
Pemodelan Identifikasi Objek Kendaraan Bermotor Menggunakan *Faster Region based Convolutional Neural Network (R-CNN)* Berbasis *Python*

Rosalia Arum Kumalasanti¹, Erma Susanti²

¹Universitas Sanata Dharma

²Universitas AKPRIND Indonesia

Email: rosalia.santi@usd.ac.id¹, erma@akprind.ac.id²

ABSTRACT

The vehicles are currently experiencing a surge in number and variation. This is evident from the kinds of vehicles that are passing through the highway area. The rise in the number of motorized vehicles will surely give a squeeze to the traffic density. The increase in the number of motor vehicles is one of the biggest factors in the impact of the congestion. The congestion can also cause damage to the highway. It's supposed to be the focus of the local government in dealing with the problem. Each road point has its own potential, so it is necessary to have a calculation in identifying the number of vehicles and the type of vehicles that are slipped on the road. Motor vehicle identification can be solved using the Faster Region based Convolutional Neural Network approach. Faster R-CNN is a deep learning architecture used to detect inside computers. Research will run at several highway points to take samples of video at a certain time, for identified the type of vehicle. Vehicle labelling will facilitate the calculation of the number of vehicles crossing the road in a given unit of time. The vehicle identification needs are used to see the density of the highway so that it can help the local government in making the right decision or solution to reduce the traffic density. The results of research such as quantitative data can be easily used to give the right picture and decision.

Keywords: *Faster R-CNN, identification, traffic, vehicles.*

INTISARI

Kendaraan bermotor saat ini mengalami lonjakan jumlah dan variasi. Hal ini terlihat dari jenis-jenis kendaraan yang berlalu-lalang di area jalan raya. Lonjakan jumlah kendaraan bermotor tentunya akan memberikan dampak pada padatnya lalu lintas. Bertambahnya jumlah kendaraan bermotor merupakan salah satu faktor terbesar dari dampak kemacetan yang terjadi. Kemacetan yang terjadi juga dapat menyebabkan kerusakan jalan raya. Hal ini sudah seharusnya menjadi fokus pemerintah daerah dalam menanggulangi permasalahan. Setiap titik jalan memiliki potensi masing-masing, maka perlu adanya perhitungan dalam mengidentifikasi jumlah kendaraan dan jenis kendaraan yang lewat di jalan tersebut. Identifikasi kendaraan bermotor dapat tersolusikan dengan menggunakan pendekatan *Faster Region based Convolutional Neural Network*. *Faster R-CNN* merupakan arsitektur *deep learning* yang digunakan untuk mendeteksi di dalam komputer. Penelitian akan berjalan pada beberapa titik jalan raya untuk diambil sampel video pada waktu tertentu, untuk diidentifikasi jenis kendaraannya. Pelabelan kendaraan akan memberikan kemudahan dalam menghitung jumlah kendaraan yang melintasi jalan dengan satuan waktu tertentu. Kebutuhan identifikasi kendaraan ini digunakan untuk melihat kepadatan jalan raya sehingga dapat membantu pemerintah daerah dalam memberikan keputusan atau solusi yang tepat untuk mengurangi kepadatan jalan. Hasil penelitian berupa data kuantitatif akan dapat dengan mudah digunakan untuk memberikan gambaran dan keputusan yang tepat.

Kata kunci: *Faster R-CNN, identifikasi, kendaraan, lalu lintas.*

PENDAHULUAN

Transportasi lalu lintas, aktivitas industri, pariwisata dan kegiatan-kegiatan masyarakat saat ini menjadi salah satu aktivitas mobilitas yang tidak dapat dihindari. Kebutuhan mobilitas yang semakin tinggi diimbangi pula dengan bertambahnya jumlah kendaraan bermotor sehingga kendaraan melebihi kapasitas jalan (Mela, et al., 2021). Hal ini berdampak pada lalu lintas jalan yang semakin padat dan juga mengakibatkan kerusakan aspal di titik-titik jalan tertentu. Kerusakan aspal sering terjadi apabila area tersebut sangat sering dilalui kendaraan bermotor dengan alat beratnya. Pemerintah daerah tak jarang melakukan pengalihan jalur dalam rangka memberikan perbaikan jalan yang rusak untuk dibenahi kembali. Namun keadaan ini akan terus terulang kembali setiap tahunnya apabila tidak ada perhitungan secara aktual terkait jumlah kendaraan bermotor yang menggunakan

