

DETEKSI KEPATUHAN ALAT PELINDUNG DIRI DI TERMINAL BAHAN BAKAR MINYAK MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORKS (CNN)

DETECTION OF PERSONAL PROTECTIVE EQUIPMENT COMPLIANCE AT FUEL TERMINALS USING CONVOLUTIONAL NEURAL NETWORKS (CNN)

Seni Patrisillia¹⁾, Gede Putra Kusuma²⁾

^{1, 2)}Master of Computer Science, Bina Nusantara University

Anggrek Campus, Jl. Kebon Jeruk Raya No. 27, Jakarta Barat, 11530 Indonesia

e-mail: seni.patrisillia@binus.ac.id¹⁾, inegara@binus.edu²⁾

ABSTRAK

Ketidakpatuhan penggunaan alat pelindung diri (APD) pada area berisiko tinggi di terminal bahan bakar umumnya masih dicatat melalui pelaporan manual, sehingga rentan terhadap kelalaian, bias pelapor, dan ketidaklengkapan bukti. Penelitian ini bertujuan untuk membandingkan tiga model deteksi objek berbasis Convolutional Neural Network (CNN) yaitu Faster R-CNN, SSD, dan YOLOv9 untuk menentukan model yang paling layak diterapkan pada zona merah terminal bahan bakar, khususnya di zona merah yang berisiko tinggi terjadi ledakan atmosfer. Metode penelitian meliputi pengumpulan 5.000 gambar dari laporan insiden dan platform data publik, diikuti dengan anotasi, pra pemrosesan, dan augmentasi untuk meningkatkan ketahanan model. Faster R-CNN dilatih dengan backbone ResNet50 dan ResNet101, SSD dengan backbone yang sama, sementara YOLOv9 diuji dengan varian YOLOv9-C dan YOLOv9-E. Kinerja model dianalisis secara kuantitatif menggunakan mean average precision (mAP50) dan grafik loss-epoch, serta secara kualitatif melalui visualisasi prediksi dalam kondisi dunia nyata seperti objek kecil, hambatan visual, dan pencahayaan rendah. Hasil menunjukkan bahwa YOLOv9-E mencapai mAP50 tertinggi sebesar 62,30%, diikuti YOLOv9-C 60,40%, SSD-ResNet50 40,04%, SSDResNet101 39,58%, Faster R-CNN-ResNet101 26,18%, dan Faster R-CNN-ResNet50 25,00%. Meskipun demikian, SSD menunjukkan stabilitas generalisasi terbaik dengan rentang mAP50 antar data latih, validasi, dan uji hanya 3,72–4,46 poin serta prediksi visual yang paling konsisten pada kondisi lapangan yang sulit. Dengan mempertimbangkan akurasi, stabilitas, dan kelayakan implementasi pada domain data nyata, SSD direkomendasikan sebagai model dasar untuk pemantauan kepatuhan alat pelindung diri (APD) di terminal bahan bakar, sedangkan YOLOv9-E merupakan kandidat dengan akurasi mentah tertinggi yang masih memerlukan adaptasi domain lebih lanjut.

Kata kunci: alat pelindung diri (APD), CNN, deteksi objek, terminal bahan bakar

ABSTRACT

Personal protective equipment (PPE) non-compliance in high-risk fuel terminal areas is still commonly recorded through manual reporting, making it vulnerable to omission, reporter bias, and incomplete evidence. This study aims to compare three Convolutional Neural Network (CNN) based object detection architectures, namely Faster R-CNN, SSD, and YOLOv9 to determine the most feasible model to be applied to the red zone of fuel terminals, especially in red zones with a high risk of atmospheric explosions. The research method included the collection of 5,000 images from incident reports and public data platforms, followed by annotation, pre-processing, and augmentation to improve the robustness of the model. Faster R-CNN was trained on ResNet50 and ResNet101 backbones, SSDs with the same backbone, while YOLOv9 was tested with YOLOv9-C and YOLOv9-E variants. Model performance was analyzed quantitatively using mean average precision (mAP50) and loss-epoch graphs, as well as qualitatively through visualization of predictions in real-world conditions such as small objects, visual barriers, and low lighting. The results showed that YOLOv9-E reached the highest mAP50 of 62.30%, followed by YOLOv9-C 60.40%, SSD-ResNet50 40.04%, SSDResNet101 39.58%, Faster R-CNN-ResNet101 26.18%, and Faster R-CNN-ResNet50 25.00%. Nonetheless, the SSD exhibits the best generalization stability with an mAP50 range between train, validation, and test data of only 3.72–4.46 points and the most consistent visual prediction in difficult field conditions. Considering the accuracy, stability, and feasibility of implementation in the real data domain, SSDs are recommended as a basic model for personal protective equipment (PPE) compliance monitoring at fuel terminals, while YOLOv9-E is the candidate with the highest raw accuracy that still requires further domain adaptation.

Keywords: personal protective equipment (PPE), CNN, object detection, fuel terminals

I. PENDAHULUAN

Terminal bahan bakar merupakan fasilitas berisiko tinggi karena aktivitas penyimpanan, pemindahan, dan distribusi bahan mudah terbakar berlangsung secara terus menerus. Penilaian risiko kuantitatif pada terminal penyimpanan bahan bakar menunjukkan bahwa skenario kecelakaan terburuk dapat berkembang menjadi kebakaran genangan (*pool fire*), nyala api cepat (*flash fire*), atau ledakan awan uap (*vapour cloud explosion*). Sementara itu, studi keselamatan pada proyek *downstream oil and gas* menunjukkan bahwa kelemahan dalam pengawasan keselamatan serta kepatuhan terhadap prosedur masih menjadi faktor penting penyebab kecelakaan [1] [2]. Oleh karena itu, pengendalian perilaku aman dan pemantauan kepatuhan pada area kerja kritis menjadi kebutuhan operasional yang sangat penting. Meskipun demikian, implementasi pemantauan kepatuhan keselamatan di area tersebut masih dihadapkan pada berbagai tantangan yang signifikan.

