

Penerapan Image Detection Menggunakan YOLO Ultralytics Framework

Rully Meidyta
Program Studi Magister Matematika
Universitas Indonesia

Abstract

Deteksi objek adalah komponen penting dalam pengolahan gambar dan penglihatan komputer, vital untuk aplikasi seperti pengawasan keamanan, kendaraan otonom, dan analisis medis. Dalam laporan ini, menjelaskan penggunaan tiga model terkini: YOLOv8, YOLO-SEG, dan RT-DETR. Ketiganya digunakan dengan dataset COCO 128 untuk melatih dan menguji performanya. Hasil evaluasi menunjukkan kinerja yang sangat baik: YOLOv8 mencapai mAP50 0.89114, RT-DETR mencapai 0.96476, dan model segmentasi mencapai 0.89114, menunjukkan kemajuan dalam deteksi objek. Penelitian ini berkontribusi pada pemahaman tentang pentingnya deteksi objek dalam pengolahan gambar dan pengenalan model-model efisien dan akurat untuk berbagai aplikasi.

I. PENDAHULUAN

Deteksi objek adalah salah satu tugas utama dalam pengolahan gambar yang bertujuan untuk mengidentifikasi dan menggambar kotak pembatas (bounding box) di sekitar objek-objek yang ada dalam gambar. Tugas ini sangat penting dalam berbagai aplikasi, termasuk pengawasan keamanan, kendaraan otonom, analisis medis, dan banyak lagi. Dengan kemajuan teknologi, model-model deteksi objek yang kuat dan efisien semakin penting untuk mencapai hasil yang akurat dan real-time.

Deteksi objek yang efisien dan akurat adalah tantangan utama dalam pengolahan gambar. Dalam beberapa tahun terakhir, kemajuan pesat dalam teknologi deep learning telah menghasilkan model-model yang lebih unggul. Dalam implementasi ini, kita akan membahas penggunaan Ultralytics YOLOv8, YOLO-SEG, dan RT-DETR untuk tugas deteksi objek. Ketiga model ini adalah representasi terbaru dalam pengolahan gambar yang memanfaatkan transformasi dan arsitektur terkini dalam dunia kecerdasan buatan.

Dengan menggunakan dataset COCO 128, sebuah subset dari dataset COCO yang berisi 80 kelas objek. Dataset ini memungkinkan kita untuk melatih dan menguji model-model ini dalam konteks tugas deteksi objek dengan dataset yang lebih ringan, namun masih mencerminkan kompleksitas tugas deteksi objek di dunia nyata.

Dengan memahami potensi model-model ini dan penggunaan dataset COCO 128, kita akan menggali lebih dalam tentang cara mengimplementasikan tiga model tersebut untuk tugas deteksi objek yang efisien dan akurat. Dan memberikan gambaran umum tentang pentingnya deteksi objek dalam pengolahan gambar dan mengenalkan model-model yang akan digunakan dalam implementasi.

II. LATAR BELAKANG

Dalam dunia pengolahan gambar dan penglihatan komputer, deteksi objek adalah salah satu tugas paling penting. Kemampuan untuk mengidentifikasi dan melokalisasi objek dalam gambar atau video memiliki implikasi luas dalam berbagai aplikasi, seperti pengawasan keamanan, kendaraan otonom, pengenalan wajah, pemrosesan gambar medis, dan masih banyak lagi. Untuk mengatasi tugas ini, model-model deteksi objek terus berkembang, dan dalam laporan ini, kami akan menjelaskan tiga model terkini yaitu, YOLOv8, YOLO-SEG, dan RT-DETR.

1) YOLOv8

YOLOv8 adalah kelanjutan dari seri model YOLO yang telah lama terkenal dalam tugas deteksi objek. Salah satu fitur utama yang membedakan YOLOv8 adalah kemampuannya untuk menggabungkan kecepatan tinggi dengan akurasi yang tinggi. Model ini dirancang untuk memproses gambar dengan efisien, menjawab kebutuhan aplikasi yang memerlukan deteksi objek dalam waktu nyata atau hampir waktu nyata. Dengan bobot yang lebih ringan, YOLOv8 mampu mengatasi masalah deteksi objek dalam berbagai situasi, termasuk dalam lingkungan visual yang beragam.

2) YOLO-SEG

YOLO-SEG adalah model yang menggabungkan deteksi objek dengan konsep segmentasi instans. Selain mengidentifikasi objek, YOLO-SEG mampu menggambarkan objek-objek dalam gambar secara individu. Kemampuan ini memungkinkan kita untuk membedakan dan melokalisasi objek-objek secara detail. YOLO-SEG memiliki aplikasi penting dalam berbagai bidang, termasuk segmentasi citra medis, di mana identifikasi struktur anatomi dan kelainan memerlukan tingkat kejelasan yang tinggi.

