

BAB II

LANDASAN TEORI

2.1 Image Processing

Image processing atau pengolahan citra adalah suatu bentuk pengolahan atau pemrosesan sinyal dengan input berupa gambar (*image*) dan ditransformasikan menjadi gambar lain sebagai keluarannya dengan teknik tertentu. *Image processing* dilakukan untuk memperbaiki kesalahan data sinyal gambar yang terjadi akibat transmisi dan selama akuisisi sinyal, serta untuk meningkatkan kualitas penampakan gambar agar lebih mudah diinterpretasi oleh sistem penglihatan manusia baik dengan melakukan manipulasi dan juga penganalisisan terhadap gambar. Operasi *Image processing* dapat dikelompokkan berdasarkan dari tujuan transformasinya, yaitu:

a. Image Enhancement

Image Enhancement atau perbaikan kualitas citra merupakan operasi yang bertujuan untuk memperbaiki kualitas citra dengan cara memanipulasi parameter-parameter citra. Dengan operasi ini, ciri-ciri khusus yang terdapat di dalam citra lebih ditonjolkan[5]. Contoh-contoh operasi perbaikan citra:

1. Perbaikan kontras gelap/terang
2. Perbaikan tepian objek (*edge enhancement*)
3. Penajaman (*sharpening*)
4. Pemberian warna semu (*pseudocoloring*)
5. Penapisan derau (*noise filtering*)

b. Image Restoration

Image restoration atau pemugaran citra merupakan operasi yang bertujuan menghilangkan/meminimumkan cacat pada citra. Tujuan pemugaran citra hampir sama dengan operasi perbaikan citra. Bedanya, pada pemugaran citra penyebab degradasi gambar diketahui[5]. Contoh-contoh operasi pemugaran citra:

1. penghilangan kesamaran (*deblurring*).

2. penghilangan derau (*noise*)

c. *Image Compression*

Image compression atau pemampatan citra merupakan operasi yang dilakukan agar citra dapat direpresentasikan dalam bentuk yang lebih kompak sehingga memerlukan memori yang lebih sedikit[5].

d. *Image Segmentation*

Image segmentation atau segmentasi citra merupakan operasi yang bertujuan untuk memecah suatu citra ke dalam beberapa segmen dengan suatu kriteria tertentu. Jenis operasi ini berkaitan erat dengan pengenalan pola[5].

e. *Image Analysis*

Image analysis atau pengorakan citra merupakan operasi yang bertujuan menghitung besaran kuantitatif dari citra untuk menghasilkan deskripsinya. Teknik pengorakan citra mengekstraksi ciri-ciri tertentu yang membantu dalam identifikasi objek. Proses segmentasi kadangkala diperlukan untuk melokalisasi objek yang diinginkan dari sekelilingnya[5]. Contoh-contoh operasi pengorakan citra:

1. Pendeteksian tepi objek (*edge detection*)
2. Ekstraksi batas (*boundary*)
3. Representasi daerah (*region*)

f. *Image Reconstruction*

Image reconstruction atau rekonstruksi citra merupakan operasi yang bertujuan untuk membentuk ulang objek dari beberapa citra hasil proyeksi. Operasi rekonstruksi citra banyak digunakan dalam bidang medis. Misalnya beberapa foto rontgen dengan sinar X digunakan untuk membentuk ulang gambar organ tubuh[5].

2.2 Histogram of Oriented Gradients (HOG)

Histogram Oriented Gradients (HOG) merupakan salah satu deskriptor ciri yang digunakan dalam pengolahan citra untuk keperluan deteksi objek, teknik yang digunakan dengan menghitung keadaan arah perubahan tingkat keabuan (gradien) pada bagian lokal dari sebuah citra. Metode ini hampir sama dengan metode *edge orientation histograms*, *scale-invariant feature transform* and *shape contexts*, tapi perbedaannya pada komputasi menggunakan teknik membagi citra ke dalam beberapa bagian (*dense grid*) dan menggunakan normalisasi *overlapping* pada kontras lokal untuk meningkatkan akurasi[6].