Isu kepatuhan penggunaan alat pelindung diri (APD) tidak dapat dilepaskan dari tingginya angka kecelakaan kerja yang masih terjadi, baik pada tingkat global maupun nasional. Data International Labour Organization (ILO) menunjukkan bahwa sekitar 2,93 juta pekerja meninggal dunia setiap tahun akibat faktor terkait pekerjaan, sementara lebih dari 395 juta pekerja mengalami cedera kerja yang tidak berakibat kematian pada periode yang sama [3]. Sejalan dengan kondisi tersebut, menurut data Kementerian Ketenagakerjaan mencatat 462.241 kasus kecelakaan kerja di Indonesia sepanjang Januari–Desember 2024, dengan nilai klaim kecelakaan kerja mencapai sekitar Rp3,4 triliun [4]. Data ini menunjukkan bahwa penguatan pengawasan keselamatan kerja, termasuk kepatuhan penggunaan APD, masih menjadi kebutuhan yang sangat penting.

Pada zona merah terminal bahan bakar, pekerja diwajibkan menggunakan alat pelindung diri (APD) minimal berupa helm keselamatan, rompi visibilitas tinggi, dan sepatu keselamatan. Dalam praktik lapangan, verifikasi kepatuhan masih banyak bergantung pada observasi manusia dan unggahan foto pada sistem pelaporan insiden. Pola ini menimbulkan tiga masalah utama yaitu keterlambatan deteksi, subjektivitas penilaian, dan terbatasnya cakupan pemantauan. Di sektor migas,

faktor psikososial, kelelahan, dan perilaku keselamatan pekerja juga terbukti berhubungan dengan kualitas kepatuhan terhadap prosedur, sehingga keberadaan sistem pemantauan yang objektif menjadi semakin penting [5]. Dengan demikian, dibutuhkan teknologi pemantauan visual otomatis yang mampu memberikan bukti deteksi secara konsisten dan terdokumentasi.

Deteksi objek berbasis visi komputer merupakan salah satu solusi yang paling relevan untuk kebutuhan tersebut. Teknologi ini memungkinkan sistem mengenali keberadaan pekerja dan atribut alat pelindung diri (APD) secara simultan, sekaligus menentukan lokasinya dalam citra. Implementasi industri serupa telah dilaporkan pada sistem monitoring APD real-time, pemeriksaan APD berbasis IoT, hingga pengawasan lingkungan pengeboran lepas pantai dan fasilitas manufaktur [6] [7] [8] [9]. Artinya, pemantauan kepatuhan APD berbasis deteksi objek tidak lagi bersifat konseptual, tetapi telah bergerak ke arah implementasi operasional.

Penelitian yang secara khusus membahas deteksi kepatuhan alat pelindung diri (APD) pada terminal bahan bakar masih sangat terbatas. Sementara, studi yang ada lebih banyak berfokus pada sektor konstruksi dan lingkungan kerja umum. Padahal, terminal bahan bakar memiliki karakteristik risiko yang berbeda karena memadukan potensi kebakaran dan ledakan dengan kebutuhan pengawasan yang lebih ketat pada area berbahaya. Kesenjangan ini penting dikaji, mengingat model yang berkinerja baik pada dataset publik belum tentu menunjukkan performa yang konsisten pada dokumentasi lapangan nyata akibat perbedaan karakter visual dan pergeseran domain data [10] [11] [1] [12].

Berdasarkan celah tersebut, penelitian ini bertujuan membandingkan dan mengevaluasi kinerja Faster R-CNN, SSD, dan YOLOv9 dalam mendeteksi kepatuhan penggunaan APD di zona merah terminal bahan bakar, serta menentukan model yang paling efektif, stabil, akurat, dan layak untuk sistem pemantauan keselamatan otomatis. Tinjauan pustaka di atas mengarahkan metode penelitian pada tiga kriteria evaluasi utama, yaitu akurasi deteksi, stabilitas kinerja antar split data, dan ketahanan visual pada objek kecil, hambatan visual, serta pencahayaan rendah. Dengan demikian, penelitian ini tidak hanya menjawab model mana yang paling akurat, tetapi juga model mana yang paling layak diimplementasikan.

Kontribusi penelitian ini terletak pada empat hal. Pertama, penelitian menggunakan kombinasi data primer terminal bahan bakar dan data sekunder publik sehingga evaluasi dapat menangkap gejala domain shift. Kedua, penelitian membandingkan model two-stage dan one-stage pada konteks zona merah yang jarang dikaji. Ketiga, analisis tidak berhenti pada mAP50, tetapi juga memasukkan *loss epoch*, validasi statistik deskriptif, dan visualisasi prediksi. Keempat, hasil penelitian diarahkan secara eksplisit pada keputusan implementasi keselamatan kerja.

II. STUDI PUSTAKA

Secara metodologis, deteksi objek telah berkembang dari pendekatan berbasis handcrafted features menuju pendekatan deep learning yang lebih adaptif. Model convolutional neural network (CNN) memungkinkan ekstraksi fitur representatif secara otomatis dari data citra, sedangkan augmentasi citra dan strategi regularisasi berperan penting untuk meningkatkan ketahanan model terhadap variasi domain, skala objek, dan pencahayaan [13] [14] [15] [16]. Perkembangan ini relevan bagi lingkungan terminal bahan bakar yang memiliki karakter visual kompleks, seperti pantulan permukaan tangki, objek kecil pada jarak menengah hingga jauh, serta fluktuasi intensitas cahaya.

