

## **BAB 2**

### **TINJAUAN PUSTAKA**

#### **2.1 Computer Vision**

Merekonstruksi suatu bentuk dari rupa nyata (3 dimensi) ke bentuk dua dimensi merupakan hal yang dilakukan oleh keilmuan *computer vision* [1]. Dengan kata lain ilmu ini mengajarkan komputer untuk memahami suatu objek nyata dan dapat di manipulasi dengan pendekatan digital, seperti halnya implementasi pada kehidupan sehari-hari antara lain deteksi kematangan buah, aplikasi deteksi untuk robot, deteksi manusia, maupun deteksi pengenalan multi objek bergerak[11].

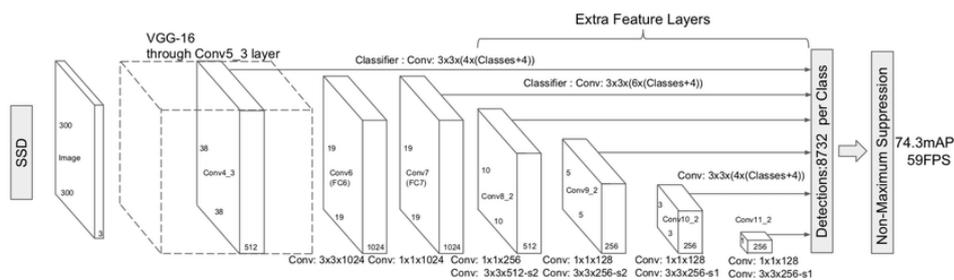
#### **2.2 Deep Learning (Deep Neural Network)**

Deep learning merupakan cabang dari machine learning yang terinspirasi dari korteks manusia dengan menerapkan jaringan syaraf buatan (neural network) yang memiliki banyak lapisan (hidden layer)[2]. Algoritma ini kiannya merupakan pengembangan dari salah satu algoritma machine learning, yaitu artificial intelligen (ANN). Banyak lapisan tersebut merupakan algoritma atau metode yang melakukan klasifikasi perintah yang diinput hingga menghasilkan output.

Untuk penjelasan lebih spesifiknya, neural network ini menggunakan apa yang telah dipelajarinya agar terlatih untuk bekerja di dunia digital, hal hal yang dikerjakan seperti untuk mengenali suatu gambar, mengenali kata kata yang diucapkan, penyakit kelainan darah, atau memberikan rekomendasi video yang akan ditonton selanjutnya, dalam dunia AI hal ini dikenal sebagai inference[7]. Untuk mendapatkan suatu inference dibutuhkan suatu training terlebih dahulu. Dengan begitu caranya mendapatkan sebagian besar pengetahuan, dilatih (training) terlebih dahulu lalu (inference) menyimpulkan suatu pengetahuan kemudian.

## 1.2 SSD (Single Shot Multibox Detector)

Pendekatan SSD didasarkan pada jaringan konvolusional umpan maju (feed-forward) yang menghasilkan kumpulan kotak pembatas dan skor ukuran tetap untuk keberadaan keadaan kelas objek dalam kotak tersebut. SSD memiliki dua komponen yaitu *Backbone Model* dan *SSD Head*. *Backbone model* adalah jaringan klasifikasi citra yang telah dilatih sebelumnya sebagai pengestrak fitur. *SSD Head* adalah kumpulan lapisan konvolusi yang ditambahkan pada *backbone* dan outputnya ditafsirkan sebagai kotak pembatas dan kelas objek di lokasi spasial aktivasi lapisan terakhir. SSD melakukan pembagian gambar bagaikan urutan sel-sel, dan setiap sel akan bertanggung jawab untuk mendeteksi objek pada wilayah gambar tersebut.



**Gambar 1** Model single shot multibox detector (SSD)[9]

## 1.3 Open Computer Vision (OpenCV)

OpenCV (*Open Source Computer Vision Library*) adalah perpustakaan komputer open source visi dan pembelajaran perangkat lunak [3]. Untuk penelitian ini modul opencv yang digunakan yaitu OpenCV “DNN” yang gunanya untuk menguji model *deep learning* pada framework yang digunakan untuk melatih model yaitu caffe model.

### 1.1 Caffe

Caffe adalah kerangka pembelajaran mendalam (Deep Learning framework) yang dikembangkan oleh BAIR. Dengan menggunakan modul kamera, kita dapat mengumpulkan data dan mendeteksi adanya hambatan dan

mengklasifikasikannya ke dalam kelas yang ditentukan dengan bantuan model yang sudah dilatih sebelumnya[5].

## 1.2 Flask

Flask adalah framework python yang tersedia di bawah lisensi BSD. Flask terinspirasi dari framework Sinatra Ruby. Flask tergantung pada toolkit WSGI Werkzeug dan template Jinja2.

Ide utama dibalik Flask adalah membantu membangun fondasi aplikasi web yang kokoh. Dari sana, developer dapat menggunakan ekstensi apapun yang mungkin developer butuhkan. Flask dipilih untuk setiap dan semua proyek. Pilihan ini ditujukan untuk proyek web yang tidak cocok dengan Django.

Desain Flask yang ringan dan modular membuatnya mudah disesuaikan dengan kebutuhan developer.

Sejak diluncurkan pada tahun 2010, Flask telah diperbarui sebanyak 27 kali. Sampai saat ini, itu tetap menjadi kerangka kerja Python yang paling banyak berkembang. Namun, banyak ekstensi Flask yang tidak lagi didukung: dokumentasinya sudah usang dan tidak lagi dikembangkan. Jika Anda mengalami ini, Anda harus ke Google untuk menemukan plugin yang didukung secara aktif dengan fungsi serupa.

## 1.3 Pre-Trained Model

*Pre-trained* model merupakan salah satu langkah mengambil model yang sudah diuji sebelumnya oleh orang lain kedalam data yang cukup besar seperti *dataset ImageNet*. *Pre-trained model* merupakan *dataset* yang sudah diakui sehingga kualitas *pre-trained* model sudah sangat baik[10].

## 1.4 VNC (Virtual Network Computing)

VNC adalah singkatan dari Komputasi Jaringan Virtual. Ini adalah sistem berbagi layar lintas platform yang dibuat untuk mengontrol komputer lain dari jarak jauh. Ini berarti bahwa layar komputer, keyboard, dan mouse dapat digunakan dari jarak jauh oleh pengguna jarak jauh dari perangkat sekunder seolah-olah mereka duduk tepat di depannya.