

Vision-based On Board Obstacle Detection System (VOODS) pada Sarana Perkeretaapian Menggunakan Pantauan Kamera

Diva Yulian Trisna¹, Sutrisno², Joko Hariyono^{3*}

^{1,2,3*} Program Studi Teknik Elektro, Pascasarjana Universitas Sebelas Maret, Surakarta, Indonesia

¹divayuliantrisna@student.uns.ac.id, ²sutrisno@uns.ac.id, ^{3*}jokohariyono@uns.ac.id

^{*)} jokohariyono@uns.ac.id

Abstrak—Industri perkeretaapian perkotaan menghadapi berbagai tantangan keselamatan, termasuk potensi tabrakan, kendaraan yang melintasi jalur rel, serta keberadaan objek asing di lintasan. Sistem inspeksi konvensional yang bergantung pada masinis memiliki keterbatasan pada kondisi jarak pandang terbatas dan kecepatan tinggi. Penelitian ini mengembangkan Vision-based On Board Obstacle Detection System (VOODS) berbasis kamera dengan algoritma YOLOv8 segmentation yang mampu mendeteksi dan mengklasifikasikan rintangan secara real-time pada jalur kereta. Sistem memanfaatkan kamera CCTV existing pada kereta, jaringan Ethernet sebagai jalur komunikasi, dan mini PC sebagai unit pemrosesan. Dataset dikumpulkan dari lingkungan operasional KRL Jabodetabek pada kondisi pagi dan siang hari, dengan total 300 gambar (240 training, 60 validasi) dan empat kelas (Rail, Train, Human, Car). Anotasi dilakukan menggunakan platform LabelMe (Anaconda) dengan pertimbangan kerahasiaan data karena seluruh dataset disimpan pada PC pribadi peneliti. Model yolov8m-seg dilatih melalui Google Colaboratory dengan parameter epochs=150, imgsz=1024, dan patience=30. Penentuan zona bahaya dilakukan secara dinamis menggunakan mask hasil segmentasi rel sebagai Region of Interest (ROI), kemudian sistem warning memeriksa overlap antara mask rel dengan mask objek menggunakan operasi bitwise AND untuk menentukan apakah objek berada on track. Pendekatan ini mengacu pada profil ruang bebas yang diatur dalam PM No. 60 Tahun 2012. Hasil evaluasi menunjukkan mAP@0.5 (Box) sebesar 0,959 untuk seluruh kelas. Pada level per-kelas, Rail mencapai 0,995, Train 0,989, Human 0,933, dan Car 0,918. Analisis Confusion Matrix menunjukkan True Positive Rate sebesar 1,00 untuk Rail, 0,96 untuk Train dan Car, serta 0,88 untuk Human. Sistem warning telah berhasil diuji pada video operasional dan dapat memunculkan peringatan ketika objek manusia masuk ke area rel. Nilai Recall yang tinggi merupakan capaian positif karena dalam konteks keselamatan kereta, Recall lebih kritis dibanding Precision. VOODS yang dikembangkan memberikan kontribusi signifikan dalam peningkatan keselamatan operasional transportasi perkeretaapian perkotaan di Indonesia.

Kata Kunci: Vision-based On Board Obstacle Detection System, YOLOv8 Segmentation, Deteksi Rintangan, Keselamatan Kereta, Rail Boundary System

Abstract—Urban railway industries face various safety challenges, including potential collisions, vehicles crossing rail tracks, and foreign objects on the track. Conventional inspection systems relying on train operators have limitations under restricted visibility and high speeds. This study develops a camera-based Vision-based On Board Obstacle Detection System (VOODS) using the YOLOv8 segmentation algorithm capable of detecting and classifying obstacles in real-time on railway tracks. The system utilizes existing CCTV cameras on trains, Ethernet network as communication channel, and a mini PC as the processing unit. A dataset was collected from the KRL Jabodetabek operational environment under morning and daytime lighting conditions, comprising 300 images (240 training and 60 validation) across four classes (Rail, Train, Human, Car). Annotation was performed using the LabelMe (Anaconda) platform, considering data confidentiality since the entire dataset is stored on the researcher's local PC. The yolov8m-seg model was trained via Google Colaboratory with epochs=150, imgsz=1024, and patience=30. Danger zone determination is performed dynamically using the rail segmentation mask as Region of Interest (ROI); a warning subsystem then checks overlap between rail mask and object mask using bitwise AND operation to determine on-track presence. This approach refers to the clearance profile defined in Indonesian Ministerial Regulation PM No. 60 Year 2012. Evaluation results show a Box mAP@0.5 of 0.959 across all classes. Per-class performance: Rail 0.995, Train 0.989, Human 0.933, and Car 0.918. Confusion Matrix analysis shows a True Positive Rate of 1.00 for Rail, 0.96 for Train and Car, and 0.88 for Human. The warning subsystem was successfully tested on operational video and can trigger alerts when a human enters the rail area. The high Recall value is positive because, in railway safety context, Recall is more critical than Precision. The developed VOODS contributes significantly to improving operational safety of urban railway transportation in Indonesia.

Keywords: Vision-based On Board Obstacle Detection System, YOLOv8 Segmentation, Obstacle Detection, Railway Safety, Rail Boundary System

1. PENDAHULUAN

Industri perkeretaapian, khususnya transportasi rel perkotaan, memiliki peranan vital dalam mendukung mobilitas masyarakat dengan karakteristik kapasitas angkut besar, stabilitas operasional tinggi, ketepatan waktu, serta efisiensi energi. Sistem transportasi ini dirancang untuk melayani pergerakan penumpang secara masif dan berkelanjutan, sehingga aspek keselamatan menjadi faktor utama yang tidak dapat ditawar [1]. Namun demikian, dalam praktik operasional masih sering dijumpai berbagai tantangan keselamatan, seperti potensi tabrakan antar kereta, kendaraan yang melintas di jalur rel, serta keberadaan objek asing berupa material konstruksi, sampah besar, atau pohon tumbang yang menghalangi lintasan. Pada konteks operasional LRT Jabodebek di mana operasional perjalanan kereta sudah menggunakan jalur *elevated*, berbagai jenis kecelakaan dan gangguan keselamatan masih berpotensi terjadi. Kondisi tersebut tidak hanya berpotensi menyebabkan gangguan operasional dan keterlambatan perjalanan, tetapi juga dapat menimbulkan risiko kecelakaan serius yang membahayakan keselamatan penumpang dan kru operasional. Kejadian berupa batang pohon yang jatuh dan membentang di jalur rel, jatuhnya baliho ke area rel, material pembangunan gedung tinggi di sekitar jalur yang terjatuh di lintasan, risiko *rear end collision*, penumpang yang terjatuh ke area rel, serta keberadaan *foreign objects* yang mengganggu ruang bebas kereta merupakan contoh nyata yang perlu diantisipasi.

Seiring meningkatnya tingkat otomatisasi dan berkurangnya intervensi manusia dalam sistem transportasi rel perkotaan, ketergantungan terhadap sistem keselamatan pasif saja tidak lagi memadai. Sistem konvensional yang mengandalkan inspeksi visual oleh masinis atau laporan dari petugas lapangan memiliki keterbatasan, terutama pada kondisi jarak pandang terbatas, kecepatan kereta yang tinggi, serta dinamika lingkungan lintasan yang berubah secara cepat. Oleh karena itu, dibutuhkan sistem keselamatan aktif yang mampu mendeteksi potensi bahaya secara dini serta memberikan peringatan yang cepat dan akurat. Perkembangan teknologi *Vision Camera*, pengolahan data berbasis *machine learning*, dan penerapan sistem cerdas di industri perkeretaapian membuka peluang penerapan *Vision-based On Board Obstacle Detection System* (VOODS) sebagai solusi strategis untuk meningkatkan keselamatan operasional. VOODS memanfaatkan sistem kamera yang terpasang pada kereta dan unit pengolahan data sebagai komputasinya. Melalui pemrosesan citra dan algoritma kecerdasan buatan, sistem ini mampu melakukan deteksi, klasifikasi, serta pelacakan objek secara *real-time*, sekaligus memberikan peringatan dini terhadap keberadaan *obstacle* yang berpotensi membahayakan perjalanan kereta. Secara prinsip, VOODS bekerja dengan membangun model visual jalur dan lingkungan lintasan dari kamera, kemudian secara dinamis menentukan area *Train Clearance Envelope* (Ruang Bebas Kereta) di depan kereta untuk mengevaluasi apakah terdapat hambatan yang masuk ke dalam ruang tersebut.

