

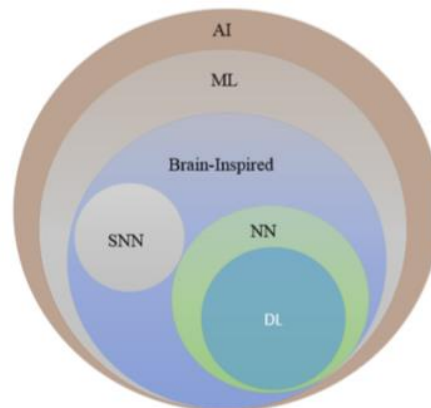
BAB III

Landasan Teori

Bab ini akan membahas berbagai teori yang melandasi penulis dalam penelitian yang akan dilakukan.

3.1. *Deep Learning*

Deep learning merupakan suatu kelas dari teknik pembelajaran mesin yang memanfaatkan banyak lapisan dalam pemrosesan informasi untuk ekstraksi dan transformasi, serta untuk analisis pola dan klasifikasi terutama untuk memahami data seperti gambar, suara, dan teks. *Deep learning* sendiri berada dititik potong antara beberapa area penelitian seperti jaringan syaraf, kecerdasan buatan, pemodelan grafis, optimisasi, pengenalan pola, dan pemrosesan sinyal. Tiga alasan penting yang mendasari populernya *deep learning* saat ini adalah karena meningkatnya kemampuan perangkat keras untuk pemrosesan komputasi, bertambah besarnya ukuran data yang digunakan untuk melatih jaringan, dan semakin berkembangnya penelitian terkait pembelajaran mesin dan pemrosesan data. Beberapa area aplikasi dari *deep learning* antara lain adalah *computer vision*, pengenalan suara, pemrosesan bahasa alami, dan area lain yang berhubungan. Ruang lingkup dari *deep learning* dapat dilihat pada ilustrasi Gambar 3.1. berikut.



Gambar 3.1. Ruang Lingkup *Deep Learning* (Vivienne, et al., 2017)

3.2. *Computer Vision*

Computer Vision merupakan salah satu bidang ilmu komputer yang mempelajari tentang bagaimana cara agar komputer dapat memahami secara mendalam citra digital maupun video. Dalam sudut pandang ilmu teknik, *computer vision* mempunyai tujuan utama untuk membangun sistem yang mampu melakukan tugas yang sistem pengelihatan manusia dapat lakukan atau bahkan dapat melebihi sistem pengelihatan manusia (Huang, 1996). Citra digital yang kompleks diharapkan dapat dipahami secara baik oleh komputer sehingga data yang didapatkan dari citra digital tersebut dapat diperoleh dan dianalisis untuk dapat menghasilkan informasi yang dapat diolah lebih lanjut oleh komputer. Aplikasi dari *computer vision* sendiri sangatlah luas, antara lain dapat digunakan untuk membantu manusia dalam identifikasi suatu tugas tertentu, proses pengendalian, deteksi suatu kejadian, proses interaksi komputer-manusia, pemodelan objek, navigasi, dan pengorganisasian informasi.

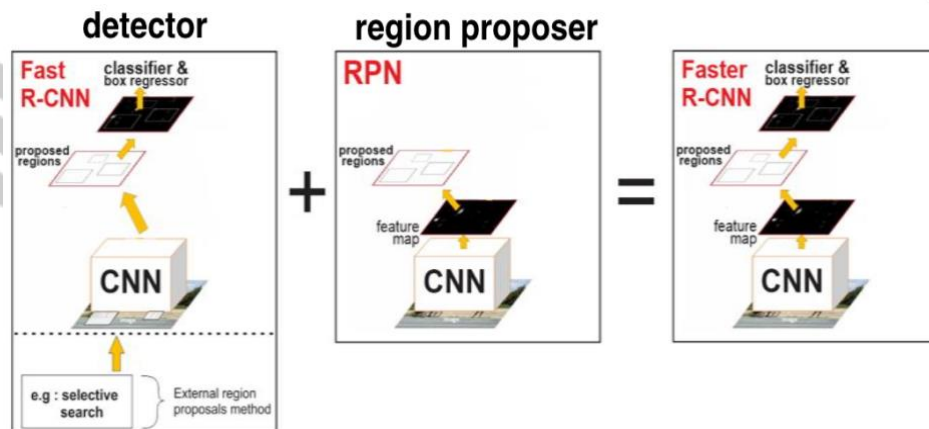
3.3. *Convolutional Neural Network (CNN)*

