

Implementasi Ensemble YOLOv8 RT DETR untuk Deteksi Kontaminan Batubara

Muhammad Dhafa¹, Giant Prakoso Amukti Wibowo², Slamet Riyadi³✉
¹⁻³Universitas Muhammadiyah Yogyakarta, Indonesia

✉Corresponding Author: riyadi@mail.umy.ac.id

ABSTRAK

Deteksi batubara dan gangue merupakan tahap penting dalam proses penambangan untuk meningkatkan kualitas hasil tambang dan efisiensi operasional. Kemiripan visual antara batubara dan gangue serta kondisi lingkungan pertambangan yang kompleks menjadi tantangan utama dalam proses deteksi otomatis, selain itu dalam proses pengolahan batubara dan gangue sering kali tercampur sehingga mengakibatkan mesin crusher cepat rusak dan dapat mempengaruhi produksi. Penelitian ini mengusulkan metode deteksi objek berbasis ensemble dengan mengombinasikan model YOLOv8 dan RT-DETR. YOLOv8 memiliki keunggulan dalam kecepatan dan akurasi deteksi objek kecil, sedangkan RT-DETR mampu menangkap konteks global citra secara lebih efektif. Hasil eksperimen menunjukkan bahwa model YOLOv8 mencapai nilai mAP@0.5 sebesar 0.901 dan mAP@0.5:0.95 sebesar 0.781, sedangkan RT-DETR mencapai nilai mAP@0.5 sebesar 0.849 dan mAP@0.5:0.95 sebesar 0.713. Penggabungan kedua model melalui pendekatan ensemble menghasilkan performa deteksi yang lebih stabil dan andal dibandingkan penggunaan model tunggal. Metode yang diusulkan berpotensi diterapkan pada sistem deteksi otomatis di lingkungan pertambangan nyata.

Kata kunci : Deteksi Batubara, RT-DETR, YOLOv8, Otomatisasi

A. Pendahuluan

Batubara merupakan salah satu sumber energi utama yang hingga saat ini masih dimanfaatkan secara luas di berbagai sektor industri. Dalam proses penambangan dan pengolahan, batubara kerap tercampur dengan material pengotor berupa gangue. Keberadaan gangue tidak hanya menurunkan kualitas batubara, tetapi juga berpotensi menyebabkan kerusakan pada peralatan pemrosesan, seperti crusher, serta meningkatkan biaya operasional secara keseluruhan [1]. Oleh karena itu, diperlukan suatu sistem deteksi yang mampu membedakan batubara dan gangue secara akurat dan andal.

Metode konvensional yang umum digunakan, seperti inspeksi manual dan pemanfaatan sensor fisik, memiliki sejumlah keterbatasan, terutama dari sisi efisiensi, konsistensi, dan ketergantungan terhadap tenaga manusia [2], [3], [4]. Seiring dengan perkembangan teknologi deep learning, khususnya pada bidang deteksi objek berbasis visi komputer, pendekatan berbasis citra menjadi solusi yang lebih efektif dan fleksibel. Berbagai model deteksi objek, seperti YOLO dan DETR, telah menunjukkan kinerja yang menjanjikan dan banyak diterapkan pada beragam aplikasi industri.

Perkembangan teknologi deep learning yang cukup pesat, khususnya pada bidang deteksi objek berbasis visi komputer, telah membuka peluang baru dalam otomatisasi proses identifikasi batubara dan gangue [5] [6]. Berbagai penelitian menunjukkan bahwa model deteksi objek seperti YOLO dan DETR mampu memberikan performa yang baik dan telah diterapkan pada beragam aplikasi industri [7]. Pendekatan model tunggal tersebut cenderung memiliki keterbatasan dalam menghadapi variasi kondisi lingkungan pertambangan, seperti perubahan pencahayaan, kepadatan objek pada conveyor belt, serta kemiripan tekstur dan

warna antara batubara dan gangue. Selain itu, kajian yang secara khusus menggabungkan model lightweight YOLOv8 dengan model berbasis transformer RT-DETR untuk meningkatkan robustnes deteksi batubara dan gangue masih relatif terbatas. Oleh karena itu, diperlukan penelitian yang mengkaji pendekatan alternatif dengan menggabungkan keunggulan dari kedua arsitektur tersebut.

Tujuan dari penelitian ini adalah untuk membuat sistem monitoring, mengembangkan dan mengevaluasi sistem deteksi batubara dan gangue berbasis ensemble object detection dengan menggabungkan model YOLOv8 dan RT-DETR. Penelitian ini menganalisis performa deteksi batubara dan gangue menggunakan model YOLOv8 dan RT-DETR secara terpisah, kemudian mengimplementasikan pendekatan ensemble untuk menggabungkan hasil prediksi kedua model guna meningkatkan akurasi dan robustnes deteksi. Selanjutnya, dilakukan perbandingan performa antara model tunggal dan metode ensemble menggunakan metrik evaluasi standar deteksi objek, yaitu Precision, Recall, mAP@0.5, dan mAP@0.5:0.95.

