Machine (SVM) adalah metode klasifikasi yang pertama kali diperkenalkan oleh Vapnik pada tahun 1998. Secara prinsip, metode ini bekerja dengan menentukan batas pemisah antara dua kelas menggunakan data yang paling dekat dengan batas tersebut. Algoritma ini beroperasi dalam ruang fitur berdimensi tinggi, yang membuat perhitungan produk skalar di ruang fitur menjadi lebih kompleks. Untuk menghitung produk skalar ini, digunakan fungsi kernel. Fungsi kernel memungkinkan perhitungan produk skalar tanpa perlu menghitung ruang fitur secara eksplisit [7].

Awalnya dirancang sebagai pengklasifikasi linear, metode SVM juga dapat digunakan untuk memetakan data nonlinier. Hal ini ditunjukkan oleh cara pemisahan data dengan metode SVM pada contoh sederhana yang ditampilkan pada Gambar 2.1.



Gambar 2.1. Pemisahan data dengan metode SVM

SVM memisahkan data berdimensi p menggunakan bidang berdimensi $p-1$, yang juga dikenal sebagai *hyperplane*. Bidang ini memaksimalkan margin pemisahan data, yang diukur sebagai jarak minimal dari sampel ke *hyperplane*.

Dengan menggunakan SVM, proses pembelajaran bertujuan untuk menghasilkan hipotesis berupa bidang pemisah terbaik yang meminimalkan risiko empiris, yaitu rata-rata kesalahan pada data pelatihan, serta memiliki kemampuan generalisasi yang baik. Generalisasi adalah kemampuan sebuah hipotesis untuk mengklasifikasikan data yang tidak termasuk dalam data pelatihan dengan benar. Untuk menjamin generalisasi ini, SVM bekerja berdasarkan prinsip *Structural Risk Minimization* (SRM). Berikut adalah algoritma SVM:

1. Fungsi kernel *(linear, polynomial, radial basis function, dan sigmoid)* digunakan untuk memetakan atribut data ke dalam ruang dimensi yang lebih tinggi.
2. Memetakan data yang telah dipisahkan ke dalam ruang dimensi tinggi untuk mencari *hyperplane* terbaik.
3. Memaksimalkan jarak *hyperplane* dengan titik data terdekat *(Support Vector)* adalah cara untuk mendapatkan *hyperplane* terbaik.

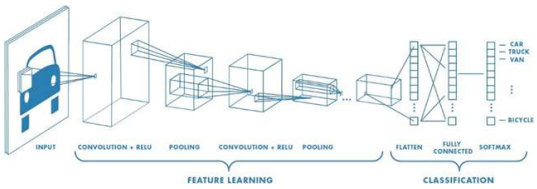
2.4 *Deep learning*

Deep learning (DL)adalah jenis model pembelajaran mesin (machine learning), dan ketika model ini disesuaikan dengan data, proses tersebut disebut sebagai DL. Jaringan neural dalam bentuk modern saat ini merupakan model pembelajaran mesin yang paling kuat dan praktis, dan sering ditemui dalam kehidupan sehari-hari. Contoh aplikasi DL termasuk penerjemahan teks menggunakan algoritma pemrosesan bahasa alami, pencarian gambar di internet melalui sistem pengenalan objek, dan percakapan dengan asisten digital melalui antarmuka pengenalan suara. Semua aplikasi ini didukung DL [8].

2.4.1 *Convolutional Neural Network*

Salah satu jenis algoritma *Deep Learning* yang dikembangkan berdasarkan *Multilayer Perception* (MLP) adalah *Convolutional Neural Network* (CNN). CNN biasanya digunakan untuk jenis data gambar. Metode CNN ini sering digunakan untuk mendeteksi dan mengidentifikasi objek pada gambar. Dalam proses ini, CNN ini menerapkan proses filtering dengan menentukan ukuran tertentu pada gambar, dan kemudian menghasilkan informasi yang memiliki representatif baru berdasarkan hasil perkalian matriks gambar dengan ukuran filter yang digunakan. Hasil yang dihasilkan sebanding dengan jumlah data gambar yang dikumpulkan dan diproses

untuk latihan [9]. CNN memiliki beberapa lapisan yang digunakan untuk melakukan proses filtering selama pemrosesannya, yang dikenal sebagai proses pelatihan. Proses pelatihan terdiri dari tiga tahap: lapisan convolutional, lapisan pooling, dan lapisan penuh terhubung, seperti yang ditunjukkan pada Gambar 2.2.



Gambar 2.2. Tahapan Proses CNN

2.5 Perhitungan Akurasi

Akurasi dalam citra merujuk pada seberapa tepat suatu sistem atau algoritma dapat mengidentifikasi, mengklasifikasikan dalam pengenalan wajah untuk menghitung parameter yang diuji seperti iluminasi, oklusi dan ekspresi. Akurasi biasanya dievaluasi dengan membagi citra yang dikenali dengan total citra yang ada. Akurasi didapat dengan menggunakan perhitungan sebagai berikut:

$$\text{Akurasi} = \left(\frac{\text{Citra yang Dikenali}}{\text{Total Citra yang Diuji}} \right) \times 100\% \tag{2.10}$$

2.6 Raspberry Pi 3 Model B+

Raspberry Pi Foundation adalah sebuah *embedded* komputer yang dibuat oleh Raspberry Pi Foundation, sebuah badan amal yang berbasis di Inggris yang didirikan dengan tujuan membantu orang belajar menggunakan komputer dan membuat pendidikan komputer lebih mudah diakses [10]. Raspberry Pi memiliki varian salah satunya yaitu Raspberry Pi 3 Model B+ yang meiliki dimensi hanya 86 x 54 mm dari luar. Casing kecil ini menampung semua komponen yang diperlukan untuk menggunakannya sebagai komputer pribadi biasa. Prosesor quad-core 64-bit Broadcom BC2837 ARM-8 ARM-8 yang dilengkapi dengan grafis Cortex -A53 adalah komponen utama perangkat ini. Model 3B + memiliki memori operasi yang ditingkatkan menjadi 1GB, dan kartu MicroSD berfungsi sebagai hard disk dalam pembaca perangkat. Dengan solusi ini, setiap pengguna dapat mengubah kapasitas kartu sesuai dengan kebutuhan dan cara komputer mikro digunakan [11]. Pada tabel 2.1 merupakan Spesifikasi Raspberry Pi 3 Model B+ yang digunakan.

Tabel 2.1. Spesifikasi Raspberry Pi 3 Model B+

Parameter	Description
Processor	1.2 GHz
Number of cores	4 Quad Core
Number of cores RAM	1 GB
Memory	MicroSD memory card
GPIO	40
USB	4
HDMI	1.4 Standard
Communication	Ethernet, Wi-Fi, Bluetooth
Protocols	UART, I2C, SPI
System	System

2.7 Dataset

Data citra wajah yang diambil secara mandiri disebut dataset mandiri. Terdapat 98 citra wajah dari 14 objek yang berbeda. Setiap citra dalam file tersebut berukuran 640 × 480 piksel dan berformat JPG. Gambar 2.3 menunjukkan ilustrasi gambar wajah dari dataset mandiri.

Gambar 2.3. Dataset Mandiri