

BAB 2

TINJAUAN PUSTAKA

1. Deep Learning

Deep Learning merupakan cabang pengetahuan dari kecerdasan buatan yang berfokus pada pembuatan model neural network besar yang mampu membuat keputusan yang akurat berdasarkan data. Deep Learning merupakan metode yang cocok untuk konteks dimana data yang diproses merupakan data kompleks dan ketersediaan data latih yang besar. Pengambilan keputusan adalah hal yang krusial. salah satu cara untuk mengambil keputusan adalah dengan menggunakan intuisi atau insting. meski demikian kebanyakan orang akan setuju bahwa cara pengambilan keputusan yang paling baik adalah dengan menggunakan data yang relevan. Deep Learning memungkinkan pengambilan keputusan berdasarkan data dengan mengidentifikasi dan mengekstraksi pola dari dataset yang besar yang secara akurat terdistribusi dari serangkaian input yang kompleks menjadi suatu pengambilan keputusan yang baik[15][17].

2. Faster RCNN

Faster R-CNN” adalah model deteksi objek yang memperluas Fast R-CNN dengan menghubungkannya dengan Region Proposal Network (RPN). RPN berbagi fitur konvolusi dari seluruh gambar dengan jaringan deteksi, memungkinkan proposal wilayah hampir tanpa biaya. RPN adalah jaringan sepenuhnya konvolusi yang secara bersamaan memprediksi batas objek dan skor objek di setiap posisi. Dengan menggabungkan RPN dan Fast R-CNN menjadi satu jaringan, kita menciptakan sistem terpadu untuk deteksi objek. Dalam terminologi yang sedang populer, komponen RPN bertindak sebagai mekanisme “perhatian”, membimbing jaringan terpadu pada area yang perlu difokuskan[5][6][14].

3. CNN(Convolutional Neural Network)

Jaringan konvolusional, atau CNN, adalah jenis jaringan saraf yang dirancang khusus untuk menangani data dengan struktur grid yang sudah dikenal. Contoh data tersebut termasuk deret waktu, yang dapat dilihat sebagai grid 1D dengan sampel yang diambil pada interval waktu yang tetap, dan data gambar, yang dapat dianggap sebagai grid 2D dari piksel. CNN telah terbukti sangat efektif dalam berbagai aplikasi praktis. Istilah "jaringan saraf konvolusional" merujuk pada penggunaan operasi matematika yang disebut konvolusi dalam jaringan ini. Konvolusi merupakan bentuk operasi linear yang khusus. CNN adalah jaringan saraf yang menggunakan konvolusi sebagai pengganti perkalian matriks umum di setidaknya salah satu lapisannya[9][10].

4. Resnet(Residual Network)

Residual Networks (ResNets) adalah *deep convolutional network* di mana ide dasarnya adalah melewati blok-blok lapisan konvolusi dengan menggunakan koneksi pintas. Blok dasar yang disebut "bottleneck" mengikuti dua aturan desain sederhana: (i) untuk ukuran peta fitur keluaran yang sama, lapisan memiliki jumlah filter yang sama; dan (ii) jika ukuran peta fitur berkurang setengah, jumlah filter digandakan. Pengambilan sampel dilakukan langsung oleh lapisan konvolusi yang memiliki stride 2 dan normalisasi batch dilakukan tepat setelah setiap konvolusi dan sebelum aktivasi ReLU.

Ketika dimensi input dan output sama, digunakan shortcut identitas. Ketika dimensinya meningkat, shortcut proyeksi digunakan untuk menyamakan dimensi melalui konvolusi 1×1 . Dalam kedua kasus, ketika shortcut melewati peta fitur dari dua ukuran, mereka dilakukan dengan stride 2. Jaringan ini diakhiri dengan lapisan fully-connected (fc) sebanyak 1.000 dengan aktivasi softmax. Total jumlah lapisan berbobot adalah 50, dengan 23.534.592 parameter yang dapat dilatih[7].

5. RPN(Region Proposal Neural Network)

Jaringan Proposal Wilayah (Region Proposal Network atau RPN) menerima gambar (dengan ukuran berapa pun) sebagai input dan menghasilkan serangkaian proposal objek berbentuk persegi panjang, masing-masing dengan skor objectness. RPN dapat dilatih secara end-to-end menggunakan backpropagation dan stochastic gradient descent (SGD). Setiap mini-batch berasal dari satu gambar yang berisi banyak anchor contoh positif dan negatif. Memungkinkan untuk mengoptimalkan fungsi kerugian untuk semua anchor, tetapi ini akan cenderung mengutamakan sampel negatif karena jumlahnya lebih banyak. Sebagai gantinya, 256 anchor dapat dipilih dalam sebuah gambar untuk menghitung fungsi kerugian dari mini-batch, di mana anchor positif dan negatif yang dipilih memiliki rasio hingga 1:1. Jika terdapat kurang dari 128 sampel positif dalam sebuah gambar, kami mengisi mini-batch tersebut dengan sampel negative[5].

6. Dataset

Dalam terminologi statistik dataset adalah kumpulan dari suatu objek yang mempunyai atribut atau variabel tertentu, di mana untuk setiap objek merupakan individu dari data yang mempunyai sejumlah atribut atau variabel tersebut. Nama lain dari objek yang sering digunakan adalah record, point, vector, pattern, event, observation, dan case. Sementara itu, baris yang menyatakan objek-objek data dan kolom disebut atribut. Atribut juga dapat disebut dengan variabel, field, fitur atau dimensi[12].

7. **Augmentasi Data**

Salah satu cara untuk meningkatkan generalisasi dari model adalah dengan melatihnya menggunakan data yang besar, namun seringkali karena adanya keterbatasan kebutuhan pada data ini tidak bisa terpenuhi. Adapun cara untuk mengatasi kebutuhan ini dalam kondisi keterbatasan yang dihadapi adalah dengan membuat data palsu secara digital, proses ini disebut dengan proses augmentasi[9]. Augmentasi citra yang merupakan penerapan satu atau lebih deformasi pada kumpulan sampel pelatihan beranotasi yang menghasilkan data pelatihan baru, tambahan, dan berpotensi non-redundan[10]. Teknik telah menjadi teknik yang efektif dalam proses klasifikasi terutama proses *object recognition*.

Operasi-operasi seperti *translating* gambar pelatihan beberapa piksel ke berbagai arah tertentu sering kali dapat meningkatkan kemampuan generalisasi model secara signifikan, meskipun model tersebut telah dirancang untuk memiliki ketahanan sebagian terhadap translasi melalui penggunaan konvolusi dan pooling. Selain itu, operasi lain seperti *rotating* atau mengubah ukuran gambar juga terbukti sangat efektif. Teknik augmentasi data terbukti dapat secara drastis mengurangi generalisasi error pada dalam teknik pembelajaran mesin[7][12].