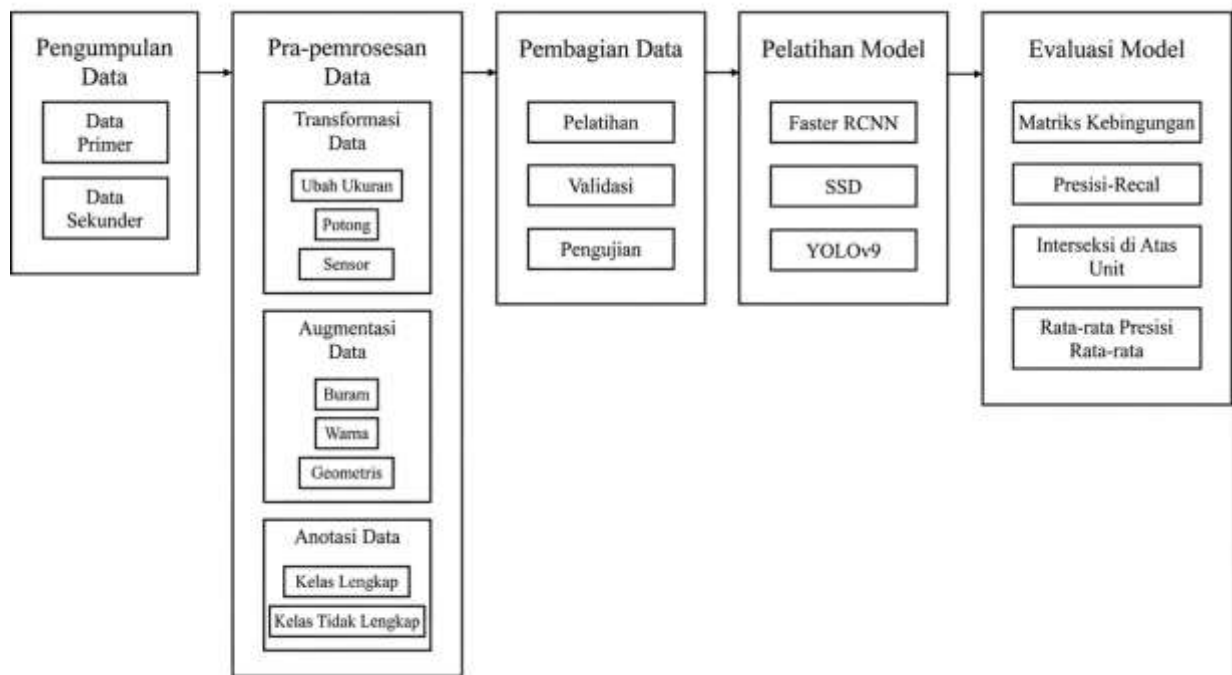
Sejumlah penelitian terkait deteksi APD telah menunjukkan bahwa pendekatan deep learning efektif untuk memantau keselamatan kerja. Nath Behzadan. (2020), Delhi Thomas. (2020), dan Wang Evison. (2021) melaporkan bahwa model berbasis YOLO mampu mendeteksi beberapa jenis APD secara real-time di lokasi konstruksi [10] [17] [18]. Liang dan Seo (2022) menunjukkan akurasi tinggi untuk deteksi helm [19], sedangkan Lee Jung Kang. (2023), Barlybayev Amangeldy. (2024), Guney Altin (2024), Utomo Sasongko (2025), serta López Suarez. (2025) menegaskan bahwa pemantauan APD semakin mengarah pada sistem yang lebih lengkap, cepat, dan siap dioperasikan pada lingkungan nyata [20] [21] [22] [23] [24]. Di sisi lain, hasil antar model tidak selalu seragam. Pada beberapa kasus, SSD tetap menunjukkan kinerja yang kompetitif untuk kebutuhan implementasi praktis maupun pada konfigurasi deteksi objek berukuran kecil, sementara model two-stage masih dapat memberikan hasil yang lebih unggul pada

konteks tertentu, khususnya pada kondisi yang memerlukan presisi lokalisasi dan akurasi deteksi yang lebih tinggi [25] [26] [27] [28].

Dari perspektif model, perbandingan antarmodel tetap penting karena setiap model detektor memiliki karakteristik trade-off yang berbeda dalam aspek akurasi, kecepatan pemrosesan, dan robustnes model. Faster R-CNN umumnya dikenal presisi tetapi lebih sensitif terhadap konfigurasi pelatihan. SSD sering dipilih karena sederhana dan stabil, sedangkan model YOLO berkembang sangat cepat dengan peningkatan fitur *back bone*, *neck*, dan strategi pelatihan [29] [30] [31]. Studi komparatif pada berbagai domain juga menunjukkan bahwa kinerja antar model relatif dapat mengalami perubahan ketika objek yang dideteksi berukuran kecil, distribusi data mengalami pergeseran, atau kondisi visual menjadi lebih kompleks [32] [33] [34].

III. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental untuk membandingkan kinerja tiga model deteksi objek, yaitu Faster R-CNN, SSD, dan YOLOv9, dalam mendeteksi kepatuhan penggunaan alat pelindung diri (APD) pada lingkungan terminal bahan bakar minyak (BBM). Fokus penelitian diarahkan pada kemampuan masing-masing model dalam mengenali objek APD pada dokumentasi visual lapangan yang merepresentasikan kondisi



Gambar 1. Alur Penelitian

operasional nyata, sehingga hasil evaluasi tidak hanya menekankan aspek akurasi, tetapi juga mempertimbangkan kecepatan inferensi dan potensi penerapan praktis pada sistem pemantauan keselamatan kerja. Alur penelitian yang digunakan dalam studi ini disajikan pada Gambar 1. Secara umum, penelitian mencakup lima tahapan utama, yaitu penyusunan dataset, anotasi dan praproses data, pembagian data, pelatihan model, serta evaluasi kinerja model.

Desain penelitian yang digunakan adalah studi komparatif eksperimental. Penelitian diawali dengan pengumpulan citra pekerja di area terminal BBM, kemudian dilakukan proses anotasi terhadap objek APD yang relevan. Dataset yang telah disusun selanjutnya diproses melalui tahapan prapengolahan data, pembagian data latih, validasi, dan uji, lalu digunakan untuk melatih tiga model deteksi objek yang berbeda. Kinerja model dibandingkan berdasarkan metrik evaluasi yang sama agar diperoleh gambaran yang objektif mengenai model yang paling sesuai untuk konteks terminal bahan bakar minyak.