Sistem deteksi rintangan pada kereta telah banyak diteliti dengan berbagai pendekatan sensor dan algoritma. Zhao et al. [1] mengembangkan metode deteksi rintangan *real-time* menggunakan *lightweight CNN* dan *improved transformer* yang mengintegrasikan kemampuan ekstraksi fitur CNN dengan pemahaman konteks spasial *transformer*, sehingga mampu mendeteksi objek kecil dan jauh secara efisien pada perangkat onboard. Li dan Zhang [2] mengusulkan kombinasi *road segmentation* dan *multi-class obstacle detection* menggunakan *encoder-decoder* untuk segmentasi jalur dan *machine learning detector* untuk deteksi *obstacle* pada kereta otonom. Xu et al. [3] melakukan kajian komprehensif berbagai metode deteksi intrusi *obstacle* pada transportasi rel, mencakup sensor kamera visual, inframerah, sensor fusion, serta pendekatan *machine learning*. Pendekatan berbasis *deep learning* juga banyak dikembangkan untuk deteksi rintangan pada rail transit, seperti *Improved Mask R-CNN* [4] dan *Improved R-CNN* [5], serta metode deteksi berbasis *deep learning* pada transportasi rel perkotaan [6]. Chen et al. [7] mengembangkan RH-Net yang menggabungkan *lightweight CNN* dengan *improved transformer* untuk deteksi rintangan kereta secara *real-time* dan akurat, sementara Liu et al. [8] mengusulkan LDCCD-Net, yaitu jaringan deteksi *dual-channel* ringan yang efisien untuk perangkat komputasi terbatas. Selain pendekatan berbasis kamera, beberapa penelitian memanfaatkan fusi sensor untuk meningkatkan keandalan deteksi, seperti fusi radar dan visual [9] serta fusi LiDAR-kamera pada area jalur kereta [10]. Pengembangan sistem deteksi rintangan cerdas pada jalur rel juga telah dilakukan melalui berbagai inisiatif [11][12], sejalan dengan perkembangan sistem *surveillance* cerdas berbasis *computer vision* dan *machine learning* secara umum [13].

Penelitian-penelitian tersebut menunjukkan kemajuan yang signifikan, namun masih memiliki beberapa keterbatasan: belum membahas integrasi sistemik ke dalam arsitektur keselamatan kereta secara komprehensif, belum menerapkan evaluasi berdasarkan *Train Clearance Envelope*, serta belum menggunakan YOLOv8 *segmentation* generasi terkini dengan dataset yang disesuaikan lingkungan perkeretaapian Indonesia.

Penelitian ini mengembangkan VOODS berbasis kamera dengan algoritma YOLOv8 *segmentation* yang secara khusus diadaptasi untuk lingkungan perkeretaapian di Indonesia. Kebaruan penelitian ini terletak pada penggunaan *instance segmentation* untuk menghasilkan mask rel yang berfungsi sebagai *Region of Interest (ROI)* dinamis, mekanisme penentuan on-track berbasis overlap antara mask rel dengan mask objek, penetapan zona bahaya yang

mengacu pada Train Clearance Envelope sesuai PM No. 60 Tahun 2012, serta penerapan platform anotasi LabelMe (Anaconda) yang memastikan kerahasiaan data karena seluruh dataset disimpan pada PC pribadi peneliti. Penelitian ini bertujuan untuk: (1) merancang arsitektur sistem VOODS yang efisien menggunakan kamera sebagai sumber data utama; (2) melatih model YOLOv8-seg untuk mendeteksi dan mengelompokkan objek manusia, kendaraan, kereta, dan rel pada citra kamera depan kereta; serta (3) membangun mekanisme warning berbasis overlap mask rel dan mask objek untuk mengidentifikasi obstacle yang berada di area lintasan.

2. METODOLOGI PENELITIAN

2.1 Arsitektur Sistem VOODS

Arsitektur sistem VOODS yang dikembangkan dirancang modular dan kompatibel dengan infrastruktur yang tersedia pada kereta, sepenuhnya memanfaatkan kamera CCTV *existing* sebagai sumber data visual utama dan jaringan Ethernet *existing* sebagai jalur komunikasi antar komponen. Aliran video dari kamera CCTV diteruskan ke Ethernet switch, kemudian diproses oleh *Controller* (mini PC) sebagai unit pemroses utama menggunakan model YOLOv8-seg. Hasil deteksi dan segmentasi selanjutnya digunakan untuk mengontrol keluaran digital menuju *alarm* atau *loop Emergency Brake*.

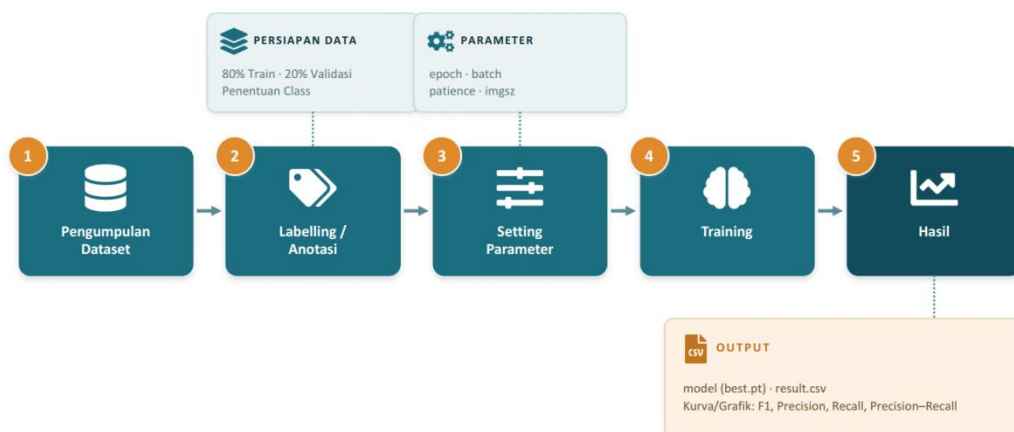
Sistem *onboard* utama terdiri dari dua unit: (a) Unit Host, terdiri dari *Power Supply Unit* (PSU), *Network Video Recorder* (NVR), dan *Ethernet Switch* (SW), dengan modul opsional *Digital Input/Output* (DIO) dan *network TRDP/MVB*; dan (b) *Front-End Perception Unit*, yang terdiri dari kamera CCTV yang dipasang di bagian depan kabin masinis. Pendekatan ini mempermudah integrasi dengan sistem kereta yang sudah ada tanpa perlu penambahan sensor mahal seperti LiDAR.

Tabel 1. Perangkat yang Digunakan dalam Sistem VOODS

No	Nama Perangkat	Fungsi
1	Kamera CCTV	Sumber data visual utama dari depan kereta
2	Ethernet Switch	Jalur komunikasi data antar komponen
3	Mini PC (Controller)	Unit pemrosesan YOLOv8-seg dan pengambilan keputusan
4	Power Supply Unit	Sumber daya seluruh komponen sistem
5	Kabel Ethernet	Koneksi fisik antar perangkat

2.2 Alur Proses Penelitian

Sebelum membahas masing-masing tahapan secara detail, Gambar 1 menyajikan alur proses penelitian pengembangan model deteksi objek YOLO yang diterapkan pada sistem VOODS. Alur ini menjadi kerangka kerja yang menghubungkan tahap pengumpulan data dengan keluaran berupa model terlatih beserta metrik evaluasinya.



