

BAB II

LANDASAN TEORI

2.1 Python

Python adalah bahasa pemrograman interpretatif multiguna dengan filosofi perancangan yang berfokus pada tingkat keterbacaan kode. Python diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif. Python bisa dibilang bahasa pemrograman dengan tujuan umum yang dikembangkan secara khusus untuk membuat *source code* mudah dibaca. Python juga memiliki *library* yang lengkap sehingga memungkinkan *programmer* untuk membuat aplikasi yang mutakhir dengan menggunakan *source code* yang tampak sederhana (Ljubomir Perkovic, 2012).

2.2 OpenCV

OpenCV (Open Source Computer Vision Library) adalah sebuah pustaka perangkat lunak yang ditujukan untuk pengolahan citra dinamis secara *real time*, yang pengembangannya diawali oleh Intel, dan sekarang didukung oleh Willow Garage dan Itseez. OpenCV dirilis dibawah lisensi permisif BSD yang lebih bebas dari pada GPL, dan memberikan kebebasan sepenuhnya untuk dimanfaatkan secara komersil tanpa perlu mengungkapkan kode sumbernya. OpenCV juga memiliki antar muka yang mendukung bahasa pemrograman C++, C+, Python dan Java, termasuk untuk sistem operasi Windows, Linux, dan Android. OpenCV didesain untuk efisiensi dalam komputasi dan difokuskan pada aplikasi *real time*.

2.3 Citra Digital

Citra digital adalah gambar dua dimensi yang dihasilkan dari gambar analog dua dimensi yang kontinu menjadi gambar diskrit melalui proses *sampling*. Dalam citra digital terdapat sebuah larik (array) yang berisi nilai-nilai *real* maupun kompleks yang di representasikan dengan deretan bit tertentu. Sebuah citra digital dapat diwakili oleh sebuah matriks dua dimensi $f(x,y)$ yang terdiri dari M kolom dan N baris, dimana perpotongan antara

kolom dan baris disebut piksel (*pixel = picture element*) atau elemen terkecil dari sebuah citra.

Berdasarkan rumus di atas, suatu citra $f(x,y)$ dalam fungsi matematis dapat dituliskan sebagai berikut :

$$f(x,y) \approx \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{bmatrix}$$

$$\leq x \leq M-1$$

$$0 \leq y \leq N-1$$

$$0 \leq f(x,y) \leq G-1$$

Dimana :

M = Jumlah piksel baris (row) pada array citra

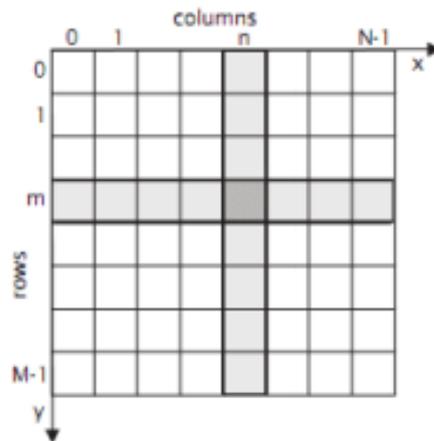
N = Jumlah piksel kolom (column) pada array citra

G = Nilai skala keabuan (graylevel)

Besarnya nilai M,N dan G pada umumnya merupakan perpangkatan dari dua.

$$M = 2^m ; N = 2^n ; G = 2^k$$

Dimana nilai m, n dan k adalah bilangan bulat positif. Interval (0,G) disebut skala keabuan (grayscale). Besar G tergantung pada proses digitalisasinya. Biasanya keabuan 0 (nol) menyatakan intensitas hitam dan 1 (satu) menyatakan intensitas putih. Untuk citra 8 bit, nilai G sama dengan $2^8 = 256$ warna (derajat keabuan).