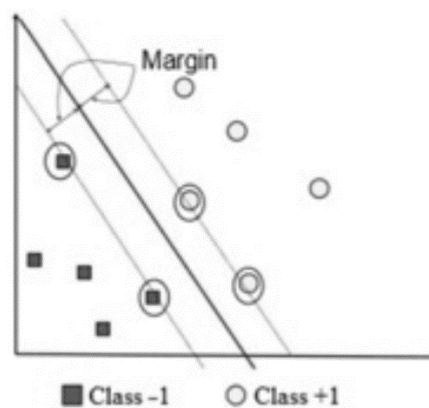
Metode HOG didasarkan pada evaluasi histogram lokal yang dinormalisasi dengan baik dari orientasi gradien gambar dalam kisi padat. Ide dasar dari metode HOG adalah tampilan dan bentuk objek lokal dapat dicirikan dengan baik oleh distribusi gradien intensitas lokal atau arah tepi, bahkan tanpa pengetahuan yang tepat tentang gradien atau posisi tepi yang sesuai. Dalam praktiknya metode HOG diimplementasikan dengan membagi jendela gambar menjadi wilayah spasial kecil ("*cell*"), untuk setiap *cell* mengumpulkan histogram arah gradien lokal 1-D atau berorientasi tepi terhadap piksel dari *cell*. Gabungan masukan histogram membentuk representasi. Untuk *invariant* yang lebih baik terhadap pencahayaan, bayangan, dan lain sebagainya, juga berguna untuk menormalkan kontras tanggapan lokal sebelum menggunakannya. Hal Ini dapat dilakukan dengan mengumpulkan ukuran histogram lokal "energi" pada wilayah spasial yang agak lebih besar ("blok") dan menggunakan hasilnya untuk menormalkan semua sel di dalam blok[1]. Histogram-histogram tersebut dijadikan fitur vektor yang akan diklasifikasikan menggunakan *linear Support Vector Machine* untuk menghasilkan keluaran berupa objek atau bukan objek[7][3].

Metode HOG dapat menganalisa apakah objek yang ditangkap oleh kamera adalah manusia atau bukan. Proses deteksi manusia pada sistem ini dilakukan per *frame* gambar dari hasil pencacahan *file* video[8].

2.3 Support Vector Machine (SVM)

SVM merupakan metode *learning machine* dengan tujuan menemukan *hyperplane* terbaik yang memisahkan dua buah *class* pada *input space*. Prinsip dasar SVM adalah *linear classifier*. Untuk mengecek apakah dalam masukan tersebut terdapat manusia atau tidak, digunakan SVM *classifier* untuk memisahkan manusia dan bukan manusia[9].

SVM adalah sistem pembelajaran yang menggunakan fungsi-fungsi linier dalam sebuah ruang fitur atau *feature space* yang memiliki dimensi tinggi. Pada ruang berdimensi tinggi, akan dicari *hyperplane* yang dapat memaksimalkan jarak antara kelas data. *Hyperplane* merupakan sebuah garis pemisah antara dua *class*. *Hyperplane* dapat ditemukan dengan mengukur jarak dan mencari titik maksimalnya. Jarak antara *hyperplane* akan dihitung dengan jarak terdekat dari masing-masing *class*. Garis pada Gambar 2.1 menunjukkan *hyperplane* yang terbaik. *Hyperplane* terbaik terletak pada tengah-tengah kedua *class*. Mencari lokasi *hyperplane* terbaik inilah merupakan inti dari proses pembelajaran pada SVM[9].



Gambar 2.1. Ilustrasi SVM untuk Menemukan *Hyperplane* terbaik yang memisahkan Kedua *Class -1* dan *+1*

2.4 Deteksi Manusia Menggunakan Metode HOG

Ekstraksi fitur pada kandidat objek menjadi kunci dalam melakukan deteksi.

Dalam penelitian ini, metode deteksi manusia dengan pendekatan klasifikasi menggunakan metode HOG dan SVM.

Ekstraksi fitur HOG diperoleh dari hasil perhitungan nilai gradien dalam daerah tertentu pada *frame* gambar. Setiap *frame* mempunyai karakteristik yang ditunjukkan oleh distribusi gradient. Karakteristik ini diperoleh dengan membagi *frame* ke dalam daerah kecil atau sel. Setiap sel disusun sebuah histogram dari sebuah gradien. Kombinasi dari histogram ini dijadikan sebagai deskriptor yang mewakili sebuah objek manusia[1].