B. Kajian Pustaka

Model You Only Look Once (YOLO) merupakan model deteksi yang banyak digunakan dalam penelitian karena keunggulannya dalam kecepatan inferensi dan efisiensi komputasi. Keandalan arsitektur YOLO dalam mengklasifikasikan berbagai objek secara akurat menggunakan deep learning telah dibuktikan secara komprehensif, termasuk dalam studi komparasi model YOLO [8]. Berbagai penelitian juga telah mengadaptasi dan mengembangkan arsitektur YOLO untuk meningkatkan akurasi deteksi pada lingkungan pertambangan batubara [9], [10]. Model deteksi batubara dan gangue berbasis YOLO dengan backbone ResNet18 untuk menghasilkan model yang lebih ringan dan sesuai untuk aplikasi real-time pada robot penyortir batubara. Dengan menerapkan strategi feature scale reduction dan unstructured pruning, model tersebut mampu mencapai nilai mAP sebesar

96,27% dengan ukuran model yang relatif kecil, sehingga cocok untuk perangkat dengan keterbatasan sumber daya komputasi [9]. Terkait penerapan pada perangkat bersumber daya terbatas, penelitian yang dipublikasikan dalam jurnal EIST juga menunjukkan bahwa performa dan portabilitas deteksi objek YOLO pada single-board computer dapat dioptimalkan secara signifikan menggunakan virtual machine dan framework spesifik, sehingga sangat mendukung implementasi real-time di lapangan [8] [11].

Seiring dengan berkembangnya YOLO generasi terbaru, beberapa penelitian mulai memanfaatkan YOLOv8 untuk meningkatkan performa deteksi. Mengembangkan model YOLOv8 yang ditingkatkan dengan penambahan Convolutional Block Attention Module (CBAM), fungsi loss EIou, serta teknik peningkatan kualitas citra menggunakan metode Laplacian. Model yang diusulkan berhasil mencapai mAP hingga 99,5% dengan ukuran model yang ringan, sehingga sangat potensial untuk sistem penyortiran batubara secara real-time. Selain itu, Zong et al. mengusulkan varian Improved YOLOv8n dengan integrasi Global Context Block (GCBlock), Context-Guided Feature Pyramid Network (CGFPN), serta GSConv-SlimNeck untuk meningkatkan deteksi objek berukuran kecil dan mengurangi kompleksitas model. Hasil eksperimen menunjukkan bahwa pendekatan ini mampu meningkatkan akurasi deteksi gangue pada kondisi pencahayaan rendah dan citra kabur, yang merupakan tantangan utama dalam lingkungan pertambangan [10].

Perkembangan model deteksi objek berbasis transformer memberikan pendekatan baru dalam identifikasi batubara dan gangue, khususnya pada lingkungan pertambangan yang memiliki kondisi visual kompleks [3]. Berbeda dengan metode konvensional dan model berbasis Convolutional Neural Network (CNN) yang menitikberatkan pada ekstraksi fitur lokal, arsitektur Detection Transformer (DETR) memanfaatkan mekanisme self-attention untuk

menangkap hubungan konteks global antarobjek secara end-to-end. Pendekatan ini dinilai lebih stabil karena tidak bergantung pada proses post-processing seperti Non- Maximum Suppression (NMS), yang sering menjadi sumber kesalahan pada deteksi objek berdekatan.

Dalam konteks deteksi batubara dan gangue, pendekatan berbasis konteks global sangat relevan karena karakteristik visual kedua objek yang sering kali memiliki tekstur, warna, dan bentuk yang saling menyerupai. Studi mengenai pengenalan batubara dan gangue berbasis citra menunjukkan bahwa informasi kontekstual dan hubungan antar fitur berperan penting dalam meningkatkan akurasi identifikasi, terutama pada kondisi pencahayaan rendah dan tingkat tumpukan material yang tinggi [12]. Selain itu, ketersediaan dataset citra batubara berskala besar juga mendorong pengembangan model deteksi berbasis deep learning yang lebih robust dan generalisasi [13], [14]. Pendekatan ensemble pada deteksi objek bertujuan untuk meningkatkan performa sistem dengan menggabungkan hasil prediksi dari beberapa model deteksi. Strategi ini efektif dalam mengurangi kesalahan deteksi yang dihasilkan oleh model tunggal, terutama pada kondisi visual yang kompleks. Penelitian sebelumnya menunjukkan bahwa ensemble berbasis penggabungan bounding box menggunakan nilai Intersection over Union (IoU) mampu menghasilkan deteksi yang lebih akurat dan stabil dibandingkan penggunaan satu model deteksi saja [15].

Metode ensemble umumnya diterapkan pada tingkat keputusan (decision-level fusion), di mana prediksi dari beberapa model disaring dan digabungkan untuk menghasilkan prediksi akhir yang lebih representatif. Pendekatan ini memungkinkan pemanfaatan keunggulan masing-masing model, seperti presisi tinggi atau kemampuan mendeteksi objek secara lebih menyeluruh, sehingga mampu meningkatkan keseimbangan antara presisi dan recall. Dalam aplikasi deteksi batubara dan gangue, pendekatan ensemble menjadi relevan karena karakteristik visual kedua objek yang serupa serta kondisi lingkungan pertambangan yang dinamis [12]. Oleh karena itu, penggabungan model berbasis CNN seperti YOLO dengan model berbasis transformer dinilai mampu meningkatkan robustness sistem deteksi dengan memanfaatkan keunggulan fitur lokal dan konteks global secara bersamaan.