A. Dataset dan Sumber Data

Penelitian ini menggunakan total 5.000 citra. Dataset terdiri atas 1.000 citra primer yang diperoleh dari sistem pelaporan insiden internal dan 4.000 citra sekunder yang diperoleh dari platform data publik, seperti Kaggle dan Roboflow. Data primer merekam personel yang bekerja di area

terminal bahan bakar dengan kondisi penggunaan APD lengkap maupun tidak lengkap. Sebagian citra insiden internal yang telah melalui proses anonimisasi turut diintegrasikan ke dalam data sekunder guna memperkaya keragaman karakteristik visual pada dataset [35] [36]. Penggunaan data primer dibatasi hanya untuk kepentingan penelitian, dan seluruh penanda identitas organisasi pada citra disamarkan melalui proses blur.

Dataset penelitian ini disusun dari citra yang berasal dari perangkat dokumentasi lapangan dan kamera pengawasan dengan spesifikasi yang beragam. Kondisi tersebut menghasilkan variasi alami pada sudut pengambilan gambar, resolusi, jarak objek, dan kualitas pencahayaan. Penelitian ini tidak melakukan standardisasi terhadap seluruh sumber kamera, melainkan mempertahankan heterogenitas tersebut agar dataset yang digunakan lebih mencerminkan kondisi implementasi nyata di lapangan [12].



Gambar 2. Contoh Dataset dari Situs Pelaporan Insiden

B. Praproses, Augmentasi, dan Anotasi Data

Tahap praproses meliputi resize, cropping seperlunya, dan anonimisasi citra. Ukuran masukan akhir disesuaikan dengan konfigurasi masing-masing model, yaitu 300×300 piksel untuk Faster R-CNN dan SSD, serta 480×480 piksel untuk YOLOv9. Dataset disimpan dalam format anotasi deteksi objek yang kompatibel dengan pipeline Roboflow, Torchvision, dan Ultralytics. Setiap entri data memuat identitas citra, koordinat bounding box, dan label kelas. Praktik penyesuaian ukuran citra dan penggunaan format anotasi yang kompatibel dengan pipeline pelatihan merupakan prosedur umum dalam pengembangan sistem deteksi objek modern [14] [37].

Untuk meningkatkan kemampuan generalisasi model terhadap variasi kondisi lapangan, diterapkan teknik augmentasi data berupa perubahan blur, pencahayaan, dan transformasi warna. Strategi ini digunakan untuk meningkatkan kemampuan generalisasi model terhadap variasi domain visual, terutama pada kondisi pencahayaan rendah, perubahan kualitas citra, dan karakteristik lingkungan industri yang dinamis [38].

Anotasi dilakukan pada tingkat individu dengan menetapkan dua kategori, yaitu lengkap dan tidak lengkap. Setiap pekerja yang teridentifikasi dalam citra direpresentasikan melalui *satu bounding box*. Oleh karena satu citra dimungkinkan memuat lebih dari satu pekerja, jumlah total *bounding box* pada dataset dapat melebihi jumlah citra. Kategori lengkap diberikan apabila pekerja teramati menggunakan helm keselamatan, rompi visibilitas tinggi, dan sepatu keselamatan secara bersamaan. Sebaliknya, apabila salah satu atau lebih komponen tersebut tidak digunakan, pekerja diklasifikasikan dalam kategori tidak lengkap. Skema anotasi berbasis kepatuhan APD seperti ini juga digunakan dalam penelitian-

penelitian terbaru mengenai deteksi APD pada lingkungan industri dan konstruksi [35] [36].

C. Pembagian Data

Data dibagi menjadi data pelatihan, data validasi, dan data pengujian dengan rasio 6:2:2. Data pelatihan dan validasi diambil dari data sekunder, sedangkan data pengujian diambil dari data primer terminal bahan bakar. Strategi ini dipilih secara sengaja agar evaluasi akhir dapat mengukur kemampuan generalisasi model pada domain lapangan yang lebih dekat dengan skenario implementasi sebenarnya. Dalam literatur terkini, rancangan evaluasi lintas domain seperti ini dipandang penting untuk menguji ketahanan model terhadap *domain shift* dan untuk menilai kelayakan implementasi nyata di lingkungan kerja yang dinamis [12] [36].

D. Pelatihan Model

Penelitian ini membandingkan tiga model detektor objek, yaitu Faster R-CNN sebagai detektor two-stage, SSD sebagai detektor one-stage, dan YOLOv9 sebagai detektor one-stage generasi mutakhir. Pemilihan ketiga model ini didasarkan pada pertimbangan bahwa ketiganya merepresentasikan *trade-off* yang berbeda antara akurasi, efisiensi komputasi, dan kesiapan implementasi, sebagaimana ditunjukkan dalam kajian mutakhir mengenai evolusi detektor objek berbasis deep learning [14] [37]. Faster R-CNN dan SSD diuji menggunakan backbone ResNet50 dan ResNet101, sedangkan YOLOv9 diuji menggunakan varian C dan E yang mewakili kompromi antara kompleksitas model dan potensi akurasi [35].

Untuk menjamin objektivitas perbandingan antarmodel, konfigurasi hiperparameter disusun secara sebanding agar perbedaan kinerja yang diamati lebih mencerminkan karakteristik model masing-masing model. Ringkasan hiperparameter yang digunakan disajikan pada Tabel 1, meliputi *batch size*, *epoch*, *learning rate*, ukuran citra masukan, *optimizer*, dan *status pre-trained model*. Faster R-CNN dioptimasi menggunakan *stochastic gradient descent* (SGD), sedangkan SSD dan YOLOv9 menggunakan AdamW. Konfigurasi yang digunakan dalam penelitian ini mengacu pada implementasi bawaan yang paling stabil pada masing-masing framework, untuk meminimalkan pengaruh faktor teknis di luar perbedaan arsitektur model [37] [14]. Seluruh model dilatih dalam lingkungan komputasi berbasis GPU dengan

jumlah *epoch*, pembagian data, dan skema augmentasi yang beragam. Oleh karena itu, perbedaan kinerja yang diperoleh dalam eksperimen ini terutama diinterpretasikan sebagai konsekuensi dari karakteristik model masing-masing model.

