

PENDETEKSIAN DAN PENGENALAN JENIS MOBIL MENGUNAKAN ALGORITMA YOU ONLY LOOK ONCE DAN CONVOLUTIONAL NEURAL NETWORK

Calvin Geraldy¹⁾ Chairisni Lubis²⁾

¹⁾²⁾ Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara
Jl. Letjen S. Parman No.1, Jakarta
email : calvin.dquaker@gmail.com¹⁾, chairisnil@fti.untar.ac.id²⁾

ABSTRACT

Car Type Detection and Recognition system is an application that is developed using You Only Look Once (YOLO) and Convolutional Neural Network (CNN) algorithm. This application purpose is to detect and recognize the car image from the data input. In this application the input image will be divided into two parts, namely the training image and test image. For the training image, the first step, the training image will be divided into two process stages, namely detecting the image of a car and searching for the unique features of a car.

To detect the image of the car, the image will be processed to detect parts of the image of the car and not the car using the YOLO method. After obtaining a part of the car image, the image of the car will be saved as a detection model. The image that has been detected will be learning the car image by the CNN method. For the test images of the stages carried out as in the training image, after the image of the car is detected, an introduction will be made based on learning that has been done with the CNN method to obtain output in the form of a car that is successfully recognized and detected will be labeled by the application.

Key words

Car type, Convolutional neural network, Detection and Recognition, You only look once.

1. Pendahuluan

Pada dasarnya, sangat mudah bagi manusia untuk mendeteksi dan mengenali jenis mobil pada gambar atau video. Bagi komputer, sulitnya mengenali dan mengklasifikasi sebuah mobil sangat tergantung pada jenis data [1]. Kondisi pencahayaan, kompleksitas, dan kontras latar dari gambar atau video merupakan salah satu tantangan utama. Selain itu, kualitas gambar atau video secara keseluruhan juga sangat berpengaruh pada hasil pengenalan. Mobil memiliki bentuk dan warna yang beragam. Beberapa model bahkan memiliki kemiripan dengan mobil lainnya.

Ciri khas yang dimiliki setiap mobil sangat unik. Antara satu mobil dengan mobil lainnya pasti memiliki perbedaan meskipun tergolong terlihat serupa. Ciri yang menjadi pembeda antara lain: bentuk bodi, logo, dimensi, jumlah pintu, dan sebagainya. Untuk dapat membedakan berbagai jenis mobil digunakan fitur unik yang terdapat pada mobil tersebut.

2. Dasar Teori

2.1. You Only Look Once

You only look once (YOLO) adalah sebuah algoritma yang dikembangkan untuk mendeteksi sebuah objek secara real-time. Sistem pendeteksian yang dilakukan adalah dengan menggunakan repurpose classifier atau localizer untuk melakukan deteksi. Sebuah model diterapkan pada sebuah citra di beberapa lokasi dan skala. Daerah dengan citra yang diberi skor paling tinggi akan dianggap sebagai sebuah pendeteksian [2].

Sebelum dilakukan proses pelatihan perlu dilakukan proses anotasi terlebih dahulu untuk membentuk dataset. Untuk setiap data memiliki nama kelas, titik koordinat X objek, titik koordinat Y objek, panjang kotak pembatas, dan lebar kotak pembatas [3].

YOLO menggunakan pendekatan jaringan saraf tiruan (JST) untuk mendeteksi objek pada sebuah citra. Jaringan ini membagi citra menjadi beberapa wilayah dan memprediksi setiap kotak pembatas dan probabilitas untuk setiap wilayah. Kotak-kotak pembatas ini kemudian dibandingkan dengan setiap probabilitas yang diprediksi. YOLO memiliki beberapa kelebihan dibandingkan dengan sistem yang berorientasi pada classifier, terlihat dari seluruh citra pada saat dilakukan test dengan prediksi yang diinformasikan secara global pada citra.

YOLO menerapkan arsitektur yang mirip seperti Convolutional neural networks. YOLO hanya menggunakan lapisan konvolusi dan lapisan pooling. Untuk lapisan konvolusi terakhir disesuaikan dengan jumlah kelas dan jumlah kotak prediksi yang diinginkan.

