

BAB II

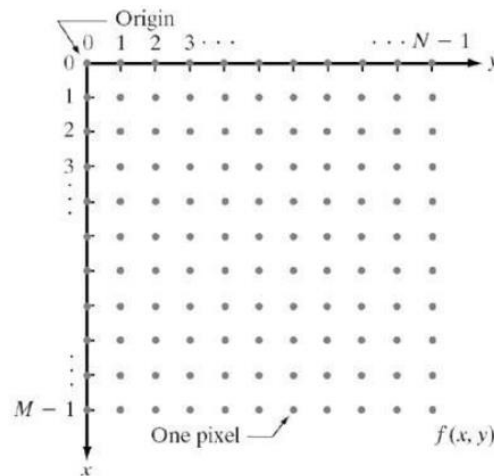
LANDASAN TEORI

Bab ini akan membahas teori dari yang telah dilakukan oleh peneliti-peneliti sebelumnya yang berhubungan dengan penjelasan sistem-sistem yang akan dipakai pada penelitian ini.

2.1. Pengolahan Citra Digital

Secara umum pengolahan citra digital menunjuk pada pemrosesan gambar dimensi dua menggunakan komputer. Dalam konteks yang lebih luas, pengolahan citra digital mengacu pada pemrosesan setiap data berdimensi dua. Citra digital merupakan sebuah larik (array) yang berisi nilai-nilai real maupun kompleks yang direpresentasikan dengan deretan bit tertentu.[6]

Suatu citra dapat didefinisikan sebagai fungsi $f(x, y)$ berukuran M baris dan N kolom, dengan x dan y adalah koordinat spasial, dan amplitudo f di titik koordinat (x,y) dinamakan intensitas atau tingkat keabuan dari citra pada titik tersebut. Apabila nilai (x,y) , dan nilai amplitudo f secara keseluruhan berhingga (finite) dan bernilai diskrit maka dapat dikatakan bahwa citra tersebut adalah citra digital. Gambar 1 menunjukkan posisi koordinat citra digital. [7]



Gambar 2. 1 Koordinat Citra Digital[7]

Citra digital dapat di tulis dalam bentuk matriks sebagai berikut.

$$f(x, y) \begin{bmatrix} f(1,1) & f(1,2) & \cdots & f(1,N) \\ f(2,1) & f(2,2) & \cdots & f(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ f(M,1) & f(M,2) & \cdots & f(M,N) \end{bmatrix} \dots\dots(1)$$

Nilai pada suatu irisan antara baris dan kolom (pada posisi x,y) disebut dengan picture elements, image elements, pels, atau pixel

2.2 Grayscale

Citra *grayscale* adalah suatu citra yang hanya memiliki warna tingkat keabuan. Warna abu-abu pada citra *grayscale* adalah warna RGB yang memiliki intensitas yang sama. Citra grayscale dapat diperoleh dari citra berwarna melalui transformasi dari ruang warna RGB ke ruang warna lain.[8]

$$K_o = (R_i + G_i + B_i) / 3 \dots\dots (2)$$

Dimana K_o adalah nilai warna output, sedangkan R_i , G_i , dan B_i adalah nilai RGB input.

2.3 Thresholding

Thresholding merupakan salah satu teknik segmentasi yang baik digunakan untuk citra dengan perbedaan nilai intensitas yang signifikan antara latar belakang dan objek utama. Region citra yang cenderung gelap akan dibuat semakin gelap (hitam sempurna dengan nilai intensitas sebesar 0), sedangkan region citra yang cenderung terang akan dibuat semakin terang (putih sempurna dengan nilai intensitas sebesar 1). Oleh karena itu thresholding berupa citra biner dengan nilai intensitas piksel sebesar 0 atau 1. Setelah citra sudah tersegmentasi atau sudah berhasil dipisahkan objeknya dengan background, maka citra biner yang diperoleh dapat dijadikan sebagai masking untuk melakukan proses cropping sehingga diperoleh tampilan citra asli tanpa background.[9]

2.4 Haar Cascade

Deteksi objek menggunakan fungsi Haar berdasarkan cascade classifier adalah metode deteksi objek Paul Viola dan Michael Jones pada tahun 2001, mereka mempresentasikan makalah yang disebut "Deteksi Objek Cepat Menggunakan Perangkat Tambahan Sederhana". Haar Cascade adalah kumpulan fungsi Haar-Like, yang digabungkan untuk membentuk pengklasifikasi. Fiturnya adalah jumlah nilai piksel putih yang dikurangkan dari nilai piksel pada area hitam.

Fungsi Haar-like-feature atau biasa disebut juga dengan Haar cascade classifier adalah fungsi persegi panjang (persegi) yang memberikan indikasi spesifik pada gambar. Pengklasifikasi Haar Cascade berasal dari gabungan piksel hitam dan piksel putih yang membentuk kotak.