Tabel 1. Pengaturan Hiperparameter untuk Setiap Model

Hiperparameter	Faster R-CNN	SSD	YOLOv9
Batch size	8	32	20
Epoch	100	100	100
Learning rate	0.01	0.01	0.01
Image size	300	300	480
Optimizer	SGD	AdamW	AdamW
Pre-trained model	Torchvision	Torchvision	True

E. Evaluasi Model

Evaluasi model dalam penelitian ini dilakukan melalui tiga tahapan utama. Tahap pertama difokuskan pada pengukuran kinerja deteksi menggunakan *mean Average Precision* pada ambang *Intersection over Union (IoU)* 0,50 (mAP50), karena metrik ini tidak hanya mencerminkan ketepatan klasifikasi, tetapi juga menunjukkan tingkat kesesuaian lokalisasi *bounding box* terhadap *ground truth* dan merupakan salah satu indikator performa yang paling luas digunakan dalam evaluasi detektor objek modern [14] [37]. Tahap kedua, proses pembelajaran model dianalisis melalui grafik *loss* terhadap *epoch* untuk mengidentifikasi pola konvergensi, kecenderungan *overfitting*, dan kestabilan validasi [37]. Tahap ketiga berupa evaluasi kualitatif yang dilakukan dengan menganalisis keluaran model pada tiga kondisi operasional yang relevan, yaitu keberadaan objek berukuran kecil, objek yang mengalami oklusi, dan kondisi pencahayaan rendah. Pemilihan ketiga kondisi tersebut didasarkan pada temuan mutakhir bahwa *small object detection*, oklusi, dan *low-light scene* merupakan tantangan utama yang dapat menurunkan konsistensi kinerja detektor objek pada lingkungan nyata [39]. Sampel *ground truth* yang digunakan sebagai acuan dalam evaluasi ini disajikan pada Gambar 6, sedangkan hasil prediksi masing-masing model ditampilkan pada Gambar 7, Gambar 8, dan Gambar 9.

Sebagai pelengkap analisis kuantitatif, penelitian ini menerapkan validasi statistik deskriptif terhadap nilai mAP50 pada data latih, validasi, dan uji. Parameter statistik yang dihitung meliputi rata-rata, simpangan baku, koefisien variasi (*coefficient of variation / CV*), dan rentang nilai antarsplit data. Karena setiap model hanya

dilatih satu kali pada partisi data yang identik, analisis statistik dalam penelitian ini bersifat deskriptif-komparatif dan tidak diarahkan untuk menghasilkan inferensi statistik. Pendekatan ini sesuai untuk studi komparatif eksperimental skala terbatas yang berfokus pada pembacaan pola performa dan stabilitas antarkonfigurasi model [37].

F. Kriteria Penentuan Model Terbaik

Penetapan model terbaik dalam penelitian ini dilakukan secara multidimensional, yaitu dengan mempertimbangkan tidak hanya nilai mAP50 tertinggi, tetapi juga kestabilan proses pelatihan, konsistensi performa pada berbagai split data, serta robustnes visual model pada kondisi lapangan yang menantang. Melalui pendekatan ini, pemilihan model diarahkan pada identifikasi model yang paling layak untuk mendukung implementasi sistem pemantauan keselamatan kerja secara nyata di terminal bahan bakar. Kriteria multidimensi seperti ini sejalan dengan rekomendasi kajian-kajian terkini yang menempatkan akurasi, stabilitas, dan kesiapan implementasi sebagai dasar evaluasi model deteksi objek untuk aplikasi industri dan keselamatan kerja [14] [37].

IV. HASIL DAN PEMBAHASAN

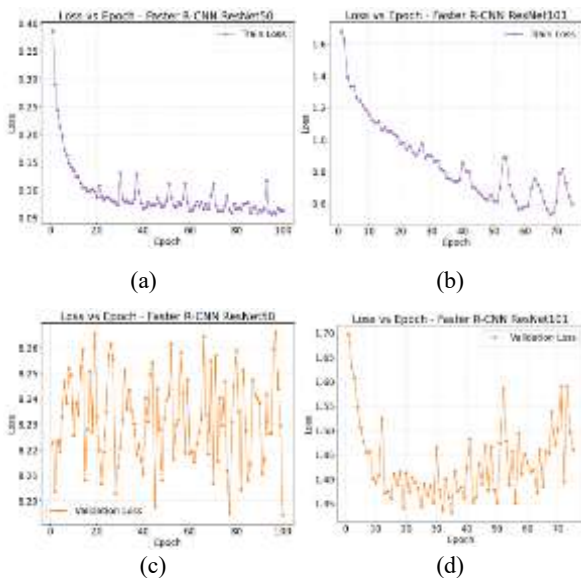
A. Hasil

Dalam konteks zona merah terminal bahan bakar, kecepatan inferensi tetap merupakan aspek yang penting, tetapi akurasi deteksi, stabilitas generalisasi, dan reliabilitas model pada kondisi lapangan yang beragam memiliki signifikansi yang lebih tinggi daripada *throughput* semata. Oleh karena itu, hasil penelitian ini dianalisis tidak hanya berdasarkan nilai mAP50, tetapi juga melalui pola *loss* selama pelatihan serta kualitas prediksi visual pada kondisi operasional yang menantang. Eksperimen dilakukan dengan membandingkan enam konfigurasi model, yaitu Faster R-CNN-ResNet50, Faster R-CNN-ResNet101, SSD-ResNet50, SSD-ResNet101, YOLOv9-C, dan YOLOv9-E. Rekapitulasi *epoch* ketika masing-masing model mencapai train loss terendah dan valid loss terendah disajikan pada Tabel 2.