Untuk menghitung ukuran keluaran pada lapisan konvolusi terakhir dapat menggunakan rumus sebagai berikut [4]:

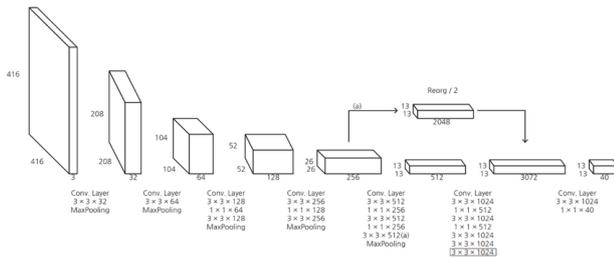
$$Y = S, S, B \times (5 + C) \tag{1}$$

2.2. Convolutional Neural Network

Convolutional neural network biasa disingkat CNN atau ConvNet adalah salah satu kelas deep feed-forward artificial neural network yang banyak digunakan pada analisis citra. CNN bisa digunakan untuk mendeteksi dan mengenali objek pada sebuah citra.

Secara garis besar CNN tidak jauh berbeda dengan neural network biasanya. CNN terdiri dari neuron yang memiliki weight, bias dan activation function. CNN terdiri atas satu lapisan masukan (input layer), suatu lapisan keluaran (output layer) dan sejumlah lapisan tersembunyi (hidden layer).

Pada arsitektur YOLOv2, hanya terdiri dari lapisan konvolusi dan lapisan pooling. Berikut adalah ilustrasi dari arsitektur YOLOv2



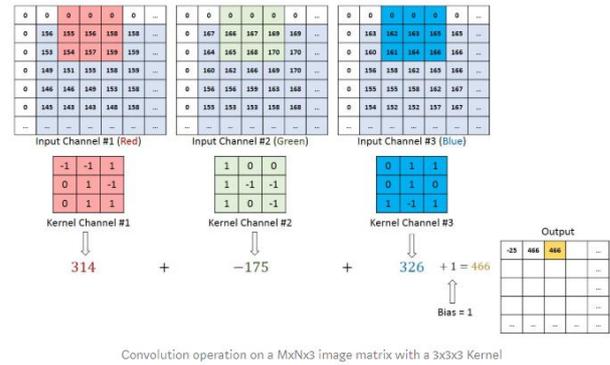
Gambar 1. Arsitektur YOLOv2

2.3. Feature Extraction Layer

Lapisan ekstraksi fitur (Feature Extraction Layer) citra terletak pada awal arsitektur yang tersusun atas beberapa lapisan dan setiap lapisan tersusun atas neuron yang terkoneksi pada daerah lokal (local region) lapisan sebelumnya. Lapisan pertama adalah lapisan konvolusi (convolution layer) dan lapisan kedua adalah lapisan pooling (Pooling Layer).

2.4. Convolution Layer

Tahap awal dalam proses CNN yang memproses input akan menghasilkan banyak peta fitur (Feature map). Input awal berupa citra akan dilakukan konvolusi dengan filter. Filter yang digunakan tergantung citra input yang dimasukan sebagai input seperti 3x3, 5x5, 7x7, dan sebagainya. Pada proses penerapan filter, digunakan parameter stride untuk menentukan pengulangan pada pergeseran filter dan proses padding untuk memanipulasi dimensi output pada lapisan konvolusi.



Gambar 2. Operasi konvolusi pada citra matriks MxNx3 dengan kernel 3x3x3

2.5 Stride

Stride adalah parameter yang menentukan berapa jumlah pergeseran filter. Jika nilai stride adalah 1, maka filter konvolusi akan bergeser sebanyak 1 piksel secara horizontal lalu vertikal. Semakin kecil stride maka akan semakin detail informasi yang diperoleh dari sebuah input, namun membutuhkan komputasi yang lebih jika dibandingkan dengan stride yang besar. Perlu diperhatikan bahwa dengan menggunakan stride yang kecil, tidak selalu didapatkan performa yang bagus.

2.6. Padding

Padding atau Zero Padding adalah parameter yang menentukan jumlah piksel berisi nilai 0 yang akan ditambahkan di setiap sisi dari input. Hal ini digunakan dengan tujuan untuk memanipulasi dimensi output dari lapisan konvolusi (Feature map).

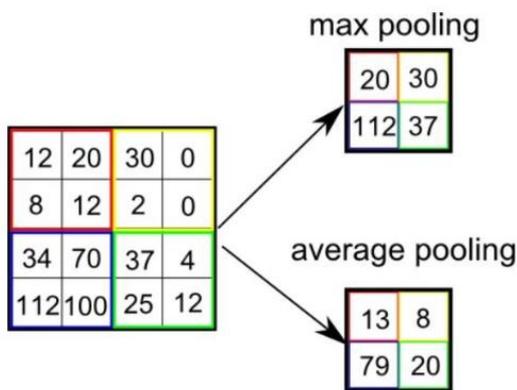
Untuk menghitung dimensi dari feature map digunakan rumus seperti dibawah ini:

$$Output = \frac{W - N + 2P}{S} + 1 \tag{2}$$

2.7. Pooling Layer

Pada lapisan pooling (Pooling Layer) peta-peta fitur yang telah di proses melalui ReLU akan dikecilkan dengan memilih ukuran window. Dengan adanya lapisan pooling, data akan menjadi kecil, mudah dikelola, dan mudah mengontrol overfitting.