jalan raya dalam satuan waktu. Menurut BPS (Badan Pusat Statistik), perkembangan jumlah kendaraan bermotor berdasarkan jenisnya memiliki angka yang fantastis. Jumlah kendaraan bermotor jenis mobil sebanyak 17.168.862, bus sebanyak 243.450, truk sebanyak 5.544.173, sepeda motor sebanyak 125.305.332, maka jumlah keseluruhan kendaraan bermotor pada tahun 2022 mencapai 148.261.817 (BPS, 2024).

Saat ini masyarakat tengah menghadapi tantangan pada meningkatnya kepadatan lalu lintas di berbagai kota di dunia. Populasi saat ini sangat bergantung pada kendaraan untuk tujuan kelancaran komutasi. Hal ini disebabkan oleh urbanisasi terus menerus di daerah pedesaan. Antisipasi yang bisa dilakukan adalah *smart traffic management*. Pemerintah daerah saat ini sedang berusaha keras dalam memberikan pelayanan transportasi yang optimal bagi masyarakat khususnya di Yogyakarta. Yogyakarta merupakan kota pelajar sekaligus menjadi primadona tempat wisata yang hingga saat ini terus bertambah. Kelayakan jalan tentunya menjadi kebutuhan pokok dalam berkendara dan hal ini berkaitan langsung dengan keselamatan dalam berkendara. Di era digital saat ini GPS sudah memberikan solusi dalam pemilihan jalur yang cepat dengan menghindari titik-titik kemacetan, namun ternyata hal ini dirasa belum cukup menjadi solusi. Tentunya dengan jumlah kendaraan yang semakin bertambah, setiap pengemudi kendaraan pasti akan mencari jalan untuk berlalu lintas yang menurut mereka aman dari kemacetan hingga sebagian besar yang menggunakan GPS berkumpul pada titik yang sama dan menurut para pengemudi, jalan tersebut adalah jalan yang tepat untuk dilalui. Masalah tersebut menjadi latar belakang dari ide penelitian ini.

Yogyakarta memiliki fasilitas berupa CCTV yang dipasang hampir di setiap sisi jalan. CCTV milik Dinas Perhubungan ini akan digunakan sebagai sumber data berupa video dan akan dijadikan data sampel. Kebutuhan sampel video ini akan diambil pada waktu-waktu tertentu yang dianggap berpotensi terjadi kemacetan, misal saat jam berangkat kantor di pagi hari dan pulang kantor di sore hari. Tren yang akan diambil adalah saat jalan raya mengalami lonjakan kendaraan bermotor sehingga didapat perhitungan yang maksimal. Diharapkan dari hasil penelitian ini, dapat membantu pemerintah daerah dalam menemukan solusi dari kemacetan jalan raya yang terjadi hampir setiap hari di jam-jam tertentu. Diharapkan pula pemerintah daerah dapat mengambil tindakan dalam memberikan solusi bagi para pengendara bermotor. Kenyamanan dan keamanan berkendara menjadi ukuran keselamatan dalam berkendara.

Di era digital saat ini tentunya teknologi dapat dimanfaatkan untuk mengurangi dampak volume kendaraan yang semakin tinggi. Salah satu bidang yang saat ini dapat dimanfaatkan adalah *machine learning*. Banyak sekali metode yang dapat digunakan dalam *machine learning*, namun saat CNN banyak digunakan untuk pemanfaatan deteksi objek. Jaringan CNN dapat memproses *dataset* sebanyak 3 dimensi sekaligus. Ada beberapa teknik CNN yang populer, terutama dilihat dari kecepatannya yaitu *Faster R-CNN*. *R-CNN* memiliki keunggulan berupa kecepatan dalam mempelajari karakteristik suatu objek. Objek berupa video CCTV akan digunakan sebagai data input untuk dipelajari *Faster R-CNN* untuk kemudian diberikan label berdasarkan jenis kendaraan. Setelah sistem berhasil memberikan label pada setiap jenis kendaraan maka akan dengan mudah untuk menghitung jumlah kendaraan berdasarkan jenis pada satuan waktu tertentu.