Tabel 2. Catatan kerugian pelatihan terendah dan kerugian valid dari setiap model

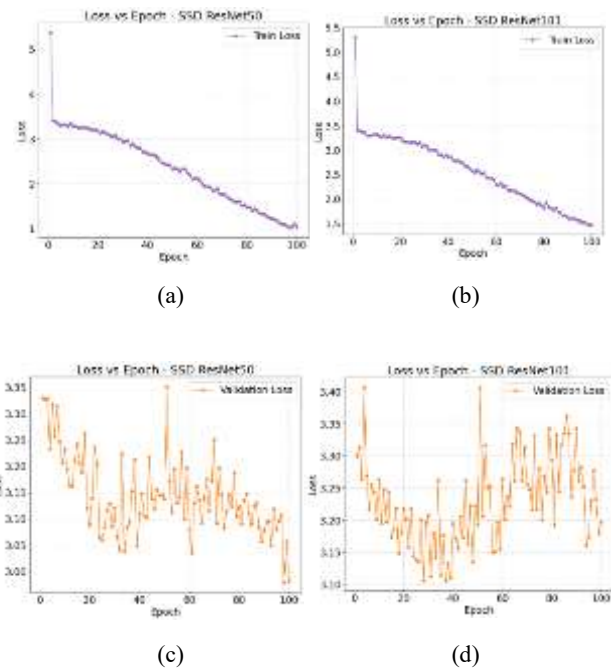
Model	Lowest Train Loss	Train Epoch	Lowest Valid Loss	Valid Epoch
Faster R-CNN ResNet50	0.0576	95	8.2031	100
Faster R-CNN ResNet101	0.5300	68	1.3310	35
SSD ResNet50	1.4326	100	2.8001	98
SSD ResNet101	1.4705	99	3.1601	94
YOLOv9-C	0.6729	100	0.6660	100
YOLOv9-E	0.8332	98	0.8237	99

Secara umum, Tabel 2 menunjukkan bahwa YOLOv9-C dan YOLOv9-E mencapai nilai loss minimum pada fase akhir pelatihan, sedangkan Faster R-CNN, khususnya varian ResNet50, memperlihatkan selisih yang besar antara train loss dan *valid loss*. Pola ini mengindikasikan bahwa kestabilan pembelajaran berbeda antararsitektur dan perlu ditafsirkan bersama dengan kurva loss pada Gambar 3 sampai dengan Gambar 5.

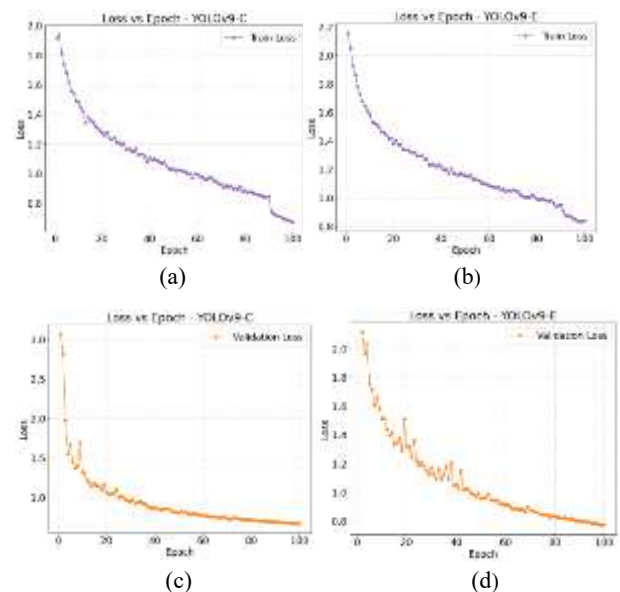


Gambar 3. Grafik kerugian versus epoch untuk Faster R-CNN: (a) Dengan ResNet50 untuk kerugian pelatihan; (b) Dengan ResNet101 untuk kerugian pelatihan; (c) Dengan ResNet50 untuk kerugian validasi; (d) Dengan ResNet101 untuk kerugian validasi.

Sebagaimana terlihat pada Gambar 3, Faster R-CNN dengan backbone ResNet50 menghasilkan train loss yang terus menurun, tetapi valid loss berfluktuasi tajam hingga akhir pelatihan. Varian Faster R-CNN-ResNet101 menunjukkan perilaku yang relatif lebih baik dibandingkan ResNet50, tetapi masih tampak lonjakan *valid loss* setelah sekitar epoch ke-35. Temuan ini menunjukkan adanya kecenderungan *overfitting* dan sensitivitas yang cukup tinggi terhadap variasi data validasi.



Gambar 4. Grafik kerugian versus epoch untuk SSD: (a) Dengan ResNet50 untuk kerugian pelatihan; (b) Dengan ResNet101 untuk kerugian pelatihan; (c) Dengan ResNet50 untuk kerugian validasi; (d) Dengan ResNet101 untuk kerugian validasi.



Gambar 5. Grafik kerugian versus epoch untuk YOLOv9: (a) Dengan YOLOv9-C untuk kerugian pelatihan; (b) Dengan YOLOv9-E untuk kerugian pelatihan; (c) Dengan YOLOv9-C untuk kerugian validasi; (d) Dengan YOLOv9-E untuk kerugian validasi

Berbeda dengan Faster R-CNN, Gambar 4 menunjukkan bahwa kedua varian SSD memiliki tren train loss yang menurun lebih stabil dan valid loss yang relatif lebih terkendali. Meskipun masih terdapat gangguan kecil pada kurva validasi, amplitudo fluktuasinya lebih rendah dibandingkan Faster R-CNN. Dengan demikian, SSD memperlihatkan kompromi yang lebih baik antara

pembelajaran pada data latih dan kemampuan mempertahankan kinerja pada data validasi.

Selanjutnya, Gambar 5 menunjukkan bahwa YOLOv9-C dan YOLOv9-E memperlihatkan pola konvergensi yang paling baik, dengan train loss dan valid loss yang sama-sama menurun hingga akhir pelatihan. Secara internal, pola tersebut mengindikasikan bahwa proses optimisasi berlangsung efisien dan bahwa kedua varian YOLOv9 memiliki kapasitas pembelajaran yang kuat terhadap distribusi data latih maupun data validasi.