Untuk proses deteksi wajah digunakan algoritma Haar Cascade. Umumnya, fungsi seperti haar digunakan untuk mendeteksi objek dalam gambar digital. Kata Haar sendiri dinyatakan sebagai fungsi matematika dalam bentuk kotak (wavelet Hhaar). Awalnya, pengolahan citra hanya didasarkan pada nilai RGB dari setiap piksel, tetapi cara pengolahan seperti itu tidak efektif. Kemudian, Viola dan Jones mengembangkan dan membentuk fungsi Haar-Like. Fitur Haar-like menangani gambar dalam kotak, di mana ada beberapa piksel dalam satu bingkai. Kemudian proses setiap kotak dan hasilkan nilai yang berbeda untuk menunjukkan area gelap dan terang. Nilai-nilai ini akan digunakan sebagai dasar untuk pengolahan citra. [5]

2.5. Deep Learning

Deep learning adalah suatu *artificial neural network* yang memiliki banyak *layer*. Pada umumnya *deep learning* sendiri memiliki lebih dari 3 *layers* dan dalam proses pembelajarannya sendiri disebut *deep neural network (DNN)*. Algoritma pada *deep learning* memiliki fitur yang unik yaitu sebuah fitur yang mampu mengekstraksi secara otomatis. Hal ini berarti algoritma yang dimilikinya secara otomatis dapat menangkap fitur yang relevan sebagai keperluan dalam pemecahan suatu masalah. Algoritma semacam ini sangat penting dalam sebuah kecerdasan buatan karena mampu mengurangi beban pemrograman dalam memilih fitur yang eksplisit, Deep learning mampu melakukan proses pembelajaran mandiri untuk pengenalan objek, ekstraksi objek dan klasifikasi [10].

2.6. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) termasuk dalam jenis deep learning karena kedalaman jaringannya. Deep learning adalah cabang dari machine learning yang dapat mengajarkan komputer untuk melakukan pekerjaan layaknya manusia, seperti komputer dapat belajar dari proses training. CNN merupakan operasi konvolusi yang menggabungkan beberapa lapisan pemrosesan, menggunakan beberapa elemen yang beroperasi secara paralel dan terinspirasi oleh sistem saraf biologis. Pada CNN setiap neuron dipresentasikan dalam bentuk 2 dimensi, sehingga metode ini cocok untuk pemrosesan dengan input berupa citra. Metode jaringan syaraf tiruan adalah pengenalan pola dengan melakukan proses pembelajaran atau pelatihan ciri fitur pada tiap masukan untuk kemudian dilakukan proses pengenalan. [10] CNN memiliki fungsi untuk melakukan ekstraksi fitur. Fitur-fitur perlu didapatkan guna proses atau tugas seperti klasifikasi, *clustering* ataupun regresi. Pada *machine learning* konvensional dilakukan ekstraksi fitur manual, artinya ditentukan terlebih dahulu fitur-fitur yang diekstraksi. Sedangkan CNN melakukan ekstraksi fitur secara otomatis pada convolutional layer, *pooling layer* dan juga aktivasi *Rectified Linear Unit* (ReLU). Sedangkan proses klasifikasi fitur-fitur dilakukan pada *Fully Connected layer* (FCL) dan aktivasi softmax.[11]

2.7. Alexnet

Arsitektur jaringan pada penelitian ini menggunakan struktur dari Alexnet. Alexnet dikembangkan Krizhevsky pada tahun 2012 yang merupakan basis dari arsitektur CNN modern yang cukup sukses untuk visual recognition. Untuk mengatasi

kekurangan pada proses training data citra resolusi tinggi, perlu pengaturan pada beberapa lapisan dan mengeksplorasi parameter untuk training data agar model CNN dapat menampilkan performa yang baik dan mencegah gradien menjadi tidak stabil, khususnya pada jaringan yang dalam.

Sedangkan Pada penelitian yang dilakukan oleh Mingjie Jiang, Xinqi Fan, dan Hong Yan , membuat sistem bernama Retina Face Mask, sebuah masker detector dengan menggunakan ResNet dan MobileNet sebagai algoritma utama untuk melakukan Deep Learningnya[12] .Di mana setiap gambar memiliki perbandingan masing-masing dari hasilnya, seperti ditunjukan pada gambar 2.2 dibawah.