3) RT-DETR

RT-DETR adalah representasi terbaru dalam dunia "real-time detection transformer." RT-DETR mengambil keunggulan arsitektur transformer, yang telah terbukti berhasil dalam pemrosesan teks, dan mengaplikasikannya dalam pemrosesan visual. Model ini memiliki kapabilitas untuk mendeteksi objek dalam waktu nyata, yang sangat penting untuk

aplikasi yang memerlukan respons cepat terhadap objek dalam gambar. Dengan RT-DETR, kita dapat menggabungkan kecerdasan buatan dan deteksi objek dalam solusi yang responsif dan efisien.

Penggunaan ketiga model ini sangat relevan dalam perkembangan teknologi saat ini, di mana aplikasi yang memerlukan pemahaman tentang objek dalam gambar semakin beragam dan menuntut respon yang cepat. Dan dengan harapan bahwa hasil penelitian ini akan memberikan kontribusi yang berharga dalam pengembangan aplikasi berbasis pengolahan gambar.

III. DATA DAN METODE

A. Data

Dalam penelitian ini, kami menggunakan dataset COCO 128, yang merupakan subset dari Microsoft Common Objects in Context (COCO) dataset. Dataset ini terdiri dari gambar-gambar berbagai objek yang telah diorganisir dalam 128 kelas berbeda. Dataset COCO 128 digunakan untuk train dan validation

B. Metode

1) Pelatihan Model

Pada tahap pelatihan model, langkah-langkah berikut diikuti:

Inisialisasi Model: Model-model YOLOv8, YOLO-SEG, dan RT-DETR diinisialisasi dengan bobot awal yang sesuai.
Pelatihan: Model-model dilatih menggunakan subset pelatihan dataset COCO 128. Pengaturan hyperparameter seperti jumlah epochs.. Monitoring Kinerja: Selama pelatihan, kinerja model dimonitor dengan metrik seperti mAP, presisi, dan recall.

2) Validasi Model

Setelah pelatihan, model-model divalidasi dengan menggunakan subset validasi dataset COCO 128 yang terpisah. Validasi melibatkan perhitungan metrik kinerja seperti mAP, presisi, dan recall untuk mengevaluasi akurasi deteksi objek.

3) Prediksi Objek

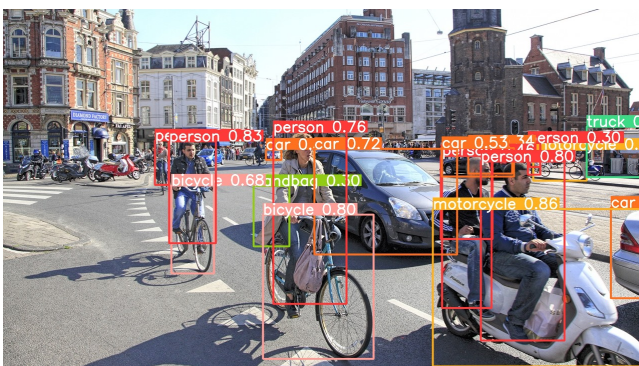
Model-model yang telah dilatih digunakan untuk melakukan prediksi deteksi objek pada gambar atau video yang diberikan. Hasil prediksi dianalisis untuk memahami sejauh mana model-model tersebut efektif dalam tugas deteksi objek.

IV. HASIL

A. Detection

Hasil evaluasi model deteksi objek menggunakan YOLO dengan metrik mAP50 sebesar 0.89114 menunjukkan bahwa model ini memiliki kemampuan yang sangat baik dalam mengenali objek-objek dalam gambar dengan tingkat kepercayaan yang minimal sebesar 50%. Selain itu, presisi model mencapai 0.92324, menunjukkan bahwa sebagian besar objek yang dideteksi oleh model adalah benar, sedangkan recall mencapai 0.83187, yang berarti bahwa model dapat mengidentifikasi sebagian besar objek yang sebenarnya ada dalam gambar.

Proses pelatihan model dilakukan selama 100 epochs dengan patience sebanyak 50 dan menggunakan batch size sebesar 16. Hasil ini menunjukkan bahwa model YOLO telah melewati pelatihan yang cukup lama untuk mencapai kinerja yang baik dalam tugas deteksi objek.