Langkah pertama yang dilakukan adalah seleksi objek manusia dengan operasi HOG. Objek citra yang terdeteksi dihitung nilai deskriptornya menggunakan HOG. Perhitungan nilai gradien setiap piksel *frame* menggunakan Persamaan 2.1 dengan I adalah *frame* skala keabuan. I_x adalah matriks terhadap sumbu x dan I_y adalah matriks terhadap sumbu y. Matrik I_x dan I_y dihitung dengan Persamaan 2.2. Parameter D_x adalah mask $[-1 \ 0 \ 1]$ dan D_y adalah mask $[-1 \ 0 \ 1]^T$ yang masing-masing dihitung dengan cara konvolusi[10].

$$|G| = \sqrt{I_x^2 + I_y^2} \quad (2.1)$$

$$I_x = I * D_x, I_y = I * D_y \quad (2.2)$$

Dalam penghitungan histogram dan orientasi, hasil gradien ditransformasikan ke dalam koordinat sumbu dengan sudut antara 0 sampai 180^0 yang disebut orientasi gradien. Orientasi gradien (θ) dinyatakan dengan Persamaan 2.3. Setelah mendapat nilai gradien, maka proses selanjutnya adalah menentukan jumlah bin orientasi yang digunakan dalam histogram. Namun sebelum menghitung gradien, *frame* dibagi menjadi beberapa sel dan dikelompokkan menjadi ukuran lebih besar yang disebut blok[10].

$$\theta = \tan^{-1}(I_y/I_x) \quad (2.3)$$

Pada proses normalisasi blok, HOG memiliki jendela deteksi dengan ukuran 64×128 yang terdiri dari 8×8 piksel. Proses normalisasi blok dengan pemisah beberapa elemen vektor ini menghasilkan fitur HOG[1]. Nilai normalisasi fitur blok diperoleh dari Persamaan 2.4.

$$b = \frac{b}{\sqrt{b^2 + e}} \quad (2.4)$$

Fitur blok dinormalisasi untuk mengurangi efek perubahan kecerahan objek pada satu blok. Variabel b merupakan nilai blok fitur dan e merupakan bilangan bernilai kecil untuk mencegah pembagian dengan 0. Penentuan objek yang dideteksi manusia atau bukan menggunakan SVM [10].

2.5 OpenCV

OpenCV adalah singkatan dari *Open Computer Vision*, yaitu suatu *library* gratis yang dikembangkan oleh Intel Corporation yang di khususkan untuk melakukan *image processing*. Tujuannya adalah agar komputer mempunyai kemampuan yang mirip dengan cara pengolahan visual pada manusia[11].

OpenCV merupakan *open source library* untuk *computer vision*. *Library* ini dapat dijalankan dengan berbagai *operating system* yaitu Linux, Windows, dan mac OS X. OpenCV didesain untuk efektivitas dalam *computer vision* dan dengan fokus aplikasi *realtime* yang kuat. Salah satu tujuan dari OpenCV adalah untuk menyediakan infrastruktur *computer vision* yang memudahkan penggunaannya sehingga dapat digunakan dengan cepat membangun aplikasi *vision* yang berharga. *Library* dari OpenCV mengandung sekitar 500 lebih fungsi yang terdapat pada berbagai area, termasuk *factory product inspection, medical imaging, security, user interface, camera calibration, stereo vision, dan robotics*. OpenCV juga mengandung *Machine Learning Library* (MLL) yang lengkap. *Sublibrary* ini difokuskan pada pengenalan pola secara statistik dan clustering. MLL sangat berguna untuk pekerjaan *vision* yang merupakan inti dari misi OpenCV dan secara umum, MLL cukup berguna untuk berbagai masalah *machine learning*[8].

2.6 Protokol RTSP (Real Time Streaming Protocol)

RTSP atau *Real Time Streaming Protocol* merupakan protokol yang ada pada level aplikasi yang berfungsi untuk mengontrol pengiriman data dengan cara *real-time*. RTSP menggunakan dan menyediakan kerangka kerja yang *extensible* untuk mengontrol pengiriman data *on-demand* secara *real-time* seperti audio dan video. Data yang berasal dari pengirim dapat mencakup *live* data dan data yang disimpan[12].

RTSP merupakan protokol yang digunakan pada layanan *streaming*, teknologi *streaming* itu sendiri merupakan sebuah teknologi yang mampu mengompresi atau menyusutkan ukuran *file* audio dan video agar mudah ditransfer melalui jaringan data. Aplikasi layanan *streaming* dibagi menjadi dua, yaitu *on-demand* dan *live streaming*. Layanan *streaming on-demand* seperti halnya musik atau video yang dijalankan langsung dari media penyimpanan server yang diakses melalui *client*, sedangkan *live streaming* seperti halnya video atau audio yang disiarkan secara *broadcast* dan langsung pada saat itu juga[12].