Convolutional Neural Network (CNN) merupakan salah satu kelas dari *deep neural network*. Seperti *neural network* pada umumnya, CNN terdiri dari banyak neuron yang mempunyai bobot dan bias yang dapat dilatih. Setiap neuron pada CNN dapat menerima beberapa *input*, melakukan operasi komputasi, menghasilkan *output*, dan terhubung ke neuron-neuron lainnya. Nama konvolusi sendiri merupakan operasi aljabar linear yang mengkalikan matriks dari filter pada citra yang akan diproses. Proses ini disebut dengan lapisan konvolusi dan merupakan salah satu jenis dari banyak lapisan yang bisa dimiliki dalam satu jaringan CNN. Meskipun begitu lapisan konvolusi merupakan lapisan utama yang paling penting untuk digunakan. Jenis lapisan lain yang biasa digunakan adalah *Pooling Layer*, *Normalization Layer*, dan *Fully Convolutional (FC) Layer*. Hal yang menjadi ciri khusus dari CNN adalah *input* awal pada jaringannya secara eksplisit selalu diasumsikan sebagai citra digital, sehingga

parameter pada arsitektur jaringannya dapat lebih mudah disesuaikan dan lebih efisien (Stanford University, 2016).

3.4. Region-based Convolutional Neural Network (R-CNN)

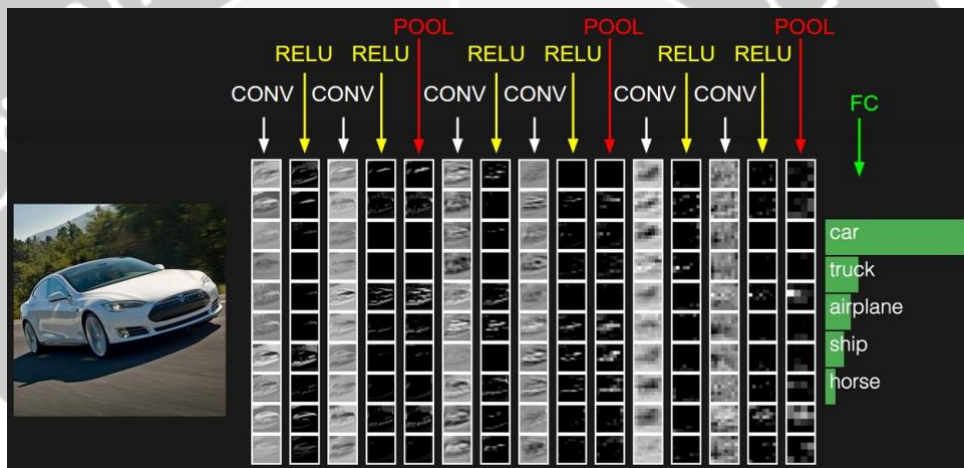
R-CNN merupakan metode pendeteksian objek yang masuk kedalam ranah *computer vision* yang berbasis jaringan konvolusi atau CNN. R-CNN sendiri awalnya pada tahun 2015 dibuat sebagai metode pendeteksian objek yang mengkombinasikan algoritma *Region Proposal Network* (RPN) dan *Convolutional Neural Network* (CNN) (Girshick et al., 2014). Seiring berjalannya waktu metode R-CNN ini terus dikembangkan untuk meningkatkan performa baik kecepatan maupun akurasi dalam pendeteksian objek. Akan tetapi R-CNN maupun *Fast R-CNN* masih mempunyai kekurangan, salah satunya adalah adanya *bottleneck* pada RPN yang tidak dapat menyamai kecepatan komputasi pada CNN. *Faster R-CNN* pun kemudian hadir dengan optimisasi pada penggunaan fitur konvolusional untuk mempercepat proses RPN sehingga dapat mengurangi terjadinya *bottleneck* (Ren et al., 2015). Gambaran mengenai arsitektur dari *Faster R-CNN* dapat dilihat pada Gambar 3.2. berikut ini.