Gambar 1. Representasi citra digital dalam 2 dimensi (Bernd, 2000)

Proses perubahan citra analog menjadi citra digital dinamakan dengan digitasi. Digitasi adalah proses mengubah gambar, tek, atau suara dari benda yang dapat dilihat ke dalam data elektronik dan dapat disimpan serta diproses untuk keperluan lainnya. Citra digital pada komputer dipetakan menjadi bentuk grid atau elemen piksel berbentuk matriks 2 dimensi. Setiap piksel-piksel tersebut memiliki angka yang mempresentasikan channel warna. Angka pada setiap disimpan secara berurutan oleh komputer dan sering dikurangi untuk keperluan kompresi maupun pengolahan tertentu.

2.3.1 Tipe Citra Digital

Citra digital dapat dikategorikan dalam beberapa jenis, yaitu :

1. Citra Biner

Citra biner (*binary image*) adalah citra yang hanya mempunyai nilai derajat keabuan yaitu hitam dan putih. Piksel-piksel objek bernilai 1 dan piksel-piksel latar belakang bernilai 0. Piksel yang bernilai 0 melambangkan warna putih dan piksel yang bernilai 1 melambangkan warna hitam pada saat menampilkan citra.

2. Citra Keabuan

Citra keabuan (*grayscale*) adalah citra yang disetiap pikselnya mengandung satu layer dimana nilai intensitasnya berada pada nilai 0 (hitam) – 255 (putih).

3. Citra Warna

Citra warna (RGB) adalah citra digital yang memiliki informasi warna pada setiap pikselnya. Sistem pewarnaan citra warna ada beberapa macam seperti RGB, CMYK, HSV, dll. RGB adalah model warna yang terdiri dari merah, hijau dan biru yang digabungkan dalam membentuk suatu susunan warna yang luas. Untuk monitor komputer, nilai rentangnya paling kecil = 0 dan paling besar = 255.

2.4 Pengolahan Citra Digital

Pengolahan citra digital (*Digital Image Processing*) adalah sebuah disiplin ilmu yang mempelajari tentang teknik- teknik mengolah citra. Citra yang dimaksud disini adalah gambar diam (foto) maupun gambar bergerak (video). Sedangkan digital disini mempunyai maksud bahwa pengolahan citra/gambar dilakukan secara digital menggunakan komputer[3]. Dalam konteks yang lebih luas, pengolahan citra digital mengacu pada pemrosesan setiap data 2 dimensi[4]. Agar dapat diolah dengan komputer digital, maka suatu citra harus dipresentasikan secara numerik dengan nilai-nilai diskrit. Representasi dari fungsi kontinyu menjadi nilai-nilai diskrit disebut digitalisasi citra. Tujuan pengolahan citra ini adalah memperbaiki kualitas suatu citra sehingga dapat diinterpretasi dengan mudah oleh manusia atau sebuah mesin (komputer).

Proses perubahan citra menjadi citra digital dinamakan digitasi. Digitasi merupakan proses mengubah sebuah gambar, teks, atau suara dari benda yang dapat dilihat ke dalam data elektronik dan dapat disimpan serta diproses untuk keperluan lainnya. Dalam konteks yang lebih luas, pengolahan citra digital lebih mengacu pada pemrosesan setiap dua data dimensi. Pengolahan citra digital adalah sebuah disiplin ilmu yang

mempelajari tentang bagaimana teknik pengolahan sebuah citra. Citra yang dimaksud disini adalah sebuah gambar diam (foto) maupun gambar bergerak (video). Berikut merupakan teknik pengolahan citra digital yang digunakan dalam penelitian ini.