2.7 Protokol UDP (User Datagram Protocol)

UDP atau *User Datagram Protocol* adalah salah satu jenis protokol pengiriman data. Melalui UDP, sebuah sistem komputer dapat mengirim pesan kepada komputer lain yang berada pada suatu jaringan yang sama tanpa perlu melalui proses komunikasi awal. UDP memiliki karakteristik utama yaitu “*connectionless*”, dimana pesan yang dikirimkan melalui UDP dapat sampai tanpa perlu proses negosiasi koneksi antara dua komputer *host* yang akan bertukar informasi. Selain itu, karakteristik UDP lainnya adalah “*unreliable*”, dimana semua pesan yang dikirim tidak mempunyai nomor urut atau pesan pemberitahuan. Jika selama pengiriman ada pesan yang hilang, maka protokol aplikasi yang berada di atas UDP harus memulihkan pesan tersebut.

Protokol UDP pada dasarnya hanya mengandung IP (*Internet Protocol*) dengan tambahan *header* singkat. Protokol UDP tidak melakukan sebuah proses kontrol alur data, kontrol kesalahan ataupun pengiriman ulang terhadap kesalahan sehingga hanya menyediakan *interface* ke protokol IP[13].

2.8 Raspberry Pi

Raspberry Pi atau biasa disebut Raspi adalah sebuah komputer papan tunggal (*single-board computer*) atau SBC yang memiliki ukuran cukup kecil hanya sebesar kartu kredit. Raspberry Pi mempunyai komponen yang hampir sama dengan komputer pada umumnya. Seperti CPU, GPU, RAM, Port USB, *Audio Jack*, HDMI, Ethernet, dan GPIO. Untuk media penyimpanan data dan sistem operasi

Raspberry Pi menggunakan Micro SD dengan kapasitas minimal 4 GB, sedangkan untuk sumber tegangan berasal dari micro USB *power* dan tegangan yang direkomendasikan sebesar 5V dengan minimal arus 700mA.

Raspberry Pi dapat digunakan layaknya komputer pada umumnya, seperti untuk mengetik dokumen atau untuk *browsing*. Tetapi Raspberry Pi juga dapat digunakan untuk membuat sebuah inovasi yang berkaitan dengan pemrograman. Jenis Raspberry Pi yang digunakan pada penelitian ini adalah Raspberry Pi 2 Model B.



Gambar 2.2. Raspberry Pi 2 Model B

2.9 Modul ESP32-CAM

ESP32-CAM merupakan modul yang memiliki fitur berupa *bluetooth*, *wifi*, kamera, dan slot microSD. Modul ini biasanya digunakan untuk *project IoT (Internet of Things)* yang membutuhkan fitur kamera. Modul ini memiliki lebih sedikit pin I/O dibandingkan modul ESP32 produk sebelumnya, yaitu ESP32 Wroom. Hal ini dikarenakan banyak pin yang digunakan secara internal untuk fungsi kamera dan fungsi slot kartu microSD. Selain itu, modul ini juga tidak memiliki *port USB* khusus untuk mengirim program dari *port USB* komputer. Jadi untuk memprogram modul ini harus menggunakan *USB to TTL* atau menambahkan modul tambahan berupa *downloader* khusus untuk ESP32-CAM.

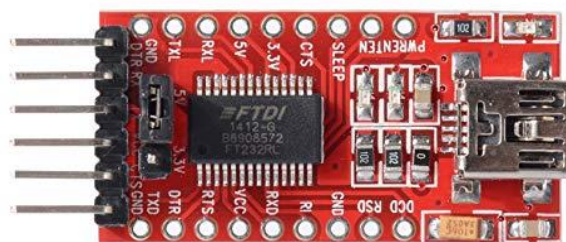


Gambar 2.3. Modul ESP32-CAM

Modul ESP32-CAM memiliki dua sisi dalam rangkaiannya. Pada bagian atas terdapat slot *microSD* dan kamera yang dapat dibongkar pasang, serta *flash* sebagai lampu tambahan untuk kamera jika diperlukan. Pada bagian belakang terdapat antena internal, konektor untuk antena eksternal, *pin male* untuk I/O dan ESP32S sebagai otaknya.

2.10 Modul FT232RL

Modul ini menggunakan *chip* produksi FTDI, yaitu FT232RL yang berfungsi sebagai *converter* data USB ke TTL/Serial. Modul ini biasa digunakan sebagai komunikator antara komputer dengan berbagai macam *device* yang tidak memiliki fungsi USB secara onboard seperti arduino pro mini, ESP32-CAM, *bluetooth* HC-05 dan lain sebagainya.