C. Metode

Penelitian ini menerapkan metode deteksi objek dengan mengombinasikan model YOLOv8 dan RT-DETR sebagai pendekatan ensemble untuk meningkatkan akurasi deteksi batubara dan gangue. Metodologi penelitian disusun ke dalam beberapa tahap, meliputi penyediaan dataset, pelatihan masing-masing model, penggabungan hasil prediksi, serta evaluasi performa menggunakan metrik.

1. Data Collection

Dataset yang digunakan adalah dataset publik [16]. Dataset berisi citra batubara dan gangue yang telah dianotasi menggunakan metode bounding box untuk keperluan deteksi objek. Seluruh citra dianotasi ke dalam dua kelas, yaitu batubara (coal) dan gangue, dengan format anotasi yang kompatibel dengan model YOLO.

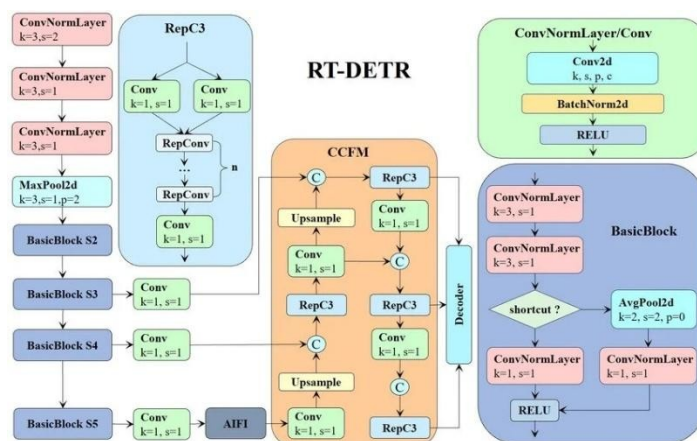
Jumlah total dataset yang digunakan dalam penelitian ini adalah 2.388 citra, yang mencerminkan variasi kondisi lingkungan pertambangan, seperti perbedaan pencahayaan, sudut pengambilan gambar, serta kondisi tumpukan material, dengan kelas coal sebanyak 14.779, gangue 2.088, dan object sebanyak 4. Variasi ini bertujuan untuk meningkatkan kemampuan generalisasi model terhadap kondisi nyata di lapangan. Dataset kemudian dibagi menjadi tiga bagian, yaitu train, val, dan test, dengan rasio pembagian train(70%), val(20%), dan test(10%).

2. Model YOLOv8

Model YOLOv8 varian nano (YOLOv8n) digunakan karena efisiensinya pada perangkat dengan sumber daya terbatas. Proses pelatihan dilakukan selama 50 epoch dengan ukuran citra 640×640 piksel, batch 8, dan workers 4. Parameter pelatihan disesuaikan untuk memperoleh keseimbangan antara akurasi dan waktu komputasi.

3. Model RT-DETR

Model RT-DETR (Real-Time Detection Transformer) digunakan sebagai model pembanding sekaligus pelengkap dalam penelitian ini. Berbeda dengan detektor satu tahap berbasis convolutional neural network (CNN) seperti YOLO yang berfokus pada ekstraksi fitur lokal, RT-DETR mengadopsi arsitektur berbasis transformer yang memanfaatkan mekanisme self-attention untuk menangkap hubungan konteks global pada citra. Pendekatan berbasis konteks global ini memungkinkan RT-DETR memodelkan relasi antar fitur secara menyeluruh tanpa ketergantungan pada proses post-processing yang kompleks. Dengan karakteristik tersebut, RT-DETR berpotensi menghasilkan deteksi yang lebih stabil pada lingkungan pertambangan dengan kondisi visual yang kompleks, seperti pencahayaan tidak merata, tumpukan material padat, serta kemiripan tekstur antara batubara dan gangue.

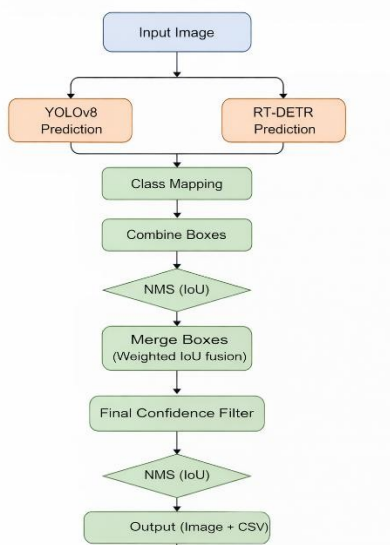


Gambar 1. Arsitektur RT-DETR [17]

Pada penelitian ini, model RT-DETR dilatih menggunakan pustaka Ultralytics dengan pretrained weights. Proses pelatihan dilakukan selama 5 epoch dengan ukuran citra 480×480 piksel, batch size 2, dan workers 4. Nilai learning rate awal ditetapkan sebesar 0,0001 ($lr_0 = 0.0001$). Pelatihan dijalankan pada perangkat CPU dan hasil pelatihan disimpan secara otomatis untuk memperoleh bobot terbaik (best weights). Konfigurasi pelatihan ini dirancang untuk menyesuaikan keterbatasan sumber daya komputasi sekaligus tetap memungkinkan model mempelajari pola visual batubara dan gangue dari dataset yang digunakan.