Hasil evaluasi kuantitatif pada Tabel 3 memperlihatkan bahwa YOLOv9-E mencapai nilai mAP50 tertinggi pada data uji, yaitu 62,30%, diikuti oleh YOLOv9-C sebesar 60,40%, SSD-ResNet50 sebesar 40,04%, SSD-ResNet101 sebesar 39,58%, Faster R-CNN-ResNet101 sebesar 26,18%, dan Faster R-CNN-ResNet50 sebesar 25,00%. Dari perspektif akurasi mentah pada data uji, kedua varian YOLOv9 jelas menunjukkan keunggulan dibandingkan model lainnya.

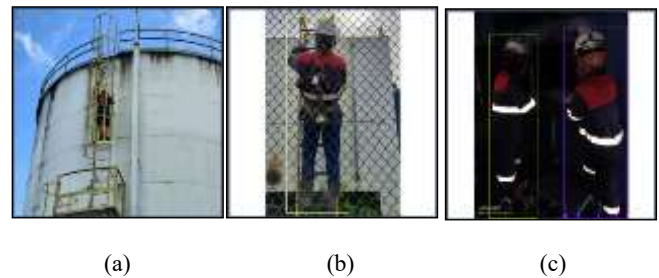
Tabel 3. Perbandingan nilai mAP50 pada dataset pelatihan, validasi, dan pengujian antara model-model

Model	mAP50 Train	mAP50 Valid	mAP50 Test
Faster R-CNN ResNet50	22.00%	18.00%	25.00%
Faster R-CNN ResNet101	24.30%	21.20%	26.18%
SSD ResNet50	43.70%	44.50%	40.04%
SSD ResNet101	41.20%	43.30%	39.58%
YOLOv9-C	79.20%	85.00%	60.40%
YOLOv9-E	75.90%	82.30%	62.30%

Namun, ketika hasil tersebut ditelaah melalui validasi statistik deskriptif antar-split data, pola yang berbeda muncul. SSD-ResNet101 memiliki koefisien variasi terendah, yaitu 4,51%, dan rentang mAP50 terkecil antara data latih, validasi, dan uji, yaitu 3,72 poin. SSD-ResNet50 menempati posisi berikutnya dengan koefisien variasi 5,56% dan rentang 4,46 poin. Sebaliknya, YOLOv9-C dan YOLOv9-E menunjukkan selisih yang lebih besar antara data validasi dan data uji, masing-masing sebesar 24,60 poin dan 20,00 poin, dengan koefisien variasi 17,18% dan 13,90%. Temuan ini menunjukkan bahwa meskipun YOLOv9 unggul dalam akurasi mentah, SSD memiliki stabilitas generalisasi yang lebih baik ketika model berpindah dari domain data sekunder atau public dataset ke data primer terminal bahan bakar.

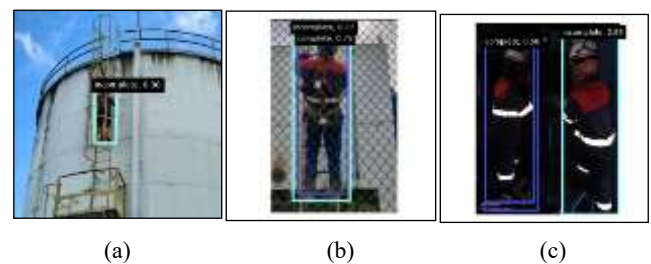
Untuk memastikan bahwa hasil kuantitatif tersebut juga relevan secara operasional, evaluasi kualitatif dilakukan pada tiga kondisi lapangan

yang representatif, yaitu objek berukuran kecil, objek yang mengalami hambatan visual, dan kondisi multiobjek dengan pencahayaan rendah sebagaimana ditunjukkan pada Gambar 6.



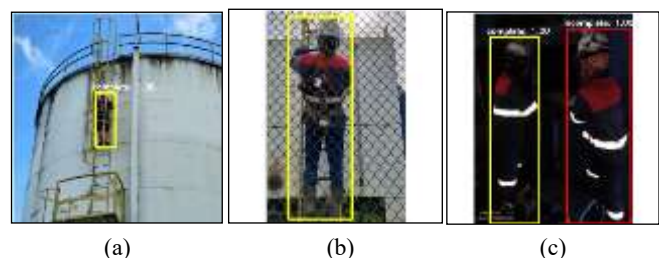
Gambar 6. Data sebenarnya dari kotak batas selama proses anotasi untuk objek dengan berbagai kondisi: (a) Ukuran kecil; (b) Hambatan yang terlihat; (c) Beberapa objek dan cahaya redup.

Hasil prediksi pada Gambar 7 menunjukkan bahwa Faster R-CNN masih menghasilkan prediksi ganda pada kondisi hambatan visual dan cahaya redup.



Gambar 7. Prediksi kotak batas yang dihasilkan oleh Faster R-CNN ResNet101 dengan label prediksi dan skor sebagai berikut: (a) Tidak lengkap 0,90; (b) Lengkap 0,75, Tidak lengkap 0,77; (c) Lengkap 0,58, Tidak lengkap 0,99.

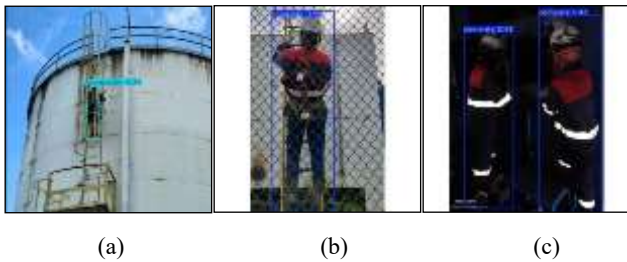
Pada Gambar 8, SSD menghasilkan label yang sesuai pada seluruh sampel dan tetap mempertahankan skor keyakinan yang tinggi.



Gambar 8. Prediksi kotak batas yang dihasilkan oleh SSD ResNet101 dengan label prediksi dan skor sebagai berikut: (a) Lengkap 0,99; (b) Lengkap 0,99; (c) Lengkap 0,99, Tidak Lengkap 0,99.