1. Resizing

Resizing merupakan proses mengubah ukuran panjang dan lebar citra digital pada arah vertikal dan/atau horizontal

2. Mean Substraction

Mean subtraction digunakan untuk membantu mengatasi permasalahan perubahan iluminasi pada gambar input yang ada pada dataset. Untuk menangani variasi intensitas dan normalisasi, dilakukan dengan cara menghitung nilai piksel rata-rata pada dataset pelatihan dan mengurangnya dari setiap gambar selama pelatihan. Tahap ini dilakukan inisialisasi 3 buah variabel yaitu : μ_R , μ_G dan μ_B . Lalu, disaat gambar siap melalui jaringan yang ada, dilakukan proses mengurangi nilai rata-rata μ untuk setiap *input channel* pada citra input :

$$R = (R - \mu_R)$$

$$G = (G - \mu_G)$$

$$B = (B - \mu_B)$$

3. Scaling

Scaling terhadap standar deviasi oleh training *set* dibutuhkan pada gambar agar dapat diproses oleh jaringan *Machine Learning* berdasarkan citra masukan serta dapat mengkoordinasikan skala *bounding box* yang diproses pada ukuran citra input, display maupun citra yang akan disimpan. *Scaling factor*, σ , yang ditambahkan dalam sebuah normalisasi:

$$R = (R - \mu_R) / \sigma$$

$$G = (G - \mu_G) / \sigma$$

$$B = (B - \mu_B) / \sigma$$

2.5 *Machine Learning*

Istilah *Machine Learning* pertama kali didefinisikan oleh Arthur Samuel ditahun 1959. Menurut Arthur Samuel, *Machine Learning* adalah

salah satu bidang ilmu komputer yang memberikan kemampuan pembelajaran kepada komputer untuk mengetahui sesuatu tanpa pemrogram yang jelas. Istilah *Machine Learning* pada dasarnya adalah proses komputer untuk belajar dari data. Tanpa adanya data, komputer tidak akan bisa belajar apa-apa. Semua pengetahuan *Machine Learning* pasti akan melibatkan data. Data bisa saja sama, akan tetapi algoritma dan pendekatannya berbeda-beda untuk mendapatkan hasil yang optimal.

Machine Learning merupakan salah satu cabang dari disiplin ilmu Kecerdasan Buatan (*Artificial Intelligence*) yang membahas mengenai pembangunan sistem yang berdasarkan pada data[5]. Dalam pembelajaran *Machine Learning*, terdapat beberapa skenario-skenario. Seperti:

1. *Supervised Learning*

Penggunaan skenario *Supervised Learning*, pembelajaran menggunakan masukan data pembelajaran yang telah diberi label. Setelah itu membuat prediksi dari data yang telah diberi label.

2. *Unsupervised Learning*

Penggunaan skenario *Unsupervised Learning*, pembelajaran menggunakan masukan data pembelajaran yang tidak diberi label. Setelah itu mencoba untuk mengelompokkan data berdasarkan karakteristik-karakteristik yang ditemui

3. *Reinforcement Learning*

Pada skenario *Reinforcement Learning* fase pembelajaran dan tes saling dicampur. Untuk mengumpulkan informasi pembelajar secara aktif dengan berinteraksi ke lingkungan sehingga untuk mendapatkan balasan untuk setiap aksi dari pembelajar

2.6 *Image Stitching*

Image Stitching adalah proses penggabungan beberapa gambar foto yang tumpang tindih dengan bidang untuk menghasilkan panorama atau gambar resolusi tinggi. Umumnya dilakukan melalui penggunaan perangkat lunak komputer, sebagian besar pendekatan untuk *Image Stitching*

memerlukan kecocokan yang paling dekat diantara gambar dan tumpang tindih untuk menghasilkan gambar yang halus. *Image Stitching* juga dikenal sebagai mosaicing[6].