4. Ensemble

Penelitian ini mengusulkan metode ensemble object detection dengan mengombinasikan model YOLOv8 dan RT-DETR untuk meningkatkan akurasi dan stabilitas deteksi batubara dan gangue. Alur kerja metode ensemble ditunjukkan pada Gambar dibawah.



Gambar 2. Flowchart Ensemble

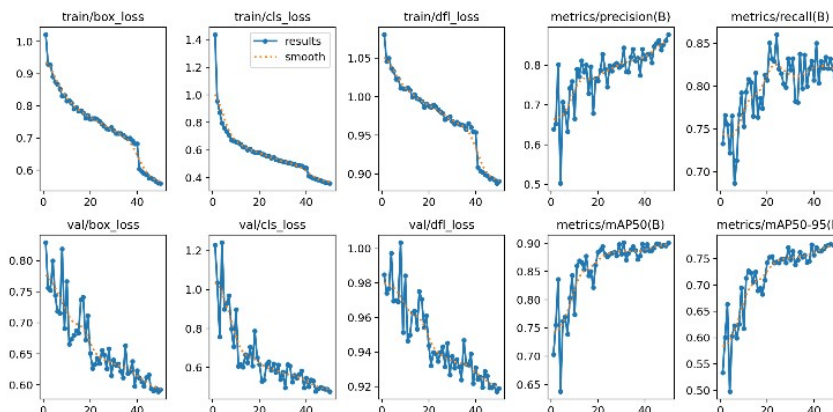
Citra uji diproses secara paralel oleh model YOLOv8 dan RT-DETR, yang masing-masing menghasilkan prediksi berupa bounding box, label kelas, dan nilai confidence. Untuk memastikan konsistensi label, dilakukan tahap class mapping guna menyelaraskan indeks kelas dari kedua model ke dalam skema kelas yang sama. Seluruh hasil prediksi kemudian digabungkan dan disaring menggunakan Non-Maximum Suppression (NMS) berbasis nilai Intersection over Union (IoU) untuk menghilangkan deteksi yang tumpang tindih. Bounding box yang masih memiliki tingkat tumpang tindih tinggi selanjutnya digabungkan menggunakan metode weighted IoU fusion, dengan koordinat akhir dihitung berdasarkan pembobotan nilai confidence. Hasil fusi kemudian difilter menggunakan ambang confidence akhir untuk menghapus deteksi dengan kepercayaan rendah. Output akhir berupa citra hasil deteksi yang telah diberi bounding box dan label kelas, serta data deteksi dalam format CSV, yang digunakan untuk analisis dan visualisasi lebih lanjut.

D. Hasil dan Pembahasan

Evaluasi performa model dilakukan untuk menilai kemampuan sistem dalam mendeteksi dan membedakan objek batubara dan gangue secara akurat. Proses evaluasi dilakukan menggunakan data uji (test set) yang tidak dilibatkan pada tahap pelatihan maupun validasi. Metrik evaluasi yang digunakan mengikuti standar pada tugas deteksi objek, yaitu Precision, Recall, mean Average Precision (mAP)_{@0.5}, dan mAP_{@0.5:0.95}.

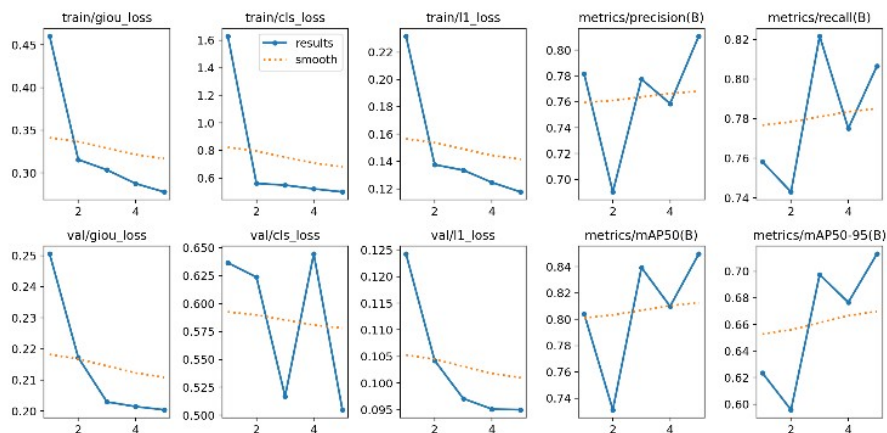
1. Evaluasi Proses Pelatihan

5



Gambar 3. Hasil Proses Evaluasi YOLOv8

Gambar hasil pelatihan menunjukkan bahwa model YOLOv8n mengalami konvergensi yang stabil selama 50 epoch. Nilai training loss dan validation loss pada komponen box loss, classification loss, dan distribution focal loss (DFL) menunjukkan tren menurun secara konsisten, yang mengindikasikan bahwa model mampu mempelajari representasi fitur objek dengan baik. Peningkatan nilai precision dan recall selama proses pelatihan juga menunjukkan peningkatan kemampuan model dalam mengidentifikasi objek batubara dan gangue. Nilai mAP@0.5 dan mAP@0.5:0.95 meningkat secara bertahap hingga mencapai kondisi stabil pada epoch akhir, menandakan performa deteksi yang optimal tanpa indikasi overfitting yang signifikan.