Gambar 1. Alur Proses Pelatihan Model Deteksi Objek (YOLO)

Proses pelatihan model dilakukan melalui lima tahap utama. Tahap pertama adalah Pengumpulan Dataset, yaitu pengambilan rekaman citra dari kamera CCTV pada kabin masinis KRL Jabodetabek sebagai sumber data mentah. Tahap kedua adalah Labelling/Anotasi, di mana setiap objek pada citra ditandai dengan kelas dan koordinat geometrinya menggunakan platform LabelMe; pada tahap ini dataset juga dibagi 80% untuk training dan 20% untuk validasi, serta ditetapkan empat kelas yaitu Rail, Train, Human, dan Car. Tahap ketiga adalah Setting Parameter, yaitu konfigurasi *hyperparameter* training meliputi *epoch*, *batch*, *patience*, dan *imgsz* yang menentukan kedalaman dan stabilitas proses pelatihan. Tahap keempat adalah Training, yaitu eksekusi pelatihan model YOLOv8m-seg di lingkungan Google Colaboratory dengan akselerasi GPU. Tahap kelima adalah Hasil, yang menghasilkan bobot model terbaik yang sering disebut *best.pt*, file *result.csv* yang merekam metrik per *epoch*, serta kurva F1, Precision, Recall, dan Precision-Recall sebagai dasar evaluasi system.

Tahap pertama dan kedua dijelaskan secara detail pada bab Pengumpulan dan Anotasi Dataset, sedangkan tahap ketiga dan keempat dijelaskan pada poin 2.4 Pemrosesan Dataset dan Training Model.

2.3 Pengumpulan dan Anotasi Dataset

Pengumpulan dataset dilakukan di lingkungan operasional Kereta Rel Listrik (KRL) Jabodetabek pada area stasiun dan jalur layanannya. Data visual diperoleh dari kamera CCTV pada kabin masinis yang mengarah ke jalur rel, mencakup kondisi pencahayaan pagi dan siang hari, dengan berbagai jenis *obstacle* seperti manusia, mobil, kereta, dan rel kereta. Pengumpulan dataset pada kondisi sore dan malam hari belum dilakukan pada tahap penelitian ini, dan menjadi rekomendasi untuk pengembangan lanjutan.

Proses anotasi dataset untuk training model dilakukan menggunakan platform LabelMe pada lingkungan Anaconda. LabelMe dipilih karena mendukung berbagai bentuk anotasi yang dibutuhkan, mencakup *rectangle*, *polygon*, *line*, dan *mask*, sehingga memungkinkan pembuatan anotasi rel berbentuk *polygon* yang mengikuti kontur jalur rel secara presisi. Selain itu, LabelMe menyimpan seluruh data anotasi pada PC pribadi peneliti, sehingga kerahasiaan data rekaman CCTV operasional dapat terjaga sepenuhnya.

Anotasi LabelMe menghasilkan file *.json* yang berisi koordinat *x*, *y*, dan kelas objek untuk setiap gambar. File tersebut kemudian dikonversi ke format *.txt* (format YOLO) untuk dapat diproses oleh model YOLOv8-seg. Dataset terdiri dari empat kelas objek utama: Rail (rel), Train (kereta), Human (manusia), dan Car (mobil), dengan total 300 gambar yang dibagi menjadi 240 gambar untuk training dan 60 gambar untuk validasi.

2.4 Pemrosesan Dataset dan Training Model

Model deteksi yang digunakan adalah YOLOv8 (*You Only Look Once*, versi 8) varian *segmentation*. YOLOv8-seg bekerja dengan prinsip *single-shot detection* di mana prediksi *bounding box*, kelas objek, dan *segmentation mask* dilakukan dalam satu proses inferensi, sehingga memungkinkan kecepatan pemrosesan tinggi untuk aplikasi *real-time*. Model menerima input citra *x* dan menghasilkan prediksi $y = f(x; \theta)$, di mana *f* adalah model YOLOv8-seg dan θ adalah parameter bobot model hasil pelatihan [14][15].

Pemilihan pendekatan *instance segmentation* dilandasi kemampuannya menghasilkan batas objek pada tingkat piksel untuk setiap *instance*, sehingga lebih representatif untuk objek berbentuk memanjang dan tidak beraturan seperti jalur rel dibandingkan representasi *bounding box* persegi yang selalu menyertakan area latar di luar [16]. YOLOv8-seg dipilih dalam penelitian ini karena tiga alasan utama: (1) kemampuan mendeteksi objek kecil, objek jauh, dan objek yang sebagian tertutup (*occluded*); (2) menghasilkan *mask* geometri presisi yang menjadi fondasi mekanisme *on-track detection* pada sistem warning; dan (3) kompatibilitas dengan berbagai perangkat *edge computing* termasuk mini PC yang digunakan sebagai *controller* dalam sistem VOODS [17].

Pra-pemrosesan dataset pada penelitian ini terbatas pada dua kegiatan: pelabelan/anotasi objek pada platform LabelMe dan seleksi gambar unik dari hasil rekaman video kamera CCTV untuk menghindari duplikasi frame yang terlalu mirip. Tidak dilakukan langkah normalisasi piksel maupun augmentasi manual di tahap pra-pemrosesan, karena augmentasi ditangani secara otomatis oleh *framework* YOLOv8 selama proses training.

Proses training model YOLOv8-seg dilakukan menggunakan platform Google Colaboratory dengan *runtime* GPU Tesla T4 untuk mempercepat komputasi. Selama training, *framework* YOLOv8 secara otomatis menerapkan augmentasi data untuk meningkatkan keberagaman dataset dan generalisasi model, mencakup *Blur* ($p=0,01$), *MedianBlur* ($p=0,01$), *ToGray* ($p=0,01$), dan *CLAHE* (Contrast Limited Adaptive Histogram Equalization, $p=0,01$). Augmentasi ini diterapkan secara stokastik per *epoch* sehingga model dapat belajar dari variasi sintesis tanpa mengubah dataset asli [18]. Perintah training yang digunakan adalah:

```
!yolo segment train data=/content/data.yaml model=yolov8m-seg.pt epochs=150
imgsz=1024 patience=30
```

Parameter training yang digunakan adalah: *task=segment*, *model=yolov8m-seg.pt* (*pretrained YOLOv8 versi medium dengan kemampuan segmentation*) [19], *epochs=150*, *imgsz=1024* (resolusi input citra), dan *patience=30* untuk *early stopping*. Penentuan jumlah *epoch* tidak memiliki nilai baku karena bergantung pada karakteristik *dataset*, seperti jumlah data, keragaman data, dan tingkat kompleksitas permasalahan. Semakin besar dan beragam *dataset* yang digunakan, umumnya semakin banyak *epoch* yang diperlukan agar model mencapai konvergensi. Sebaliknya, pada *dataset* yang relatif kecil, penggunaan *epoch* yang terlalu banyak dapat meningkatkan risiko *overfitting*, sehingga diperlukan pemilihan jumlah *epoch* yang sesuai [20]. *Patience* pada YOLO digunakan sebagai mekanisme *early stopping*, yaitu menghentikan proses pelatihan secara otomatis apabila performa model pada data validasi tidak menunjukkan peningkatan selama sejumlah *epoch* yang telah ditentukan [21]. Penggunaan parameter ini bertujuan untuk mencegah pelatihan yang berlebihan (*overfitting*) dan menghemat waktu komputasi. *Batch size* awal diset 16, namun secara otomatis diturunkan menjadi 8 oleh sistem akibat keterbatasan memori GPU. Arsitektur YOLOv8m-seg yang digunakan terdiri dari 192 *layer* dengan total 27.241.964 parameter dan kebutuhan komputasi 104,7 GFLOPs. Hasil *training* menghasilkan model bobot terbaik (*best.pt*) yang digunakan untuk tahap pengujian dan implementasi.