Gambar 2.2. Arsitektur *Faster R-CNN* (CS231n Stanford University, 2016)

3.5. Model Jaringan CNN

Sebuah jaringan CNN terbentuk dari beberapa lapisan atau *layers*. Tiap *layer* tersebut berfungsi untuk menerima sebuah *input* untuk kemudian diproses dengan fungsi tertentu yang telah didefinisikan pada *layer* tersebut untuk kemudian menjadi *output*. Tumpukan berbagai *layer* pada suatu arsitektur jaringan CNN inilah yang disebut dengan model jaringan CNN. Contoh model jaringan CNN sederhana untuk deteksi objek dapat dilihat pada Gambar 3.3. berikut ini.



Gambar 3.3. Arsitektur Model Jaringan CNN Sederhana untuk Deteksi Objek (CS231n Standford University, 2016)

Beberapa model jaringan CNN untuk deteksi objek telah dipatenkan namanya. Beberapa yang paling dikenal antara lain adalah AlexNet, ZF Net, GoogLeNet, VGGNet, dan ResNet. Model jaringan tersebut umumnya sudah teruji kemampuannya dalam proses deteksi objek.

3.6. Framework Deep Learning

Dengan sulitnya menulis kode program untuk GPU yang efisien dan performa yang maksimal maka para peneliti perlu mendefinisikan struktur alur kerja untuk menghindari penulisan kode baru untuk setiap model atau algoritma baru. Hal tersebut dapat dilakukan dengan membuat sebuah library perangkat lunak yang dapat melakukan operasi yang memerlukan performa tinggi seperti

operasi konvolusi dan perkalian matriks, untuk kemudian didefinisikan dalam bentuk model kerangka kerja yang dapat memanggil operasi melalui library tersebut (Goodfellow et al., 2016). Pendekatan tersebut juga dapat memudahkan dukungan untuk beragam jenis perangkat keras yang berbeda. Sebagai contoh, suatu program yang dibangun dengan kerangka kerja Theano dapat berjalan pada CPU maupun GPU tanpa perlu mengubah kode program tersebut. Saat penelitian ini dilakukan, telah ada beberapa kerangka kerja untuk *deep learning* seperti Theano, TensorFlow, Torch, Microsoft Cognitive Toolkit, Neon, Chainer, dan lain sebagainya yang juga menyediakan fitur serupa. Semua kerangka kerja tersebut mempunyai tujuan umum untuk memproses komputasi yang intensif, jumlah data yang besar, dan melibatkan proses iterasi yang cukup banyak.

Tiap kerangka kerja *deep learning* tersebut juga terus dikembangkan dengan mengoptimalkan berbagai aspek yang berbeda seperti dukungan untuk berbagai macam algoritma yang berbeda, arsitektur *deep learning* yang beragam, kemudahan dalam penulisan dan pemodelan kode, serta peningkatan kecepatan komputasi untuk menarik pengguna agar menggunakan kerangka kerja tersebut.

3.7. Tensorflow

Tensorflow merupakan *framework machine learning* yang dibangun secara *opensource* dengan dukungan dari perusahaan raksasa teknologi Google. Secara umum Tensorflow menyediakan antarmuka untuk mengekspresikan algoritma-algoritma *machine learning* secara fleksibel dan dapat berjalan pada beragam sistem (Abadi et al., 2016). Pada aplikasinya dalam pengenalan objek Tensorflow telah mendukung berbagai macam algoritma, termasuk proses *training* dan *inference* untuk berbagai model *deep neural network*. Tensorflow juga termasuk sebagai *library* untuk *data science* paling populer dengan jumlah pengembang dan dukungan komunitas yang besar (M. Li, 2017).