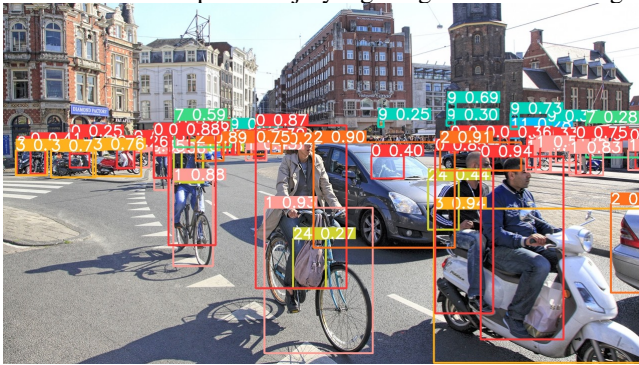


B. RT-DETR

Hasil evaluasi model RT-DETR yang menggunakan YOLO menunjukkan kinerja yang sangat baik. Metrik mAP50 mencapai nilai 0.96476, menunjukkan bahwa model ini mampu mengidentifikasi objek-objek dalam gambar dengan tingkat kepercayaan minimal sebesar 50%. Selain itu, presisi model sebesar 0.95243 menunjukkan bahwa sebagian besar objek yang dideteksi oleh

model adalah benar, sementara recall sebesar 0.91691 menunjukkan bahwa model dapat mengidentifikasi sebagian besar objek yang ada dalam gambar.

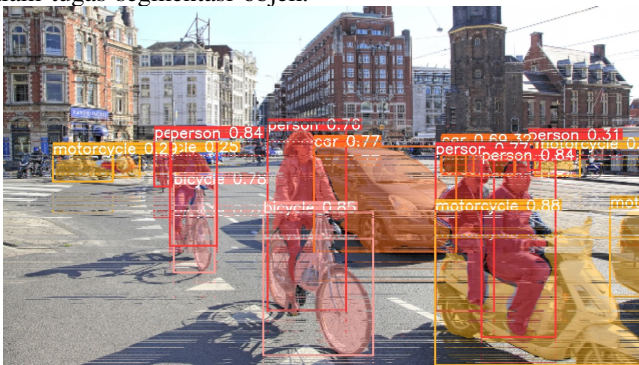
Proses pelatihan model dilakukan selama 100 epochs dengan patience sebanyak 50 dan menggunakan batch size sebesar 16. Hasil ini menunjukkan bahwa model RT-DETR yang menggunakan YOLO telah mengalami pelatihan yang cukup lama dan cermat untuk mencapai kinerja yang sangat baik dalam tugas deteksi objek secara real-time.



C. Segmentation

Hasil evaluasi model segmentasi menggunakan YOLO menunjukkan bahwa metrik mAP50 mencapai 0.89114, yang mengindikasikan kemampuan model dalam mengidentifikasi objek dengan tingkat kepercayaan minimal sebesar 50%. Dalam konteks segmentasi, presisi model mencapai 0.92324, menunjukkan bahwa sebagian besar objek yang dihasilkan oleh model sesuai dengan objek sebenarnya, sementara recall mencapai 0.83187, menunjukkan bahwa model mampu mengidentifikasi sebagian besar objek yang ada dalam gambar.

Proses pelatihan model dilakukan selama 100 epochs dengan sabar (patience) sebanyak 50 dan menggunakan batch size sebesar 16. Hasil ini menunjukkan bahwa model YOLO telah melalui pelatihan yang cukup lama untuk mencapai kinerja yang baik dalam tugas segmentasi objek.



V. KESIMPULAN

Hasil evaluasi model-model deteksi objek, RT-DETR, dan segmentasi menggunakan YOLO adalah sebagai berikut:

- Model deteksi objek YOLO menunjukkan kemampuan yang sangat baik dalam mengenali objek dalam gambar dengan tingkat kepercayaan sekitar 50%. Presisi model tinggi, sehingga sebagian besar objek yang ditemukan adalah benar, dan model dapat mengidentifikasi sebagian besar objek yang ada dalam gambar.
- Model RT-DETR menggunakan YOLO menunjukkan kinerja yang sangat baik dalam tugas deteksi objek. Tingkat kepercayaan minimal sekitar 50% sangat tinggi, dan presisi model tinggi, artinya objek yang ditemukan sebagian besar adalah benar. Model mampu mengidentifikasi sebagian besar objek dalam gambar.
- Model segmentasi YOLO baik dalam mengidentifikasi objek dengan tingkat kepercayaan sekitar 50%. Presisi tinggi, sehingga objek yang dihasilkan sesuai dengan objek sebenarnya, dan model mampu mengidentifikasi sebagian besar objek dalam gambar.