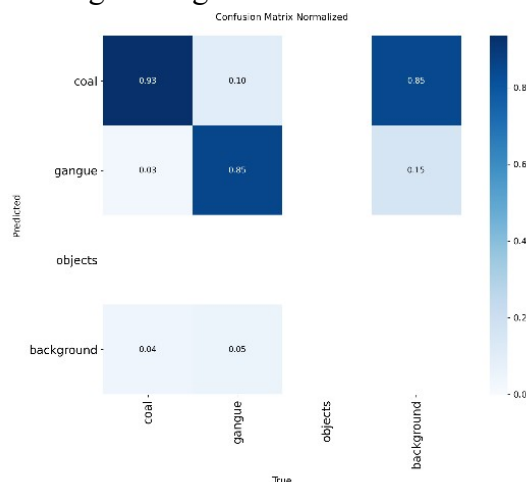
Gambar 4. Hasil Proses Evaluasi RT-DETR

Sebaliknya, proses pelatihan RT-DETR dilakukan dengan jumlah epoch yang lebih sedikit, yaitu 5 epoch, menyesuaikan dengan keterbatasan sumber daya komputasi. Meskipun demikian, grafik pelatihan menunjukkan tren penurunan loss yang konsisten serta peningkatan nilai mAP, yang mengindikasikan bahwa model tetap mampu mempelajari pola dasar objek batubara dan gangue.

2. Confusion Matrix

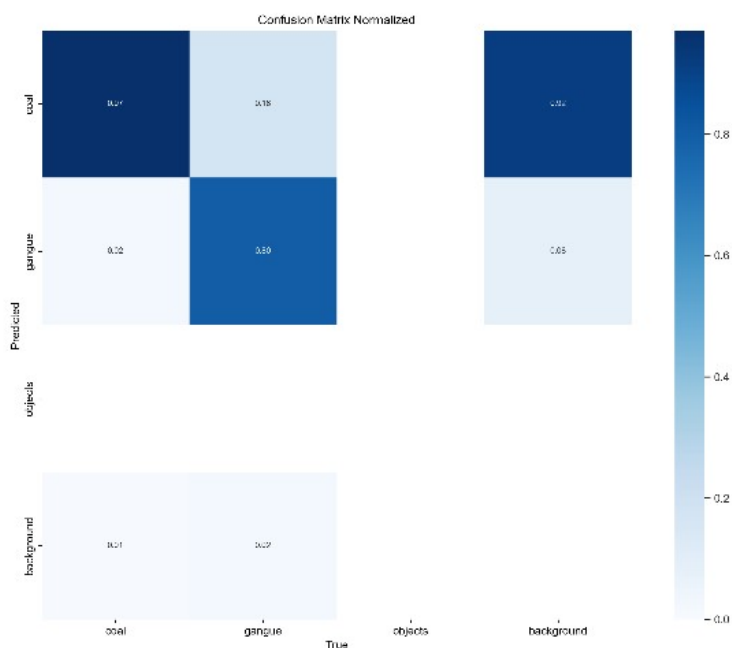
6

Analisis confusion matrix dilakukan untuk memahami distribusi prediksi benar dan kesalahan klasifikasi pada masing-masing kelas.



Gambar 5. Confusion Matrix YOLOv8

Berdasarkan confusion matrix ternormalisasi pada model YOLOv8, sebagian besar objek batubara (coal) berhasil diklasifikasikan dengan baik, dengan nilai recall kelas batubara mencapai lebih dari 90%. Hal ini menunjukkan kemampuan YOLOv8 yang kuat dalam mengenali objek batubara secara konsisten. Meskipun demikian, masih terdapat kesalahan klasifikasi antara kelas batubara dan gangue, yang terlihat dari adanya proporsi gangue yang diprediksi sebagai coal dan sebaliknya. Kesalahan ini terutama disebabkan oleh kemiripan visual antara kedua objek, seperti tekstur permukaan dan tingkat reflektansi yang serupa pada kondisi pencahayaan tertentu.



Gambar 6. RT-DETR Confusion Matrix

Confusion matrix pada model RT-DETR menunjukkan bahwa model ini masih mampu mengenali objek batubara dengan cukup baik, namun performa pada kelas gangue relatif lebih rendah dibandingkan YOLOv8. Hal ini terlihat dari meningkatnya jumlah objek gangue yang salah diklasifikasikan atau tidak terdeteksi. Kondisi tersebut mengindikasikan bahwa RT-DETR masih mengalami kesulitan dalam membedakan objek dengan karakteristik visual yang sangat mirip, terutama pada kelas minoritas. Hal ini sejalan dengan karakteristik arsitektur transformer yang umumnya memerlukan data pelatihan lebih banyak dan iterasi pelatihan yang lebih panjang untuk mencapai performa optimal.