METODE

Deteksi Objek

Data diambil dari CCTV milik Dinas Perhubungan yang diletakkan di sisi jalan untuk memantau dan merekam lalu lintas. CCTV atau kamera yang digunakan saat ini terbilang merupakan sensor serbaguna dalam mendeteksi objek kendaraan dari berbagai sudut dan lebih dari satu lajur. Deteksi objek merupakan teknik *computer vision* dalam menemukan lokasi objek pada video yang terekam pada CCTV (Jasman & Hendri, 2022).

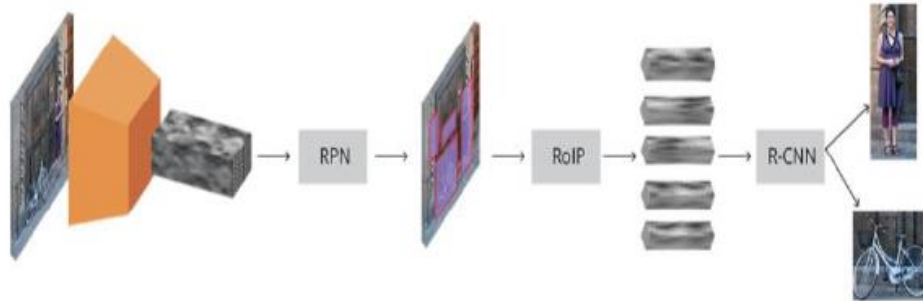
Deteksi objek dalam pembelajaran mesin merupakan proses yang melibatkan identifikasi dan klasifikasi objek dalam citra atau video. Beberapa langkah di dalam deteksi objek meliputi pengumpulan data, *preprocessing*, pemilihan model, pelatihan model, evaluasi model, penyempurnaan model dan *deploy* model.

Faster R-CNN

Permasalahan dalam deteksi objek pada penelitian ini adalah berupa variasi kendaraan bermotor. Deteksi objek melibatkan deteksi kelas yang berbeda-beda pada setiap video sehingga membutuhkan *machine learning* dalam identifikasi (Moch, et al., 2021). Secara konvensional, strategi yang digunakan adalah untuk mendistribusikan klasifikator pada dua kelas (objek dan non objek), namun penelitian ini membutuhkan lebih dari itu. CNN merupakan arsitektur yang cukup mumpuni dalam melakukan klasifikasi di dalam *deep learning*. CNN merupakan jaringan yang memiliki banyak *layer* dan di setiap *layer* yang terdiri dari beberapa dimensi, namun penelitian ini

memiliki kebutuhan lebih dari sekedar pelabelan seperti deteksi objek pada umumnya. Arsitektur jaringan yang lebih kompleks dibutuhkan untuk memberikan performa yang optimal.

Faster R-CNN hadir dengan keunggulannya yaitu kecepatan dalam mempelajari objek. Arsitektur *Faster R-CNN* terdiri dari tiga bagian utama yaitu *Convolutional Neural Network*, *Region Proposal Network*, dan *ROI Pooling Layer* (Jasman & Hendri, 2022). Penelitian ini terdiri dari dua fase penting yaitu fase *training* dan fase *testing*. Fase *training* merupakan tahap untuk mempersiapkan *dataset* untuk dipelajari oleh *Faster R-CNN* dalam mendapatkan karakteristik dari setiap objek, sedangkan fase *testing* adalah tahap dimana *machine learning* akan membandingkan hasil pelatihan dengan *dataset testing* (Sunario & Wulan, 2020). Gambar 1. merupakan arsitektur *Faster R-CNN*.



Gambar 1. Arsitektur *Faster R-CNN* (Mohan, et al., 2021)

HASIL DAN PEMBAHASAN

Dataset

Dataset berupa video kendaraan bermotor diambil dari CCTV secara langsung yang berlokasi di titik kepadatan jalan raya. Sampel yang diambil untuk penelitian ini adalah jalan raya area Jogja yang memiliki intensitas kepadatan kendaraan yang signifikan. Beberapa titik diambil untuk digunakan sebagai sampel dan pembandingan. *Dataset* berupa video diambil dan akan dikonversi ke dalam *frame by frame* untuk diolah lebih lanjut. Data citra yang didapat dari hasil konversi kemudian akan diolah di dalam pemodelan *Faster R-CNN*. Penelitian ini akan membagi objek menjadi 4 kelas yaitu, mobil, bus, truk dan sepeda motor. Gambar 2. merupakan sampel citra pada saat pelabelan.