2.5 Mekanisme On-Track Warning berbasis Mask Overlap

Train Clearance Envelope (Ruang Bebas Kereta) merupakan profil dua dimensi yang mendefinisikan zona ruang minimum aman di sekeliling kereta yang harus bebas dari segala rintangan dan benda penghalang. Dalam regulasi perkeretaapian Indonesia, konsep ini diatur secara resmi melalui Peraturan Menteri Perhubungan Republik Indonesia Nomor PM 60 Tahun 2012 tentang Persyaratan Teknis Jalur Kereta Api [22]. Berdasarkan regulasi tersebut, jalur kereta api harus memiliki pengaturan ruang yang terdiri dari ruang bebas dan ruang bangun. Ruang bebas didefinisikan sebagai ruang di atas jalan rel yang senantiasa harus bebas dari segala rintangan, sedangkan ruang bangun adalah ruang di sisi jalan rel yang harus bebas dari segala bangunan tetap, dengan batas yang diukur dari sumbu jalan rel pada ketinggian 1 meter sampai 3,55 meter.

Regulasi ini mengatur dimensi ruang bebas untuk dua lebar jalan rel yang berlaku di Indonesia, yaitu 1067 mm dan 1435 mm, dan mengklasifikasikan empat kategori batas ruang bebas: Batas I untuk jembatan kecepatan sampai 60 km/jam, Batas II untuk viaduk dan terowongan, Batas III untuk viaduk baru dan bangunan lama, dan Batas IV untuk lintas kereta listrik. Pada konteks penelitian ini yang berfokus pada KRL Commuter Line Jabodetabek, Batas IV menjadi acuan utama karena seluruh sarana yang dilayani merupakan kereta listrik. Dalam implementasi VOODS, profil ruang bebas ini menjadi acuan pembentukan Region of Interest (ROI) pada bidang citra kamera CCTV. Implementasi awal pada penelitian ini berbasis mask segmentasi rel sebagai ROI dinamis per frame, dengan pengembangan lanjutan menuju proyeksi ruang bebas penuh sesuai dimensi regulasi.

Mekanisme penentuan apakah suatu objek berada *on-track* (di area jalur rel sehingga berbahaya) menggunakan operasi geometri pada *segmentation mask* hasil YOLOv8-seg, bukan polygon ROI manual statis. Pendekatan ini memberikan dasar ROI yang adaptif terhadap citra per frame, mengacu pada profil ruang bebas yang diatur dalam Peraturan Menteri Perhubungan Republik Indonesia Nomor PM 60 Tahun 2012 tentang Persyaratan Teknis Jalur Kereta Api [22]. Implementasi awal mengikuti area lintasan rel sesuai lebar 1067 mm pada lintas KRL Jabodetabek. Algoritma *on-track warning* beroperasi dalam tiga tahap untuk setiap frame: (1) seluruh deteksi dengan *confidence* di atas threshold (default 0,492) diidentifikasi kelasnya, dan untuk setiap kelas Rail dibentuk *rail_mask* kumulatif menggunakan *cv2.fillPoly*; (2) untuk objek non-rel (Human, Car, Train), dibentuk *obj_mask* individual, kemudian dihitung *intersection* antara *rail_mask* dan *obj_mask* menggunakan operasi *cv2.bitwise_and*; jika jumlah piksel overlap melebihi *overlap_thresh* (default 10 piksel), objek tersebut diberi flag *on track* dan *warning_active* di-set *True*; (3) sistem me-render visualisasi dengan warna berbeda — objek *on track* diberi warna merah, objek di luar rel diberi warna oranye, dan area rel diberi mask merah transparan.

Status global frame ditampilkan pada bagian atas citra sebagai teks "Status: WARNING — [class] ([confidence]%) *on track!*" ketika ada objek *on track*, atau "Status: Clear" ketika tidak ada objek pada area rel. Mekanisme ini dirancang sebagai dasar untuk pemecuan sinyal pengereman darurat melalui keluaran *Digital I/O* pada mini PC kontroler.

2.6 Mekanisme Peringatan Bertingkat

Berdasarkan lokasi hambatan, jenisnya, dan hubungannya dengan kecepatan kereta secara dinamis, sistem melakukan manajemen peringatan bertingkat melalui logika respons yang dikonfigurasi. Sistem mendukung kategorisasi tingkat bahaya objek: (1) kelas hambatan yang dikenali seperti manusia, kendaraan, hewan, dan peralatan; (2) hambatan tidak dikenal (*Unknown Object*); dan (3) bukan hambatan seperti rambu atau kotak listrik di pinggir rel. Ketika potensi ancaman terdeteksi *on track*, sistem mengeluarkan peringatan *real-time* kepada masinis melalui perangkat audio-visual, dan antarmuka kontrol pengereman darurat disediakan sebagai cadangan.

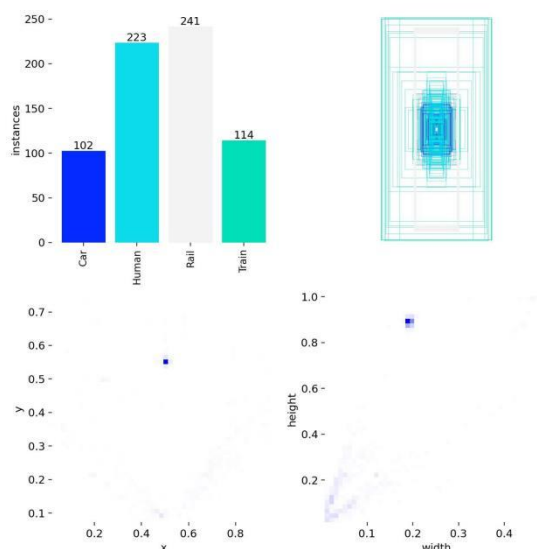
3. HASIL DAN PEMBAHASAN

3.1 Pertimbangan Pemilihan Platform Anotasi

Proses anotasi dataset merupakan tahap fundamental dalam pengembangan model YOLOv8-seg, di mana kualitas anotasi secara langsung mempengaruhi kemampuan model dalam mengenali pola visual obstacle pada citra kamera kereta. Pada tahap eksplorasi awal penelitian, anotasi dilakukan menggunakan platform Roboflow karena antarmuka berbasis web yang intuitif, integrasi langsung dengan ekosistem YOLOv8 melalui API, serta kemudahan kolaborasi dalam pengelolaan dataset. Project Roboflow yang dibuat pada tahap ini disetup dengan tipe object detection menggunakan bounding box rectangle pada objek manusia, mobil, dan kereta.

Pada perkembangan penelitian, kebutuhan berubah ketika diidentifikasi bahwa anotasi jalur rel memerlukan bentuk polygon yang mengikuti kontur jalur agar dapat digunakan sebagai dasar Region of Interest (ROI) yang presisi. Roboflow sebenarnya mendukung anotasi polygon untuk *instance segmentation*, namun karena project Roboflow yang sudah dijalankan telah dikonfigurasi sebagai object detection, mengubah tipe project ke instance segmentation berarti project harus disetup ulang dari awal dan seluruh anotasi yang telah dikerjakan harus diulang. Pertimbangan ini menjadi salah satu alasan untuk mengevaluasi alternatif platform.

Pertimbangan kedua yang lebih substantif adalah aspek kerahasiaan data. Roboflow menyimpan seluruh dataset di *cloud server* platform tersebut, sementara dataset penelitian ini bersumber dari rekaman CCTV operasional KRL Jabodetabek yang memerlukan jaminan kerahasiaan data. Penyimpanan dataset pada *cloud third-party* menimbulkan risiko terkait kontrol akses, kepemilikan data, dan kepatuhan terhadap aturan internal operator perkeretaapian. Berdasarkan dua pertimbangan tersebut kebutuhan untuk memulai project segmentation dari awal dan kebutuhan jaminan kerahasiaan data penelitian memutuskan untuk beralih sepenuhnya ke platform LabelMe yang dijalankan pada lingkungan Anaconda di PC pribadi peneliti. LabelMe mendukung berbagai bentuk anotasi yang lebih fleksibel, termasuk *rectangle*, *polygon*, *line*, dan *mask*. Seluruh dataset, file anotasi (.json), dan hasil konversi ke format YOLO (.txt) disimpan di *local storage* PC peneliti, sehingga kontrol akses dan kerahasiaan data terjamin sepenuhnya. Dataset final yang digunakan untuk training model YOLOv8m-seg pada penelitian ini sepenuhnya dianotasi menggunakan LabelMe; project Roboflow yang dibuat pada tahap eksplorasi awal tidak menjadi sumber dataset final. Pengalaman empiris pada tahap awal anotasi rel menggunakan bentuk *line* (garis tipis) menunjukkan bahwa YOLOv8 tidak mampu mengenali pola garis yang terlalu tipis. Berdasarkan temuan ini, dilakukan perubahan strategi anotasi rel dengan mengganti bentuk *line* menjadi *polygon* yang melingkupi area jalur rel secara penuh. Perubahan ini meningkatkan kemampuan model dalam mempelajari pola visual jalur rel karena YOLOv8 secara arsitektural dirancang untuk mengenali bentuk *rectangle* dan *polygon*, bukan garis tipis. Dataset akhir hasil anotasi LabelMe mencakup 300 gambar dengan empat kelas objek. Distribusi instances per kelas ditampilkan pada Gambar 2.