Gambar 2.4. Modul FT232RL

2.11 Modul ESP8266 NodeMCU

NodeMCU merupakan sebuah *platform* IoT (*Internet of Things*) yang bersifat *open source*. Terdiri dari perangkat keras berupa *System on Chip* ESP8266. NodeMCU menggabungkan ESP8266 ke dalam sebuah *board* yang menjalankan fungsi layaknya mikrokontroler dengan kemampuan dapat mengakses Wifi dan *chip* komunikasi USB *to* serial. Sehingga untuk memprogramnya hanya diperlukan kabel data USB yang menggunakan konektor *mini* USB.



Gambar 2.5. Modul ESP8266 NodeMCU

2.12 Motor Servo

Motor servo adalah sebuah aktuator putar (motor) yang dirancang dengan sistem kontrol umpan balik *loop* tertutup (servo), sehingga dapat diatur untuk menentukan posisi sudut dari poros *output* motor. Motor servo merupakan perangkat yang terdiri dari motor DC, serangkaian *gear*, rangkaian kontrol dan potensiometer. Serangkaian *gear* yang berada pada poros motor DC akan memperlambat putaran poros dan meningkatkan torsi motor servo, sedangkan potensiometer dengan perubahan resistansinya saat motor berputar berfungsi sebagai penentu batas posisi putaran poros motor servo. Penggunaan sistem kontrol *loop* tertutup pada motor servo berguna untuk mengontrol gerakan dan posisi akhir dari poros motor servo. Jenis motor servo yang digunakan pada penelitian ini adalah Tower Pro SG90.



Gambar 2.6. Motor Servo Tower Pro SG90

2.13 Spesifikasi Perangkat Keras

Berikut adalah tabel spesifikasi perangkat keras yang dibutuhkan untuk membangun alat pada penelitian ini:

Tabel 2.1. Tabel Spesifikasi Perangkat keras

No	Perangkat Keras	Jumlah	Spesifikasi
1	Raspberry Pi 2 Model B	1	<ul style="list-style-type: none"> • 900MHz quad-core ARM Cortex-A7 CPU • 1GB RAM • 100 Base Ethernet • 4 USB ports • 40 GPIO pins • Full HDMI port • Combined 3.5mm audio jack and composite video • Camera interface (CSI) • Display interface (DSI) • Micro SD card slot • VideoCore IV 3D graphics core
2	Modul ESP32-CAM	1	<ul style="list-style-type: none"> • 802.11b/g/n Wi-Fi

			<ul style="list-style-type: none"> • Bluetooth 4.2 <i>with</i> BLE • UART, SPI, I2C <i>and</i> PWM interfaces • Clock speed up to 160 MHz • Computing power up to 600 DMIPS • 520 KB SRAM <i>plus</i> 4 MB PSRAM • Supports WiFi Image Upload • Multiple Sleep modes • Firmware Over the Air (FOTA) upgrades possible • 9 GPIO ports • Built-in Flash LED • Camera resolution 2MP
3	Modul FT232RL	1	<ul style="list-style-type: none"> • Chipset FT232RL • Switch jumper 3.3V or 5V • USB over current protection self restore fuse 500mA • Pin: TX, RX, VCC, GND, CTS, DTR • Dimension : 36 mm x 17 mm • Interface: Mini USB
4	Modul ESP8266 NodeMCU	1	<ul style="list-style-type: none"> • ESP8266 ESP-12 • USB-TTL converter CH340 • Operating voltage: 3.3v • Input voltage: 5V • GPIO Pin 11 • 1 Analog Pin I/O

			<ul style="list-style-type: none"> • <i>PWM interfaces</i> • <i>Flash Memory 4MB</i> • <i>RAM 128Kb</i> • <i>Clock Frequency 80MHz</i> • <i>USB Micro USB</i> • <i>Dimension: 47mm x 31mm</i> • <i>Weight: 5g</i>
5	Motor Servo MG966R	2	<ul style="list-style-type: none"> • <i>Weight: 9g</i> • <i>Dimension: 22mm x 11.5mm x 22.5mm</i> • <i>Stall torque: 1.8 kg/cm (4.8 V)</i> • <i>Degree: max 180 degrees servo rotation</i> • <i>Operating speed: 0.1 s/60° (4.8 V)</i> • <i>Operating voltage: 4.8 V</i> • <i>Dead bandwidth: 1us</i>