3. Perbandingan Kinerja

Evaluasi kinerja dilakukan untuk membandingkan kemampuan deteksi model YOLOv8 dan RT-DETR dalam mengidentifikasi objek batubara dan gangue. Metrik yang digunakan meliputi Precision, Recall, F1-score, mAP@0.5, dan mAP@0.5:0.95, yang merupakan standar evaluasi pada tugas deteksi objek.

Tabel 1 Perbandingan Kinerja Model

Model	Precision	Recall	F1-score	mAP50	mAP50-95
YOLOv8	0.878	0.815	0.84	0.901	0.781
RT-DETR	0.811	0.806	0.81	0.849	0.713

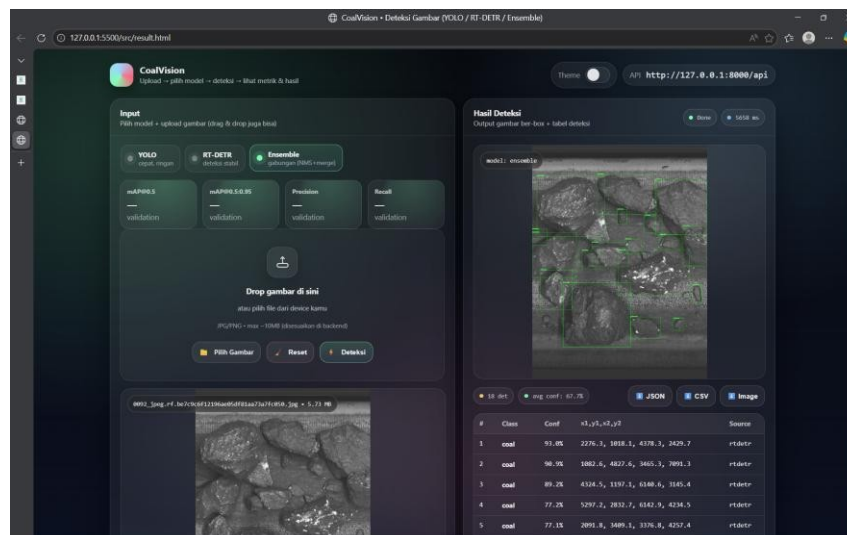
Hasil evaluasi menunjukkan bahwa YOLOv8 dan RT-DETR memiliki karakteristik kinerja yang saling melengkapi. YOLOv8 unggul dalam aspek ketepatan dan stabilitas deteksi, dengan nilai precision 0.878, F1-score 0.84, serta mAP@0.5 sebesar 0.901. Keunggulan ini menjadikan YOLOv8 efektif dalam mendeteksi objek berukuran kecil dan objek dengan jarak antarobjek yang rapat, yang umum dijumpai pada citra lingkungan pertambangan.

Sebaliknya, RT-DETR menunjukkan performa yang cukup kompetitif, dengan precision 0.811, recall 0.806, dan mAP@0.5 sebesar 0.849. Arsitektur berbasis transformer memungkinkan RT-DETR menangkap konteks global citra dengan lebih baik, sehingga berpotensi meningkatkan konsistensi deteksi pada kondisi visual yang kompleks. Namun, performa RT-DETR masih dipengaruhi oleh jumlah epoch pelatihan yang relatif terbatas dan kebutuhan data yang lebih besar untuk mencapai konvergensi optimal.

Berdasarkan perbedaan karakteristik tersebut, pendekatan ensemble diterapkan dengan menggabungkan hasil prediksi YOLOv8 dan RT-DETR. Strategi ini memanfaatkan keunggulan YOLOv8 dalam deteksi objek kecil dan kecepatan inferensi, serta kemampuan RT-DETR dalam memahami konteks global. Hasil penggabungan prediksi melalui proses Non-Maximum Suppression (NMS) dan penyaringan berbasis confidence menghasilkan deteksi yang lebih stabil dan akurat dibandingkan penggunaan model tunggal. Secara keseluruhan, pendekatan ensemble terbukti mampu meningkatkan keseimbangan antara precision dan recall, serta memberikan performa deteksi yang lebih robust pada skenario pertambangan dengan kondisi objek yang padat dan kompleks.

4. Implementasi Model ke Website

Sebagai bentuk implementasi dari model deteksi yang diusulkan, penelitian ini mengembangkan sebuah aplikasi berbasis web. Aplikasi ini dirancang untuk memvisualisasikan hasil deteksi batubara dan gangue menggunakan tiga pendekatan, yaitu model YOLOv8, model RT-DETR, dan metode ensemble yang menggabungkan kedua model tersebut.



Gambar 7. Implementasi Website

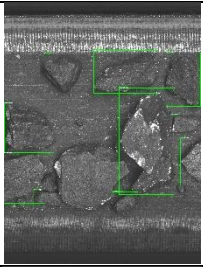
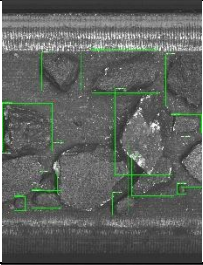
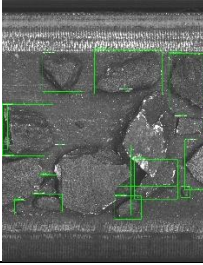
Pengguna dapat memilih model deteksi yang diinginkan, mengunggah citra uji, serta menampilkan hasil deteksi dalam bentuk bounding box pada citra. Selain visualisasi, sistem juga menyediakan informasi kuantitatif berupa jumlah objek terdeteksi, nilai rata-rata

confidence, serta tabel detail hasil deteksi yang mencakup kelas objek, skor confidence, dan koordinat bounding box. Hasil deteksi dapat diekspor dalam format JSON dan CSV.