Gambar 2. Distribusi instances per kelas dan posisi bounding box pada dataset training

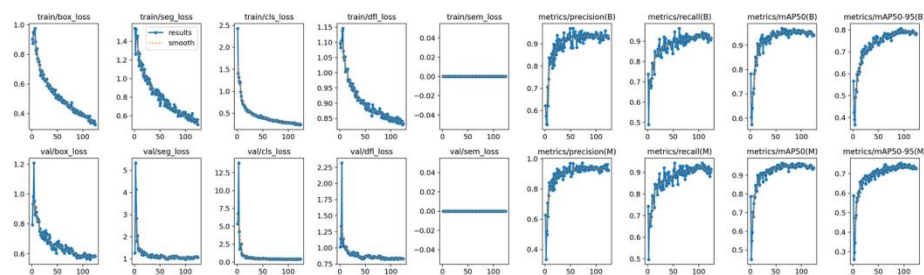
Distribusi instances pada Gambar 2 menunjukkan dataset yang tidak seimbang (imbalanced), dengan Rail sebagai kelas terbanyak (241 instances), diikuti Human (223 instances), Train (114 instances), dan Car (102 instances) — total 680 instances pada 300 gambar. Ketidakseimbangan ini berdampak pada performa per-kelas yang akan dibahas pada Sub-bab 3.2, di mana kelas dengan instances lebih sedikit cenderung memiliki performa relatif lebih rendah [23], meskipun perbedaannya tidak besar berkat representasi yang tetap memadai.

3.2 Hasil Training dan Validasi Model YOLOv8-seg

Evaluasi kinerja model dilakukan menggunakan beberapa metrik standar. *Confusion Matrix* (CM) memberikan informasi rinci mengenai kesalahan klasifikasi melalui elemen *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [24],[25]. Dari CM diperoleh metrik turunan: $Precision = TP/(TP+FP)$ yang mengukur ketepatan deteksi positif, $Recall = TP/(TP+FN)$ yang mengukur kemampuan model menemukan seluruh objek positif, dan $F1-Score = 2 \times (Precision \times Recall) / (Precision + Recall)$ sebagai keseimbangan keduanya.

Pada konteks model deteksi objek, metrik utama yang digunakan adalah *mean Average Precision* (mAP) yang menghitung area di bawah kurva Precision-Recall pada berbagai threshold Intersection over Union (IoU) [25]. $mAP@0.5$ mengukur akurasi pada $IoU \geq 0,5$, sedangkan $mAP@0.5-0.95$ merupakan rata-rata mAP pada range IoU dari 0,5 hingga 0,95 dengan step 0,05, mencerminkan ketepatan lokalisasi yang lebih ketat. Pada konteks keselamatan kereta, *Recall* lebih kritis dibanding *Precision* — kegagalan mendeteksi obstacle (*False Negative*) berdampak fatal pada keselamatan operasional, sedangkan alarm palsu (*False Positive*) hanya berdampak pada kenyamanan operasi.

Proses training model YOLOv8m-seg dilakukan pada Google Colaboratory dengan konfigurasi yang telah dijelaskan pada bab pemrosesan dataset dan training model. Training mencapai kurang lebih 125 epoch sebelum *early stopping* aktif dengan *patience* 30 epoch. Kurva training dan validasi pada berbagai metrik ditampilkan pada Gambar 3.

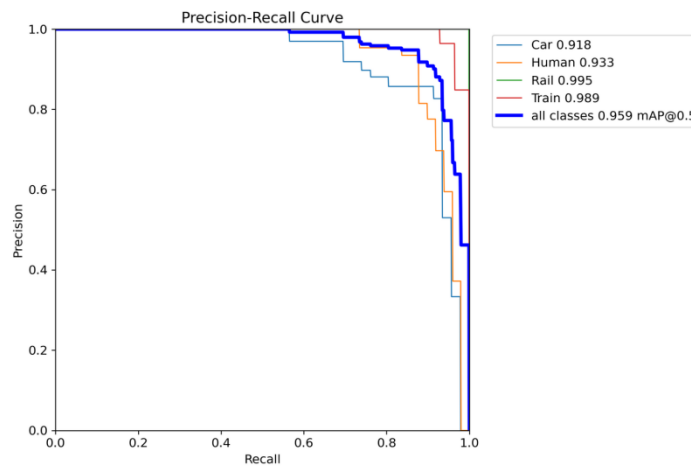


Gambar 3. Kurva loss training/validasi dan metrik Precision, Recall, mAP@0.5, dan mAP@0.5-0.95 selama training

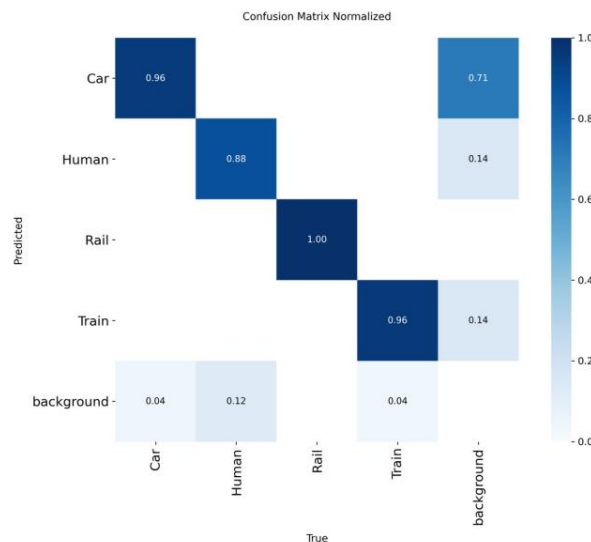
Gambar 3 menunjukkan bahwa nilai loss training dan validasi mengalami penurunan konsisten sepanjang epoch, mengindikasikan model berhasil belajar dari data. Pada sisi metrik, Precision, Recall, mAP@0.5, dan mAP@0.5-0.95 mengalami peningkatan stabil dan mencapai konvergensi setelah epoch ke-60, baik pada level deteksi (B = Box) maupun segmentasi (M = Mask). Tidak terlihat gejala overfitting signifikan karena kurva loss validasi tetap mengikuti tren loss training. Pada akhir training, nilai metrik konvergen pada Precision di kisaran 0,93–0,94, Recall di kisaran 0,92–0,93, mAP@0.5 mendekati 0,96, dan mAP@0.5-0.95 di kisaran 0,78–0,80 untuk level Box.

Evaluasi akhir model pada level Box menghasilkan capaian mAP@0.5 sebesar 0,959 untuk seluruh kelas. Kurva Precision-Recall per kelas ditampilkan pada Gambar 4.

Kurva Precision-Recall pada Gambar 4 menunjukkan kinerja per-kelas dengan capaian mAP@0.5 sebagai berikut: Rail = 0,995; Train = 0,989; Human = 0,933; dan Car = 0,918. Kelas Rail mencapai nilai mendekati sempurna karena objek tersebut memiliki bentuk visual yang konsisten dan menjadi fokus utama dalam dataset. Train juga mencapai nilai sangat tinggi karena pola visual lokomotif yang relatif khas. Human dan Car memiliki nilai relatif lebih rendah namun tetap berada di atas 0,91, mengindikasikan model mampu mengenali kedua kelas dengan baik meski terdapat variasi pose dan ukuran objek.