Backbone	Transfer Learning	Face		Mask	
		Precision	Recall	Precision	Recall
MobileNet	Imagenet	80.5%	93.0%	82.8%	89.0%
		79.0%	92.8%	78.9%	89.1%
	Wilder Face	83.0%	95.6%	82.3%	89.1%
		82.5%	95.4%	82.4%	89.3%
ResNet	Imagenet	91.0%	95.8%	93.2%	94.4%
		91.5%	95.6%	93.3%	94.4%
	Wilder Face	91.9%	96.3%	93.4%	94.5%
		91.9%	95.7%	92.9%	94.8%

Gambar 2. 2 Perbandingan algoritma [12]

Pada alexnet terdapat beberapa arsiteksektur jaringannya yang terdiri dari , Konvolusi, Pooling Layer, ReLu dan Softmax.

2.7.1 Konvolusi

Konvolusi didefinisikan sebagai proses untuk memperoleh suatu piksel didasarkan pada nilai piksel itu sendiri dan tetangganya dengan melibatkan suatu matriks yang

disebut kernel yang merepresentasikan pembobotan, Operasi ini menerapkan fungsi output sebagai Feature Map dari masukan citra. [13]

2.7.2 Pooling Layer

Pooling layer merupakan layer yang digunakan untuk downsampling input untuk memperkecil dimensi. Pooling layer yang digunakan adalah max pooling dan average pooling. Lapisan Pooling bekerja di setiap tumpukan Feature Map dan mengurangi ukurannya. Bentuk lapisan Pooling yang paling umum adalah dengan menggunakan filter berukuran 2x2 yang diaplikasikan dengan langkah sebanyak 2 dan kemudian beroperasi pada setiap irisan dari input.[14]

2.7.3 ReLU

ReLU (Rectified Linear Unit) merupakan lapisan aktivasi pada model CNN yang mengaplikasikan fungsi $f(x)=\max(0,x)$ yang berarti fungsi ini melakukan thresholding dengan nilai nol terhadap nilai piksel pada input citra.[14]

2.7.4 Softmax

Aktivasi Softmax atau Softmax Classifier merupakan bentuk lain dari algoritma Logistic Regression yang dapat digunakan untuk mengklasifikasi lebih dari dua kelas. Standar klasifikasi yang biasa dilakukan oleh algoritma Logistic Regression adalah tugas untuk klasifikasi kelas biner. Softmax juga memberikan hasil yang lebih intuitif dan juga memiliki interpretasi probabilistik yang lebih baik dibanding algoritma klasifikasi lainnya. Softmax memungkinkan menghitung probabilitas untuk semua label.[13]

2.8 Matlab

Matlab adalah sebuah bahasa dengan kemampuan tinggi untuk komputasi teknis. Dan menggabungkan komputasi, visualisasi, dan pemrograman dalam satu kesatuan yang mudah digunakan dimana masalah dan penyelesaiannya diekspresikan dalam notasi matematik yang sudah dikenal, Matlab sebagai salah satu *tools* pemrograman untuk membantu bidang pendidikan dan penelitian telah menyediakan bermacam-macam *toolbox* yang disesuaikan dengan bidang keilmuan masing-masing, salah satunya adalah *Image Processing Toolbox*. Dengan memanfaatkan *toolbox* tersebut, pengguna dapat dengan mudah melakukan penelitiannya. Matlab atau matrix laboratory merupakan salah satu bahasa pemrograman, yaitu media yang digunakan untuk berinteraksi antara manusia dengan komputer. Matlab dikembangkan sebagai bahasa pemrograman sekaligus alat visualisasi yang menawarkan banyak kemampuan untuk menyelesaikan berbagai kasus yang berhubungan langsung dengan disiplin keilmuan matematika, seperti bidang rekayasa teknik, fisika, statistika, komputasi, dan modeling.[15]

Sedangkan untuk menghentikan proses yang sedang berjalan pada matlab dapat dilakukan dengan menekan tombol Ctrl-C. lalu untuk keluar dari matlab dapat dilakukan dengan menuliskan perintah exit atau quit pada command window atau dengan menekan menu exit pada bagian menu file dari menu bar.

2.9 Human Interface

Human Interface merupakan teknologi antarmuka manusia untuk meningkatkan kemampuan seseorang saat melakukan tugas yang membutuhkan alat

atau perangkat pada pengguna, dengan memberikan pemahaman tentang interaksi manusia dengan alat, mesin, komputer, atau instrumen lain dalam dunia nyata[16]

Desain *Human Interface* adalah seperangkat alat atau elemen yang digunakan untuk memanipulasi objek digital. Untuk menyatakan bahwa sebuah *Human Interface* dianggap bagus apabila dapat berfungsi dengan baik, tidak hanya mempertimbangkan aspek estetik saja. Dengan kata lain, dalam menentukan bentuk desain *Human Interface*, tidak hanya dibutuhkan aspek estetik visual, namun juga harus mempertimbangkan aspek fungsinya.