5. Perbandingan Hasil Deteksi YOLOv8, RT-DETR, dan Ensemble

Gambar perbandingan menunjukkan perbedaan karakteristik keluaran deteksi pada tiga pendekatan, yaitu YOLOv8, RT-DETR, dan metode ensemble. Secara visual, ketiga model mampu mendeteksi objek utama berupa batubara dan gangue pada citra conveyor, namun kualitas prediksi yang dihasilkan berbeda dari sisi kelengkapan deteksi, konsistensi bounding box, serta jumlah prediksi yang muncul pada area padat objek.

Tabel 2 Perbandingan Hasil Deteksi

Model	Hasil
YOLOv8	
RT-DETR	
Ensemble	

Pada hasil YOLOv8, bounding box yang dihasilkan tampak lebih “rapi” dan stabil pada objek-objek besar. YOLOv8 juga cenderung menghasilkan prediksi yang lebih selektif sehingga jumlah kotak yang muncul tidak berlebihan. Hal ini menunjukkan bahwa YOLOv8 efektif dalam menangkap fitur lokal dan tepi objek, terutama pada area batubara yang dominan dan memiliki tekstur kuat. Namun, pada beberapa bagian citra yang padat atau memiliki objek kecil/menumpuk, terlihat potensi terjadinya missed detection (objek terlewat) atau sebagian objek terdeteksi tetapi bounding box belum sepenuhnya mengikuti kontur objek yang kompleks.

Sementara itu, hasil RT-DETR menunjukkan kecenderungan model untuk menangkap konteks yang lebih luas. Pada beberapa objek yang saling berdekatan, RT-DETR mampu memberikan bounding box yang mencakup area objek secara lebih menyeluruh. Akan tetapi, pada kondisi tertentu, RT-DETR dapat menghasilkan prediksi yang lebih sensitif sehingga muncul bounding box tambahan pada area yang ambigu. Hal ini wajar karena pendekatan

transformer memanfaatkan hubungan global antar fitur, namun dapat memerlukan pelatihan lebih panjang agar lebih konsisten pada objek kecil dan detail tekstur yang mirip.

Hasil ensemble memperlihatkan peningkatan kualitas secara umum karena menggabungkan kekuatan kedua model. Ensemble mampu mempertahankan deteksi stabil dari YOLOv8 sekaligus memanfaatkan kemampuan RT-DETR dalam menangkap konteks global. Dampaknya, beberapa objek yang sebelumnya terlewat pada salah satu model dapat muncul pada hasil gabungan. Selain itu, proses penggabungan prediksi (melalui penyamaan kelas, penggabungan bounding box, serta penyaringan IoU/NMS) membantu mengurangi prediksi yang tumpang tindih dan menghasilkan keluaran akhir yang lebih konsisten untuk kebutuhan aplikasi.

Secara keseluruhan, visualisasi ini mendukung hasil evaluasi kuantitatif bahwa pendekatan ensemble memberikan deteksi yang lebih robust, terutama pada kondisi lapangan yang kompleks seperti objek bertumpuk, perbedaan ukuran material, dan kemiripan tekstur batubara–gangué. Dengan demikian, pendekatan ensemble dinilai lebih sesuai untuk implementasi sistem deteksi otomatis di lingkungan pertambangan yang dinamis.

E. Simpulan

Penelitian ini mengusulkan sistem deteksi batubara dan gangue berbasis ensemble object detection dengan mengombinasikan model YOLOv8 dan RT-DETR. Hasil evaluasi menunjukkan bahwa YOLOv8 memiliki performa terbaik secara individual dengan nilai mAP@0.5 sebesar 0.901 dan mAP@0.5:0.95 sebesar 0.781, didukung oleh precision dan recall yang tinggi. Sementara itu, RT-DETR menunjukkan performa yang kompetitif dengan keunggulan pada pemodelan konteks global, meskipun akurasinya sedikit lebih rendah akibat keterbatasan konfigurasi pelatihan dan jumlah data.

Pendekatan ensemble mampu menggabungkan keunggulan kedua model, yaitu ketepatan deteksi objek kecil dari YOLOv8 dan pemahaman konteks global dari RT-DETR, sehingga menghasilkan deteksi yang lebih stabil dan robust pada kondisi visual pertambangan yang kompleks. Hasil visualisasi menunjukkan bahwa metode ensemble mampu mengurangi kesalahan klasifikasi serta meningkatkan konsistensi deteksi pada tumpukan material yang padat. Selain itu, implementasi sistem berbasis web membuktikan bahwa metode yang diusulkan dapat diterapkan secara praktis untuk mendukung proses monitoring dan pengolahan batubara secara real-time. Dengan demikian, penelitian ini memberikan kontribusi dalam pengembangan sistem deteksi otomatis batubara dan gangue yang akurat, efisien, dan aplikatif di lingkungan industri pertambangan.