2.7 Deteksi Objek

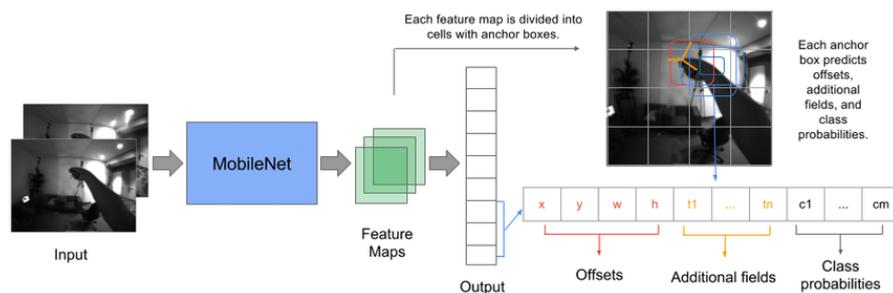
Deteksi objek bertujuan untuk memisahkan objek atau foreground dari citra latar. Biasanya untuk melakukan deteksi objek, akan dibuat sebuah model latar pada sebuah citra bergerak berdasarkan nilai pixel yang masuk seiring waktu[7]. Untuk mendeteksi sebuah objek harus memberikan informasi yang efektif dan efisien seperti untuk mendeteksi gambar dan citra digital. Disaat manusia melirik pada sebuah gambar, otak manusia dapat langsung mengenali objek yang berada pada gambar, letak objek tersebut, dan kondisi interaksi yang terjadi. Sistem visual manusia cepat dan akurat, memungkinkan kita untuk melakukan tugas-tugas kompleks. Algoritma deteksi objek yang cepat dan akurat akan memungkinkan komputer dapat melakukan hal serupa hingga berpotensi untuk menyelesaikan tugas secara umum. Deteksi objek dalam digital *image processing* adalah suatu proses yang digunakan untuk menentukan keberadaan objek tertentu di dalam citra.

Proses deteksi tersebut dapat dilakukan dengan berbagai macam metode yang umumnya melakukan pembacaan fitur-fitur dari seluruh objek pada citra input. Fitur dari objek pada citra input tersebut akan dibandingkan dengan fitur dari model yang digunakan atau template. Hasil perbandingan tersebut dapat digunakan untuk menentukan apakah suatu objek terdeteksi sebagai template yang dimaksud atau tidak. Sistem deteksi objek perlu melatih dan menguji dataset dengan bounding box dan diberi label untuk kelas per setiap objek untuk proses pengenalan. Adapun metode deteksi objek yang digunakan pada skripsi ini :

2.7.1 Mobile Net

MobileNet, merupakan salah satu arsitektur *Convolutional Neural Network* (CNN) yang dapat digunakan untuk mengatasi kebutuhan akan sumber komputasi berlebih. Para peneliti dari Google membangun MobileNet atas kebutuhan arsitektur CNN yang

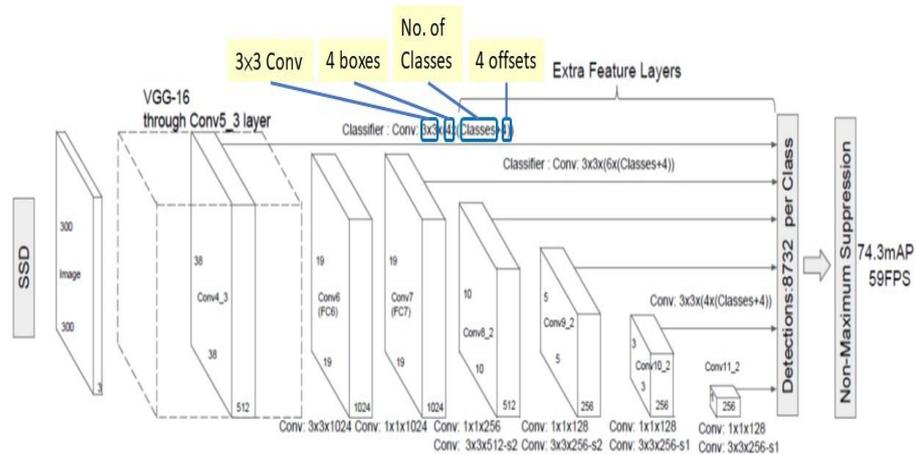
dapat digunakan untuk ponsel maupun sistem tertanam. Perbedaan mendasar antara arsitektur MobileNet dan arsitektur CNN pada umumnya adalah penggunaan lapisan atau layer konvolusi dengan ketebalan filter yang sesuai dengan ketebalan dari input image. MobileNet membagi konvolusi menjadi *depthwise convolution* dan *pointwise convolution*. Arsitektur MobileNet memanfaatkan *Batch Normalization* (BN) dan *Rectified-Linear unit* (ReLU) untuk *depthwise convolution* dan *pointwise convolution*. MobileNet dibangun di atas arsitektur jaringan yang efisien dengan menggunakan konvolusi yang dapat dipisahkan secara mendalam untuk menghasilkan *Deep Neural Network* yang ringan. Alur kerja MobileNet dapat dilihat pada Gambar 2. Alur Kerja MobileNet