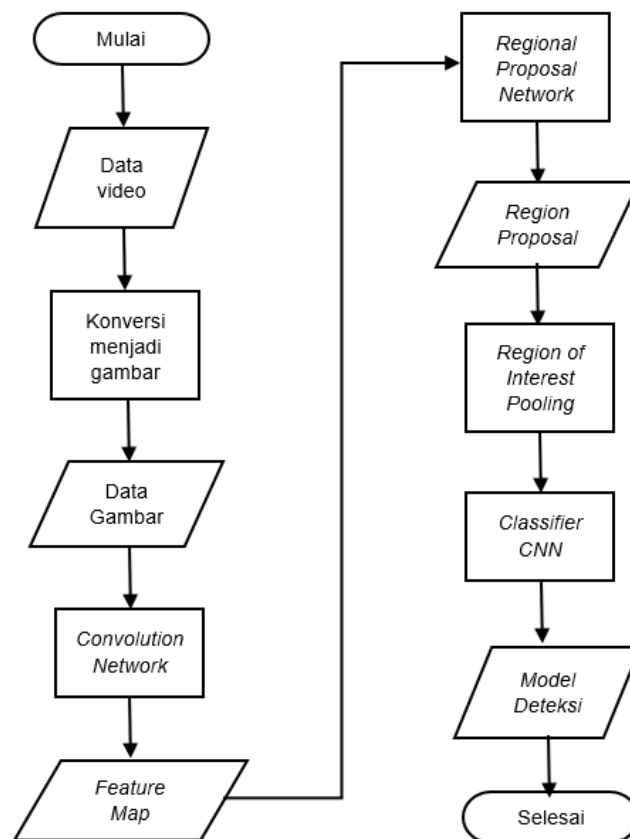


Gambar 2. *Dataset* sebelum pelabelan

Training

Penelitian ini memiliki dua fase penting di dalam pemodelan, yaitu fase *training* dan *testing*. Fase *training* pada penelitian ini melibatkan *dataset* yang sudah diambil dari rekaman CCTV untuk diubah menjadi citra. Parameter-parameter berupa *epoch*, *optimizer*, *learning rate* dan *batch size* menjadi hal penting dalam pencapaian akurasi. Adapun alur *training* di dalam penelitian ini, meliputi pengumpulan *dataset* berupa video hasil rekaman CCTV untuk dikonversi ke dalam *frame by frame*

sehingga menjadi *dataset* citra atau gambar. Data berupa citra akan diolah di dalam model yaitu menggunakan *Convolutional Network* dan kemudian akan didapat *feature map*. *Feature map* di sini merupakan hasil dari operasi konvolusi antara input dan *kernel* (filter). Proses ini menghasilkan peta fitur yang menyoroti pada fitur-fitur tertentu dari input asli seperti *edge*, tekstur, atau pola lainnya. Setiap *kernel* mempelajari fitur secara spesifik. Hasil dari *feature map* kemudian akan diproses oleh *Regional Proposal Network* (RPN). RPN merupakan komponen penting dalam model jaringan syaraf tiruan. RPN menghasilkan regional proposal atau proposal wilayah yang mungkin berisi objek dalam gambar. Proposal wilayah akan dilanjutkan oleh *Region of Interest* (RoI) untuk mengekstrak fitur dengan ukuran yang tetap. Tahap berikutnya hasil ekstraksi akan diklasifikasikan oleh CNN. CNN dirancang secara otomatis dan adaptif mempelajari fitur hierarki dari data input melalui penggunaan lapisan konvolusi, *pooling* dan *fully connected*. Fase *training* ini dilakukan hingga mendapatkan model yang tepat dalam pengklasifikasian. Alur fase *training* dapat dilihat pada Gambar 3.