Gambar 4. Kurva Precision-Recall pada level Bounding Box untuk seluruh kelas



Gambar 5. Confusion Matrix ternormalisasi hasil evaluasi model YOLOv8m-seg

Confusion Matrix ternormalisasi pada Gambar 5 memberikan informasi yang lebih rinci tentang performa per-kelas. Nilai diagonal menunjukkan *True Positive Rate* (Recall) untuk setiap kelas: Rail mencapai TPR sempurna 1,00, Car dan Train masing-masing 0,96, dan Human 0,88. Pada kolom background, terjadi misklasifikasi background sebagai objek target: 71% background terklasifikasi sebagai Car, 14% sebagai Human, dan 14% sebagai Train. Sebaliknya, sebagian objek terklasifikasi sebagai background (*False Negative*): 12% Human, 4% Car, dan 4% Train.

Tingginya *False Positive* kelas Car dari background (0,71) menunjukkan model cenderung mendeteksi objek menyerupai mobil pada area yang sebenarnya hanya background — kemungkinan disebabkan oleh keberadaan kendaraan parkir, bayangan, atau struktur menyerupai mobil di area lintasan. Hal ini menjadi catatan untuk pengembangan dataset selanjutnya, di mana penambahan contoh background tegas tanpa objek target dapat membantu model membedakan dengan lebih baik. Sementara itu, nilai TPR Human sebesar 0,88 sudah cukup baik untuk konteks keselamatan kereta model berhasil mendeteksi sebagian besar manusia di area uji, sejalan dengan prinsip bahwa *Recall* lebih kritis dibanding *Precision* pada sistem keselamatan.

Tabel 2. Target Performa Deteksi Sistem VOODS

Parameter	Target Kinerja
<i>False Alarm Rate</i> (Tingkat Alarm Palsu)	< 0,1%
<i>Miss Detection Rate</i> (Tingkat Kegagalan Deteksi)	< 0,1%
Akurasi Deteksi (kondisi normal)	> 90%
Akurasi Deteksi (kondisi <i>degraded</i> kamera)	> 70%
Jangkauan Deteksi Efektif (objek manusia)	> 60 meter
Jangkauan Deteksi Efektif (objek kendaraan)	> 100 meter

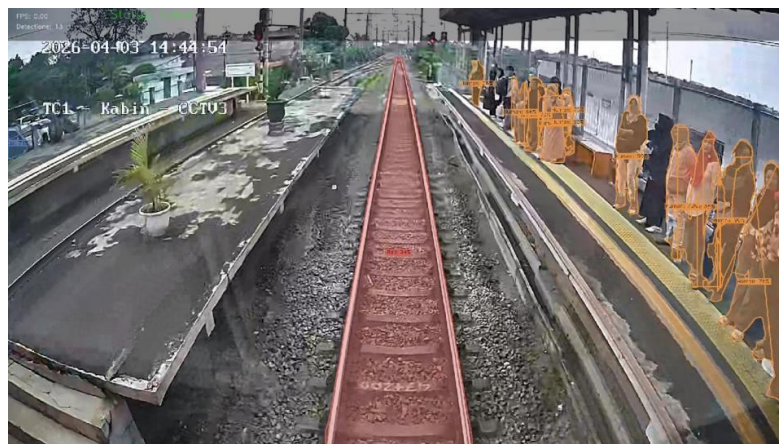
Tabel 3. Hasil Evaluasi Kinerja Model YOLOv8m-seg per Kelas

Kelas Objek	mAP@0.5 (Box)	TPR (Confusion Matrix)	Jumlah Instances
Rail	0,995	1,00	241
Train	0,989	0,96	114
Human	0,933	0,88	223
Car	0,918	0,96	102
All classes (rata-rata)	0,959	—	680

Tabel 3 merangkum capaian per-kelas pada metrik mAP@0.5 (Box) dan TPR (Confusion Matrix), beserta jumlah instances di dataset. Konsistensi tinggi antara mAP dan TPR untuk seluruh kelas (semua di atas 0,88) menunjukkan bahwa model telah belajar dengan baik terlepas dari ketidakseimbangan distribusi instances. Capaian mAP@0.5 keseluruhan sebesar 0,959 sudah melampaui target akurasi deteksi >90% pada kondisi normal yang ditetapkan pada Tabel 2.

3.3 Implementasi *Rail Boundary System* sebagai *Region of Interest* dan Pengujian Warning

Implementasi mekanisme *on-track warning* dilakukan menggunakan program Python berbasis OpenCV yang memuat model best.pt hasil training dan memproses video kamera depan kereta secara *per frame*. Skrip warning mendukung sumber video dari file, folder gambar, kamera USB, atau *Picamera*, dengan parameter konfigurasi mencakup *confidence threshold* (default 0,492) dan *overlap_thresh* minimal jumlah piksel overlap antara mask rel dan mask objek untuk memicu warning (default 10 piksel). Pengujian dilakukan pada video operasional dari kabin masinis KRL Jabodetabek dengan kondisi lintasan lurus. Sistem berhasil menampilkan visualisasi mask rel berwarna merah transparan yang mengikuti kontur jalur rel sesuai hasil segmentasi YOLOv8-seg, sebagaimana ditampilkan pada Gambar 6.



Gambar 6. Visualisasi mask rel (merah) dan deteksi penumpang di area peron (oranye) pada lingkungan stasiun. Status: Clear

Pada Gambar 6 yang diambil di lingkungan stasiun, sistem berhasil membedakan area rel dari peron. Penumpang yang berada di peron terdeteksi sebagai Human dengan mask berwarna oranye, namun tidak memicu warning karena tidak ada overlap dengan mask rel. Status global frame "Clear" terkonfirmasi pada sudut kiri atas.



Gambar 7. Visualisasi pada perlintasan jalan dengan deteksi kendaraan dan pejalan kaki di luar area rel

Pada Gambar 7 yang menampilkan area perlintasan, sistem mendeteksi mobil dan pejalan kaki di luar area rel dengan mask berwarna biru-oranye, namun tidak memicu warning karena objek-objek tersebut tidak overlap dengan mask rel. Hal ini menunjukkan bahwa mekanisme *mask overlap* berfungsi efektif sebagai filter spasial untuk menekan *false positive* di lingkungan kompleks.



Gambar 8. Pengujian skenario penumpang masuk ke area rel — Status: WARNING — Human (96%, 95%) on track



Gambar 9. Pengujian skenario penumpang tunggal di area rel — Status: WARNING — Human (94%) on track

Gambar 8 dan gambar 9 menampilkan capaian utama sistem warning. Ketika satu atau lebih objek Human terdeteksi memiliki overlap mask dengan rail_mask di atas threshold, sistem memunculkan teks "Status: WARNING — Human (*confidence%*) on track!" di bagian atas frame dan mask objek *on track* dirender berwarna merah. Pada Gambar 8, terdapat dua orang yang teridentifikasi *on track* dengan *confidence* 96% dan 95%, sementara dua penumpang lain di sebelah kiri area rel tetap dirender oranye dan tidak memicu warning. Pada Gambar 9, satu objek *on track* dengan *confidence* 94% terdeteksi, sementara objek lain di luar area rel diabaikan oleh logika overlap. Hasil pengujian ini mengonfirmasi bahwa mekanisme on-track warning berbasis mask overlap berfungsi pada skenario operasional nyata: (1) sistem dapat menghasilkan keputusan biner per objek (*on track* / tidak) berdasarkan geometri mask, bukan hanya posisi bounding box; (2) sistem dapat menangani banyak objek bersamaan dalam satu frame dengan klasifikasi independen; (3) keputusan visualisasi (warna merah vs oranye) konsisten dengan keputusan warning global.