F. Daftar Pustaka

- [1] R. Pahlevi, S. Thamrin, I. Ahmad, and F. B. Nugroho, “Masa Depan Pemanfaatan Batubara sebagai Sumber Energi di Indonesia,” *J. En. Baru & Terbarukan*, vol. 5, no. 2, pp. 50–60, Jul. 2024, doi: 10.14710/jebt.2024.22973.
- [2] P. Yan *et al.*, “Intelligent detection of coal gangue in mining Operations using multispectral imaging and enhanced RT-DETR algorithm for efficient sorting,” *Microchemical Journal*, vol. 207, p. 111789, Dec. 2024, doi: 10.1016/j.microc.2024.111789.
- [3] L. Wang *et al.*, “Intelligent real-time ash content detection for coal flotation concentrate using multi-source data fusion,” *Fuel*, vol. 406, p. 137132, Feb. 2026, doi: 10.1016/j.fuel.2025.137132.
- [4] L. Madhuanand, P. Sadavarte, A. J. H. Visschedijk, H. A. C. Denier Van Der Gon, I. Aben, and F. B. Osei, “Deep convolutional neural networks for surface coal mines determination from sentinel-2 images,” *European Journal of Remote Sensing*, vol. 54, no. 1, pp. 296–309, Jan. 2021, doi: 10.1080/22797254.2021.1920341.

- [5] A. R. Bohari, M. K. Hidayat, and D. A. Andianingsari, "Review: Identifikasi Batubara dan Gangue Menggunakan Machine Learning dan Deep Learning," *imtech*, vol. 5, no. 1, Jan. 2024, doi: 10.31294/imtechno.v5i1.3013.
- [6] G. E. Zagr , M. Gamache, R. Labib, and V. Shlenchak, "Machine learning algorithms for real-time coal recognition using monitor-while-drilling data," *International Journal of Mining, Reclamation and Environment*, vol. 38, no. 1, pp. 27–52, Jan. 2024, doi:10.1080/17480930.2023.2243783.
- [7] S. Bi, L. Shao, Z. Qi, Y. Wang, and W. Lai, "Prediction of coal mine gas emission based on hybrid machine learning model," *Earth Sci Inform*, vol. 16, no. 1, pp. 501–513, Mar. 2023, doi:10.1007/s12145-022-00894-5.
- [8] M. A. Haq, L. N. Q. Huy, and N. Fahriani, "Improving YOLO Object Detection Performance on Single-Board Computer using Virtual Machine," *EIST*, vol. 5, no. 1, May 2024, doi:10.18196/eist.v5i1.22486.
- [9] P. Yan, W. Wang, G. Li, Y. Zhao, J. Wang, and Z. Wen, "A lightweight coal gangue detection method based on multispectral imaging and enhanced YOLOv8n," *Microchemical Journal*, vol. 199, p. 110142, Apr. 2024, doi: 10.1016/j.microc.2024.110142.
- [10] G. Zong, Y. Yue, and W. Shan, "Optimization Study of Coal Gangue Detection in Intelligent Coal Selection Systems Based on the Improved Yolov8n Model," *Electronics*, vol. 13, no. 21, p. 4155, Oct. 2024, doi: 10.3390/electronics13214155.
- [11] Q. Liu, J. Li, Y. Li, and M. Gao, "Recognition Methods for Coal and Coal Gangue Based on Deep Learning," *IEEE Access*, vol. 9, pp. 77599–77610, 2021, doi: 10.1109/ACCESS.2021.3081442.
- [12] H. Bi, J. Chen, C. Jiang, and Q. Lin, "Study on mixed combustion characteristics and interaction mechanism of coal gangue/peanut shell under isothermal conditions," *Renewable Energy*, vol. 256, p. 124082, Jan. 2026, doi: 10.1016/j.renene.2025.124082.
- [13] Z. Lv et al., "A large-scale open image dataset for deep learning-enabled intelligent sorting and analyzing of raw coal," *Sci Data*, vol. 12, no. 1, p. 403, Mar. 2025, doi: 10.1038/s41597-025-04719-0.
- [14] A. M. Mayya and N. F. Alkayem, "Triple-stage crack detection in stone masonry using YOLO-ensemble, MobileNetV2U-net, and spectral clustering," *Automation in Construction*, vol. 172, p. 106045, Apr. 2025, doi: 10.1016/j.autcon.2025.106045.
- [15] S. A. Moghaddas, S. Ajayi, X. Wang, Y. Liu, Y. Huang, and Y. Bao, "Automatic multi-anomaly detection of pipelines with ensemble deep learning-based computer vision," *Advanced Engineering Informatics*, vol. 70, p. 104201, Mar. 2026, doi: 10.1016/j.aei.2025.104201.
- [16] "Coal&Gangue > Browse," Roboflow. Accessed: Mar. 31, 2026. [Online]. Available: <https://universe.roboflow.com/coalgangue/coal-gangue-nl309>
- [17] L. Liu, X. Yue, M. Lu, and P. He, Development of a RT-DETR-based model for detecting two wheelers *helmet*. 2024. doi: 10.21203/rs.3.rs-4242292/v1.