Gambar 2. Alur Kerja MobileNet

2.7.2 Single Shot MultiBox Detector (SSD)

Perbedaan utama antara pelatihan detektor klasik seperti R-CNN dan pelatihan SSD adalah bahwa detektor objek klasik menggunakan proposal wilayah dan metode SSD menggunakan data *groundtruth* yang harus ditetapkan untuk hasil yang ditentukan dalam rangkaian hasil detektor yang jelas. Setelah penugasan ini ditentukan, *loss function* dilakukan dari ujung ke ujung. Metode SSD melakukan pencocokan terhadap objek dengan *default bounding box* melalui beragam skala dan rasio untuk setiap *feature map location*. Setiap elemen dari *feature map* memiliki sejumlah kotak yang saling berhubungan. Setiap bounding box dengan IoU lebih besar dari 0,5 dianggap cocok[8].



Gambar 3. Arsitektur Single Shot MultiBox (SSD)

Single Shot Detector juga mempunyai sebuah arsitektur dimana seperti yang dilihat dari Gambar 6 di atas, arsitektur SSD dibangun di atas arsitektur VGG-16, tetapi membuang lapisan yang terhubung sepenuhnya.

Alasan VGG-16 digunakan sebagai jaringan dasar adalah karena kinerjanya yang kuat dalam tugas-tugas klasifikasi gambar berkualitas tinggi dan popularitasnya untuk masalah-masalah di mana transfer membantu dalam meningkatkan hasil. Alih-alih VGG asli sepenuhnya terhubung lapisan, satu set lapisan konvolusional tambahan (dari conv6 dan seterusnya) ditambahkan, sehingga memungkinkan untuk mengekstrak fitur pada berbagai skala dan secara progress mengurangi ukuran *input* ke setiap lapisan berikutnya.

2.8 Pemetaan objek

Pemetaan adalah suatu proses penyajian informasi muka bumi yang fakta (dunia nyata), baik bentuk permukaan buminya maupun sumbu alamnya, berdasarkan skala peta, sistem proyeksi peta,serta simbol-simbol dari unsur muka bumi yang disajikan (Jatmiko,2011). Pemetaan digital merupakan suatu dalam pembuatan peta, baik untuk keperluan pencetakan maupun dalam format peta digital (Ronny,2011).

Menurut Dickinson (1975) yang dikutip oleh Hanum (2013), berapa alasan suatu data dapat dipetakan antara lain :

1. Melalui peta dapat menimbulkan daya tarik yang lebih besar terhadap objek yang ditampilkan.
2. Melalui peta dapat memperjelas, menyederhanakan, dan menerangkan suatu aspek yang dipentingkan.
3. Melalui peta dapat menonjolkan pojok-pojok batasan dalam tulisan atau pembicaraan. Melalui peta dapat dipakai sebagai sumber data bagi yang berkepentingan.
4. Peta sebagai alat komunikasi antara membuat peta dengan pengguna dimana akan memudahkan dalam penyampaian informasi.