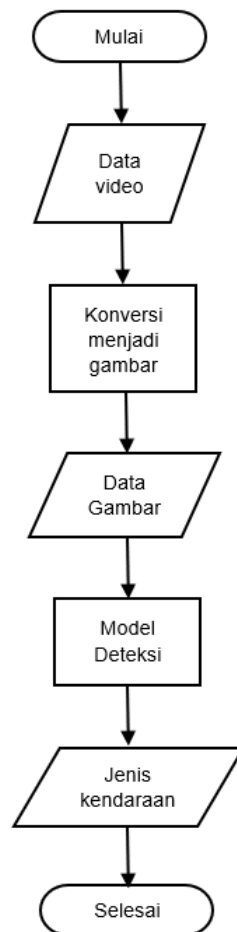


Gambar 3. Alur fase *training*

Testing

Fase *testing* merupakan fase utama kedua setelah fase *training*. Fase ini membutuhkan *dataset* yang diperlakukan sama seperti saat berada di fase *training*. *Dataset* berupa video dikonversi ke dalam citra untuk diberikan proses *Convolution Network* hingga mendapatkan *feature map*. Hasil dari *testing* kemudian akan dibandingkan dengan hasil *training*. Citra akan dideteksi dan mendapatkan akurasi. Apabila akurasi masih belum maksimal maka perlu melakukan pemodelan lagi dalam rangka mendapatkan model yang tepat dan memberikan akurasi yang optimal.

Fase *testing* dalam CNN adalah langkah di mana model yang sudah dilatih menggunakan data yang belum pernah dilihat oleh sistem. Tujuannya adalah untuk menilai kinerja model dalam hal akurasi, generalisasi, dan kemampuan untuk menangani data yang belum pernah ditemui. Alur fase *testing* dapat dilihat pada Gambar 4.



Gambar 4. Alur fase *testing*

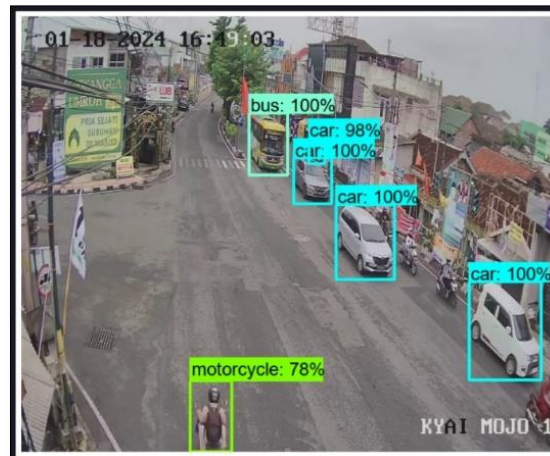
Hasil dan Pembahasan

Implementasi R-CNN pada *dataset* video rekaman CCTV membutuhkan banyak preparasi karena pada saat menentukan video yang akan digunakan, tentu perlu menentukan sisi pengambilan gambar, waktu dan kepadatan. *Dataset* yang digunakan adalah pagi dan sore pada saat volume kendaraan meningkat. Beberapa hal yang perlu diperhatikan adalah apabila terdapat video yang memiliki beberapa *frame* dengan objek yang kosong maka data tersebut tidak digunakan. Ada pula kendaraan yang tidak perlu diolah, misal kendaraan yang di luar dari label yang sudah ditentukan. Keberhasilan pelabelan akan memberikan perhitungan aktual terkait jumlah kendaraan pada satuan waktu tertentu. Sampel video yang digunakan memang dengan sengaja diambil pada saat jalur lalu lintas padat karena itu adalah titik optimal dari perhitungan, sehingga ini sangat membantu untuk mengambil keputusan pada saat *worst case*.

Pada penelitian ini R-CNN memiliki dua fase penting yaitu fase *training* dan fase *testing*, maka akurasi bisa didapatkan dari pemodelan tersebut. Parameter-parameter berupa *epoch*, *optimizer*, *learning rate* dan *batch size* menjadi hal penting dalam menentukan model yang tepat. Memerlukan beberapa kali eksperimen untuk mendapatkan model tersebut. Hasil pelabelan terdiri dari 4 target yaitu mobil, motor, bus dan truk. Apabila di dalam video terdeteksi kendaraan lain makan objek tersebut akan diabaikan. Anotasi dilakukan menggunakan CVAT (*Computer Vision Annotation Tool*) secara manual. Akurasi berupa persentase akan muncul pada kotak setiap label kendaraan. Apabila akurasi dirasa belum optimal maka pemodelan dapat disesuaikan kembali, terutama pada parameter yang mempengaruhi.