Meskipun demikian, terdapat keterbatasan teknis yang perlu disampaikan secara transparan. Pertama, pengujian saat ini baru dilakukan pada lintasan lurus, sehingga performa sistem pada kondisi tikungan atau lintasan dengan radius lengkung tertentu belum teruji secara empiris. Pada kondisi tikungan, mask rel akan bergeser secara horizontal dalam bidang pandang kamera mengikuti kontur lengkung lintasan, dan walaupun ROI berbasis mask segmentasi bersifat dinamis (lebih adaptif dibanding ROI polygon manual statis), perilakunya pada radius tikungan tertentu memerlukan verifikasi tersendiri. Kedua, ROI yang digunakan saat ini baru mencakup area lintasan rel (rail footprint) dan belum mencakup proyeksi ruang bebas penuh sesuai PM No. 60 Tahun 2012 dengan lebar 2,15 meter kiri-kanan as rel dan tinggi 3,55 meter. Ketiga, parameter *overlap_thresh* default 10 piksel masih bersifat empiris dan belum dikalibrasi terhadap jarak deteksi efektif yang dipersyaratkan (>60 meter untuk manusia, >100 meter untuk kendaraan).

3.4 Evaluasi Ketahanan terhadap Kondisi Lingkungan

Evaluasi ketahanan sistem terhadap variasi kondisi lingkungan dilakukan pada tiga aspek utama yang mencerminkan tantangan operasional nyata di lingkungan perkeretaapian: kondisi pencahayaan, kondisi cuaca, dan kondisi getaran mekanis selama kereta berjalan.

3.4.1 Kondisi Pencahayaan

Pengujian dilakukan pada dua kondisi waktu yang tersedia di dataset penelitian: pagi hari dengan intensitas cahaya rendah hingga sedang, dan siang hari dengan intensitas cahaya tinggi. Pada kedua kondisi ini, sistem menunjukkan hasil deteksi yang konsisten sebagaimana terlihat pada visualisasi hasil pengujian video di Sub-bab 3.3. Pengujian pada kondisi sore dan malam hari belum dilakukan pada tahap penelitian ini karena keterbatasan ketersediaan dataset pada kondisi tersebut. Hal ini menjadi rekomendasi penting untuk pengembangan lanjutan, di mana penambahan dataset kondisi sore dan malam hari, serta pertimbangan penggunaan kamera dengan kemampuan low-light yang lebih baik atau integrasi dengan sensor inframerah, dapat memperluas cakupan kondisi operasional sistem.

3.4.2 Kondisi Cuaca

Pengujian sistem pada variasi kondisi cuaca — khususnya hujan, kabut, dan pencahayaan rendah akibat cuaca mendung — belum dilakukan pada tahap penelitian ini. Dataset yang digunakan untuk *training* dan validasi model YOLOv8m-seg terbatas pada kondisi cuaca cerah normal, sehingga performa sistem pada kondisi cuaca ekstrem belum dapat dievaluasi secara empiris. Hal ini menjadi keterbatasan penelitian yang perlu ditindaklanjuti pada pengembangan lanjutan, mencakup: (1) pengumpulan dan anotasi dataset yang merepresentasikan variasi kondisi cuaca di lingkungan operasional KRL Jabodetabek; (2) penerapan teknik augmentasi khusus yang mensimulasikan efek hujan, kabut, dan penurunan kontras citra selama proses *training*; dan (3) pengujian sistem pada rekaman video operasional yang diambil pada kondisi cuaca beragam. Langkah-langkah tersebut penting untuk memastikan *robustness* sistem VOODS terhadap perubahan kondisi lingkungan yang lazim dijumpai pada operasional perkeretaapian di Indonesia.

3.4.3 Kondisi Getaran

Perangkat VOODS yang terinstal pada kereta diuji dalam kondisi kereta berjalan dengan kecepatan rendah hingga kecepatan operasional normal. Hasil uji dinamis menunjukkan bahwa posisi kamera tetap stabil dan sistem mampu melakukan deteksi obstacle dengan akurat meskipun dalam kondisi getaran. Hal ini membuktikan bahwa pemasangan kamera pada kabin masinis cukup stabil untuk mendukung operasi sistem deteksi secara *real-time*, sehingga tidak diperlukan sistem stabilisasi tambahan pada level hardware.

3.5 Analisis Perbandingan dengan Penelitian Terdahulu

Sistem VOODS yang dikembangkan dalam penelitian ini memiliki beberapa keunggulan dibandingkan penelitian sebelumnya, baik dari sisi metodologi maupun kontekstualisasi terhadap lingkungan operasional Indonesia. Berbeda dari Zhao et al. [1] yang berfokus pada akurasi deteksi *real-time* dengan kombinasi *lightweight* CNN dan *improved transformer* tanpa membahas integrasi ke sistem keselamatan kereta secara komprehensif, penelitian ini merancang VOODS secara utuh mencakup integrasi kamera, pemrosesan onboard, mekanisme on-track warning berbasis mask overlap, dan integrasi keluaran ke *Emergency Brake*.

Dibandingkan dengan Li dan Zhang [2] yang menggunakan kombinasi *encoder-decoder* untuk segmentasi jalur, penelitian ini menggunakan YOLOv8m-seg yang menyatukan kemampuan detection dan segmentation dalam satu framework, serta menerapkan konsep ROI berbasis *Train Clearance Envelope* yang merupakan turunan langsung dari regulasi PM No. 60 Tahun 2012.

Selain itu, penelitian ini mengembangkan dataset lokal yang disesuaikan dengan karakteristik lingkungan perkeretaapian Indonesia, khususnya KRL Jabodetabek dengan lebar rel 1067 mm. Penggunaan platform LabelMe (Anaconda) untuk seluruh anotasi dataset memastikan kerahasiaan data karena disimpan pada PC pribadi peneliti — pertimbangan yang relevan untuk pengembangan sistem keselamatan operasional yang menggunakan data rekaman CCTV operator perkeretaapian. Tabel 4 menyajikan ringkasan perbandingan berdasarkan beberapa aspek kunci.

Tabel 4. Perbandingan dengan Penelitian Terdahulu

Aspek	Zhao et al. [1]	Li & Zhang [2]	Penelitian Ini
Model/Metode	Lightweight CNN + Transformer	Encoder-Decoder + DL Detector	YOLOv8m-seg
Platform Pemrosesan	Onboard GPU	Embedded system	Mini PC (VS Code/Google Colab)
Strategi Anotasi	Single platform	Single platform	LabelMe (Anaconda) — data lokal/PC pribadi
ROI/Mekanisme On-Track	Tidak	Danger Zone parsial	Mask overlap dinamis (PM 60/2012)
Dataset Lokal Indonesia	Tidak	Tidak	Ya (KRL Jabodetabek)
Integrasi Emergency Brake	Tidak	Tidak	Ya (DIO output)
Evaluasi Kondisi Ekstrem	Terbatas	Tidak	Ya (lighting, cuaca, getaran)