Cara mengecilkan peta fitur yaitu dengan menggunakan fungsi MaxPooling. MaxPooling yaitu fungsi yang memilih nilai maksimum dari daerah window tersebut lalu direpresentasikan sebagai piksel baru. MaxPooling bergerak dari pojok kiri atas ke kanan dengan ukuran pooling yang telah ditentukan dan pergeseran pooling yang telah ditentukan.



Types of Pooling

Gambar 3 Contoh Max Pooling dan Average Pooling

2.8. Softmax

Softmax berfungsi untuk menghitung probabilitas dari setiap kelas target atas semua kelas target yang memungkinkan dan membantu menentukan kelas target untuk *input* yang diberikan. Rumus untuk menghitung softmax sebagai berikut [5].

$$f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}} \quad (3)$$

3. Hasil Percobaan

Pengujian yang dilakukan yaitu pengujian pendeteksian dan pengenalan jenis mobil. Pengujian pendeteksian mobil menggunakan dataset citra mobil dari *Stanford Car Dataset* dari *Stanford University* yang berisi 16.185 citra yang terdiri atas 196 kelas mobil yang berbeda yang terbagi atas 8.144 citra latih dan 8.041 citra uji.

Untuk dataset latih dari *Stanford Car Dataset* digunakan sebanyak 100 kelas dengan 50 citra setiap kelasnya. Sedangkan untuk dataset uji digunakan sebanyak 100 kelas dengan 10 citra setiap kelasnya.



Gambar 4 Contoh Citra Mobil dari Stanford Car Dataset from Stanford University

Pengujian pendeteksian dan pengenalan berguna untuk mengukur tingkat keakurasian hasil dari pelatihan dataset yang telah dilatih. Pengujian ini menggunakan dataset *Stanford Car Image Dataset* yang

diambil dari 100 kelas dengan 10 citra pertama setiap kelas. Hasil dari pengenalan mobil adalah sebagai berikut:

Tabel 1 Tingkat Akurasi dari Hasil Pengujian

Learning rate	Epoch	Akurasi	
		Pendeteksian	Pengenalan
0.000005	100	88,1 %	84,5 %
0.00001	100	87,1 %	80,3 %
0.00001	80	85,8 %	70,5 %
0.00001	50	83,8 %	58,8 %

4. Kesimpulan

Berdasarkan pengujian program “Pendeteksian dan Pengenalan Jenis Mobil Menggunakan Algoritma You Only Look Once dan Convolutional Neural Network”, maka didapatkan kesimpulan sebagai berikut:

- Metode You only look once dan Convolutional neural network dapat digunakan dengan baik untuk pendeteksian dan pengenalan mobil. Pada uji deteksi dan pengenalan mobil didapat keberhasilan sebesar 88,1%. Pengujian pengenalan mobil yang paling optimal menggunakan learning rate 0.000005 dan epoch 100.
- Pada pengujian, aplikasi dapat mendeteksi dan mengenali lebih dari 1 jenis mobil pada 1 citra uji.
- Intensitas cahaya, jumlah mobil dalam citra, dan modifikasi eksterior mobil pada citra dapat mempengaruhi hasil pendeteksian dan pengenalan.

REFERENSI

- [1] Viktoria Plemakova., 2018, “Vehicle Detection Based on Convolutional Neural Networks”, University of Tartu, Estonia.
- [2] Sisco Jupiyandi, dkk., “Pengembangan Deteksi Citra Mobil Untuk Mengetahui Jumlah Tempat Parkir Menggunakan Cuda Dan Modified YOLO”, Universitas Brawijaya, Malang.
- [3] Mao, Nils Tijtgat, “How to train YOLOv2 to detect custom objects”, Github.
- [4] Matthijs Hollemans, “Real-Time Object Detectiob With YOLO”, Machinethink.net.
- [5] Dataaspirant, Difference Between Softmax Function And Sigmoid Function, Dataaspirant.

Calvin Geraldy, seorang mahasiswa pada proram studi Fakultas Teknologi Informasi di Universitas Tarumanagara

Dra. Chairisni Lubis, M.Kom., memperoleh gelar Dra. Dari Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia dan gelar M.Kom dari Universitas Indonesia. Saat ini sebagai Dosen program studi Teknik Informatika, Universitas Tarumanagara.