Setelah akurasi pada setiap label didapatkan maka langkah berikutnya untuk mencapai tujuan penelitian ini adalah menghitung jumlah kendaraan bermotor pada masing-masing jenis. Perhitungan ini dapat menggunakan perhitungan secara manual dalam satu satuan waktu. Kepadatan lalu lintas akan dihitung dari volume kendaraan yang melintasi titik jalan tersebut. Beberapa eksperimen dapat divariasikan berdasarkan waktu, sisi pengambilan video (depan,

samping kiri, samping kanan, atau belakang) dan pencahayaan. Perhitungan jumlah kendaraan ini membutuhkan pemilihan CCTV yang strategis dalam pengambilan gambar terbaik. Hasil pelabelan dapat dilihat pada Gambar 5.



Gambar 5. Dataset setelah pelabelan

KESIMPULAN

Rancangan arsitektur *Faster R-CNN* digunakan untuk deteksi objek berupa citra, yaitu hasil konversi video rekaman CCTV. Pemodelan yang tepat akan memberikan akurasi yang optimal. Hasil pelabelan memberikan perhitungan untuk setiap variasi motor pada setiap satuan waktu, sehingga bisa dihitung jumlah motor, mobil, truk dan bus di setiap menitnya. Jumlah kendaraan ini yang digunakan untuk memberikan solusi terhadap dampak kemacetan. Volume kendaraan yang didapat bisa digunakan pemerintah daerah untuk memberikan kebijakan terkait perubahan jalur, tambahan jalur ataupun kebijakan lainnya yang tentunya dengan tujuan meminimalkan kemacetan jalan dan kecelakaan.

DAFTAR PUSTAKA

- BPS, 2024. *Badan Pusat Statistik*. [Online] Available at: <https://www.bps.go.id/id/statistics-table/3/VjJ3NGRGa3dkRk5MTIU1bVNFOTVVbmQyVURSTVFUMDkIMw==/jumlah-kendaraan-bermotor-menurut-provinsi-dan-jenis-kendaraan-unit-.html?year=2022> [Accessed 1 Juli 2024].
- Jasman, P. & Hendri, H., 2022. Deteksi Objek Kereta Api menggunakan Metode *Faster R-CNN* dengan Arsitektur *VGG 16*. *Multimedia Artificial Intelligent Networking Database*, 7(1), pp. 21-36.
- Mela, T. A., Fitri, U. & Dahnia, S., 2021. Sistem Deteksi dan Klasifikasi Jenis Kendaraan berbasis Citra dengan Menggunakan Metode *Faster-RCNN* pada *Raspberry Pi 4B*. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 5(2), pp. 814-819.
- Moch, D. L. Y., Wawan, S. & Yaya, W., 2021. Deteksi Sepeda Motor di Jalan Raya Menggunakan *Faster R-CNN* Berbasis *VGG16*. *Jurnal Aplikasi dan Teori Ilmu Komputer*, 4(2), pp. 10-13.
- Mohan, K. K., Sowmya, A., Jerusha, D. & Susmitha, D., 2021. Comparative Study of Vehicle Detection using *SSD* and *Faster R-CNN*. *International Journal of Computer Science and Mobile Computing*, 10(7), pp. 28-33.
- Sunario, M. & Wulan, S. L., 2020. Deteksi Spoofing Wajah Menggunakan *Faster R-CNN* dengan Arsitektur *Resnet50* pada Video. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 9(3), pp. 261-267.
- Viky, P. S., Ulinnuha, L. & Ibrahim, 2023. Simulasi Detection Counter pada Objek Kendaraan Motor dan Mobil Menggunakan Metode *Convolutional Neural Network* Berbasis *Python*. *Jurnal Ilmiah Wahana Pendidikan*, 9(16), pp. 760-766.