4. KESIMPULAN

Penelitian ini berhasil mengembangkan Vision-based On Board Obstacle Detection System (VOODS) berbasis kamera menggunakan algoritma YOLOv8 segmentation untuk meningkatkan keselamatan operasional perkeretaapian perkotaan di Indonesia. Sistem yang dikembangkan memanfaatkan infrastruktur kamera CCTV existing pada kereta, jaringan Ethernet, dan mini PC sebagai unit pemrosesan, sehingga dapat diimplementasikan tanpa penambahan sensor mahal seperti LiDAR. Dari hasil penelitian diperoleh beberapa kesimpulan penting: (1) Pemilihan platform LabelMe (Anaconda) untuk anotasi dataset memberikan dua keunggulan: dukungan bentuk anotasi *polygon* yang sesuai untuk kontur jalur rel, dan jaminan kerahasiaan data karena seluruh dataset disimpan pada PC pribadi peneliti; (2) Model YOLOv8m-seg yang dilatih dengan 240 gambar training dari lingkungan operasional KRL Jabodetabek mencapai mAP@0.5 (Box) sebesar 0,959 untuk seluruh kelas, melampaui target akurasi >90% pada kondisi normal; (3) Pada level per-kelas, Rail mencapai mAP@0.5 sebesar 0,995, Train 0,989, Human 0,933, dan Car 0,918, dengan True Positive Rate dari Confusion Matrix sebesar 1,00 untuk Rail, 0,96 untuk Train dan Car, serta 0,88 untuk Human; (4) Mekanisme *on-track warning* berbasis *mask overlap* antara mask rel dan mask objek berhasil dikembangkan dan diuji pada video operasional, terbukti dapat memunculkan peringatan "WARNING — Human on track" ketika objek manusia berada di area rel dan tidak memunculkan alarm palsu ketika objek berada di luar rel. Untuk pengembangan ke depan, disarankan: (1) penambahan dataset training yang lebih bervariasi, terutama untuk kelas Human dan Car yang masih memiliki performa relatif lebih rendah, serta penambahan dataset background tegas tanpa objek target untuk mengurangi tingginya False Positive kelas Car dari background; (2) perluasan dataset ke kondisi sore dan malam hari yang saat ini belum tersedia, sehingga sistem dapat dievaluasi pada rentang operasional waktu yang lebih lengkap; (3) pengujian dan kalibrasi sistem pada kondisi lintasan tikungan/lengkung dengan berbagai radius; (4) pengembangan ROI dari area lintasan rel saat ini menjadi proyeksi Train Clearance Envelope penuh sesuai PM No. 60 Tahun 2012, mencakup lebar 2,15 meter kiri-kanan as rel dan tinggi 3,55 meter; (5) penambahan dataset kondisi cuaca buruk; serta (6) integrasi penuh dengan Train Control and Monitoring System (TCMS) untuk sistem peringatan dan pengereman darurat yang terotomatisasi. Penelitian ini memberikan kontribusi ilmiah signifikan dalam pengembangan sistem keselamatan transportasi perkeretaapian berbasis computer vision dan kecerdasan buatan yang adaptif terhadap kondisi lingkungan operasional Indonesia.

REFERENSI

- [1] Z. Zhao, J. Kang, Z. Sun, T. Ye, and B. Wu, "A real-time and high-accuracy railway obstacle detection method using lightweight CNN and improved transformer," *Measurement*, vol. 238, p. 115380, Jul. 2024, doi: 10.1016/j.measurement.2024.115380.
- [2] D. Li and J. Zhang, "Intelligent road segmentation and obstacle detection for autonomous railway vehicle," *Advances in Mechanical Engineering*, vol. 16, no. 1, p. 16878132231225312, Jul. 2024, doi: 10.1177/16878132231225312.
- [3] X.-Y. Xu, S.-M. Wang, W.-Q. Liu, and Y.-Q. Ni, "Advancements in Obstacle Intrusion Detection Methods for Rail Transit: A Comprehensive Review," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–34, Jul. 2025, doi: 10.1109/TIM.2025.3612624.
- [4] D. He *et al.*, "Improved Mask R-CNN for obstacle detection of rail transit," *Measurement*, vol. 190, p. 110728, Jul. 2022, doi: 10.1016/j.measurement.2022.110728.
- [5] D. He *et al.*, "Urban rail transit obstacle detection based on Improved R-CNN," *Measurement*, vol. 196, p. 111277, 2022, doi: 10.1016/j.measurement.2022.111277.
- [6] D. He, Z. Zou, Y. Chen, B. Liu, X. Yao, and S. Shan, "Obstacle detection of rail transit based on deep learning," *Measurement*, vol. 176, p. 109241, Jul. 2021, doi: 10.1016/j.measurement.2021.109241.
- [7] F. Guo, Y. Qian, Y. Wu, Z. Leng, and H. Yu, "Automatic railroad track components inspection using real-time instance segmentation," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, pp. 362–377, Mar. 2021, doi: 10.1111/mice.12625.
- [8] M. Zhang *et al.*, "LDCD-Net: lightweight dual-channel detection network for obstacles detection on rail transit," *Signal Image Video Process.*, vol. 19, no. 4, p. 281, Jul. 2025, doi: 10.1007/S11760-025-03845-X.
- [9] X. Zhang, M. Zhou, P. Qiu, Y. Huang, and J. Li, "Radar and vision fusion for the real-time obstacle detection and identification," *Industrial Robot: the international journal of robotics research and application*, vol. 46, no. 3, pp. 391–395, Jul. 2019, doi: 10.1108/IR-06-2018-0113.



- [10] Z. Nan *et al.*, “LiDAR-Camera joint obstacle detection algorithm for railway track area,” *Expert Systems with Applications*, vol. 275, p. 127089, Jul. 2025, doi: 10.1016/j.eswa.2025.127089.
- [11] A. Chernov, M. Butakova, A. Guda, and P. Shevchuk, “Development of Intelligent Obstacle Detection System on Railway Tracks for Yard Locomotives Using CNN,” 2020, pp. 33–43. doi: 10.1007/978-3-030-58462-7_3.
- [12] D. Ristic-Durrant, M. A. Haseeb, M. Franke, M. Banic, M. Simonovic, and D. Stamenković, “Artificial Intelligence for Obstacle Detection in Railways: Project SMART and Beyond,” 2020, pp. 44–55. doi: 10.1007/978-3-030-58462-7_4.
- [13] S. W. Ibrahim, “A comprehensive review on intelligent surveillance systems,” *Communications in Science and Technology*, vol. 1, no. 1, Jul. 2016, doi: 10.21924/cst.1.1.2016.7.
- [14] G. Jocher, A. Chaurasia, and J. Qiu, “Ultralytics YOLOv8,” Jul. 15, 2023.
- [15] J. Terven and D. Cordova-Esparza, “A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS,” *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680–1716, Jul. 2023, doi: 10.3390/make5040083.
- [16] J. M. Molina, J. P. Llerena, L. Usero, and M. A. Patricio, “Advances in instance segmentation: Technologies, metrics and applications in computer vision,” *Neurocomputing*, vol. 625, p. 129584, Jul. 2025, doi: 10.1016/j.neucom.2025.129584.
- [17] L. Rey *et al.*, “A Performance Analysis of You Only Look Once Models for Deployment on Constrained Computational Edge Devices in Drone Applications,” *Electronics*, vol. 14, no. 3, p. 638, Jul. 2025, doi: 10.3390/electronics14030638.
- [18] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *Journal of Big Data*, vol. 6, no. 1, p. 60, Jul. 2019, doi: 10.1186/s40537-019-0197-0.
- [19] E. Pandilova, M. Petrov, V. Spasev, I. Dimitrovski, and I. Kitanovski, “Transfer Learning with Yolo for Object Detection in Remote Sensing,” B. Risteska Stojkoska and S. Janeska Sarkanjac, Eds., Springer Nature Switzerland, 2025, pp. 121–135. doi: 10.1007/978-3-031-86162-8_9.
- [20] J. Jonathan and D. Hermanto, “Penentuan Epochs Hasil Model Terbaik: Studi Kasus Algoritma YOLOv8,” *Digital Transformation Technology*, vol. 4, no. 2, pp. 792–798, Jul. 2024, doi: 10.47709/digitech.v4i2.4640.
- [21] L. Prechelt, “Early Stopping — But When?,” in *Neural Networks: Tricks of the Trade: Second Edition*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds., Berlin, Heidelberg: Springer, 2012, pp. 53–67. doi: 10.1007/978-3-642-35289-8_5.
- [22] PUSDATIN, “PM 60 TAHUN 2012,” Jul. 15, 2026.
- [23] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal Loss for Dense Object Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318–327, Jul. 2020, doi: 10.1109/TPAMI.2018.2858826.
- [24] F. Jupiter, E. S. Negara, Y. Kunang, and M. Herdiansyah, “Implementasi Algoritma CNN dan YOLO untuk Mendeteksi Jenis Kendaraan pada Jalan Raya,” *Explore: Jurnal Sistem Informasi dan Telematika*, vol. 14, p. 110, 2023, doi: 10.36448/jsit.v14i2.3259.
- [25] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. Da Silva, “A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit,” *Electronics*, vol. 10, no. 3, p. 279, Jul. 2021, doi: 10.3390/electronics10030279.