Sementara itu, Gambar 9 menunjukkan bahwa YOLOv9 mampu mendeteksi kondisi normal dengan baik, tetapi lebih mudah mengalami kesalahan pada objek berukuran kecil dan kondisi pencahayaan rendah. Secara praktis, kombinasi

antara stabilitas antar split dan konsistensi visual inilah yang menjadikan SSD lebih layak sebagai model dasar implementasi, meskipun nilai mAP50 uji YOLOv9 lebih tinggi.



Gambar 9. Prediksi kotak batas yang dihasilkan oleh YOLOv9E dengan label prediksi dan skor sebagai berikut: (a) Tidak lengkap 0,31; (b) Lengkap 0,80; (c) Lengkap 0,43, Lengkap 0,41.

B. Pembahasan

Temuan penelitian ini menunjukkan bahwa pemilihan model untuk lingkungan berisiko tinggi tidak dapat didasarkan semata-mata pada mAP50 data uji. Pada dataset ini, YOLOv9 menghasilkan akurasi mentah tertinggi, tetapi SSD memperlihatkan kestabilan pembelajaran dan stabilitas antar split yang lebih baik. Dalam konteks sistem pengawasan keselamatan, model yang stabil terhadap perubahan domain sering kali lebih bernilai daripada model yang hanya unggul pada satu partisi pengujian.

Perbedaan tersebut dapat dijelaskan oleh rancangan dataset. Data latih dan validasi seluruhnya berasal dari sumber sekunder/public dataset, sedangkan data uji berasal dari data primer terminal bahan bakar. Rancangan ini bermanfaat untuk menguji generalisasi ke kondisi nyata, tetapi sekaligus menciptakan domain shift yang cukup kuat. Gap mAP50 YOLOv9 antara validasi dan uji yang mencapai 20,00–24,60 poin menunjukkan bahwa model ini sangat baik menyesuaikan diri pada distribusi data latih validasi, namun performanya menurun ketika dihadapkan pada karakter visual terminal yang berbeda. Sebaliknya, SSD menunjukkan penurunan yang jauh lebih kecil, sehingga lebih tangguh terhadap bias sumber data.

Analisis bias dataset juga penting pada level anotasi. Pertama, kelas yang digunakan hanya dua, yaitu lengkap dan tidak lengkap. Skema biner ini memudahkan pelatihan, tetapi mengompresi beragam jenis pelanggaran misalnya tanpa helm, tanpa rompi, atau tanpa sepatu ke dalam satu label yang sama. Kedua, citra berasal dari perangkat kamera yang heterogen, dengan variasi sudut

pandang, resolusi, pantulan tangki, dan intensitas cahaya. Ketiga, satu citra dapat memuat lebih dari satu pekerja, sehingga kepadatan objek berubah-ubah. Literatur menunjukkan bahwa *small-object detection*, *low-light scenes*, dan heterogenitas citra dapat memperbesar risiko bias dan kesalahan generalisasi [38] [34] [39].

Hasil ini sejalan dengan studi PPE detection yang menempatkan YOLO sebagai kandidat kuat untuk akurasi dan efisiensi pada lingkungan industri dinamis [35] [21] [22] [24] [23], tetapi juga konsisten dengan laporan bahwa SSD dan model *one stage* tertentu dapat lebih stabil pada deployment praktis atau objek berukuran kecil ketika konfigurasi dan domain datanya sesuai [25] [26] [28]. Di sisi lain, temuan overfitting Faster R-CNN mendukung kajian yang menyatakan bahwa model *two stage* dapat sangat presisi, namun sensitif terhadap ukuran data, konfigurasi regularisasi, dan variasi visual [29] [33].

Kontribusi utama penelitian ini adalah memperluas studi deteksi APD ke konteks terminal bahan bakar, khususnya zona merah yang jarang dilaporkan dalam literatur. Selain membandingkan arsitektur, penelitian ini menunjukkan bahwa keputusan implementasi seharusnya mempertimbangkan tiga aspek secara simultan: akurasi, stabilitas antar split, dan ketahanan visual pada kondisi ekstrem. Dengan kerangka ini, SSD dapat diposisikan sebagai model dasar implementatif, sedangkan YOLOv9-E sebagai kandi-dat akurasi tinggi yang memerlukan adaptasi domain lebih lanjut.

Keterbatasan penelitian ini terletak pada tidak dilakukannya pelatihan berulang dengan seed berbeda, belum dilaporkannya distribusi kelas dan rekap bounding box per sumber data secara terpisah, serta masih digunakannya label biner yang menyederhanakan jenis pelanggaran APD. Penelitian lanjutan disarankan menambah data primer terminal bahan bakar, menerapkan domain adaptation atau synthetic/domain randomization, menguji augmentasi sintetik dan regularisasi lanjutan, memecah label tidak lengkap menjadi subkategori APD, serta mengevaluasi inferensi real-time pada edge device. Arah ini sejalan dengan penelitian terkini tentang dataset PPE, augmentasi, dan implementasi industri untuk sistem monitoring keselamatan [36] [9].

V. KESIMPULAN

Penelitian ini menunjukkan bahwa kinerja model deteksi kepatuhan alat pelindung diri (APD)

di terminal bahan bakar tidak dapat ditentukan hanya berdasarkan nilai akurasi tertinggi. YOLOv9-E memperoleh nilai mAP50 tertinggi pada data uji, tetapi SSD, khususnya SSD-ResNet101, menunjukkan kestabilan pembelajaran, konsistensi performa antarsplit data, dan ketahanan visual yang lebih baik pada kondisi lapangan yang menantang. Temuan ini menunjukkan bahwa SSD lebih layak dipertimbangkan sebagai model dasar untuk implementasi awal sistem pemantauan kepatuhan APD di terminal bahan bakar, sedangkan YOLOv9-E tetap berpotensi untuk dikembangkan lebih lanjut melalui perluasan data primer dan penyesuaian model.

Penelitian ini masih memiliki keterbatasan pada jumlah data primer dan cakupan pengujian, sehingga penelitian selanjutnya perlu melibatkan dataset yang lebih besar, lebih beragam, dan pengujian lintas lokasi agar generalisasi model dapat dievaluasi secara lebih komprehensif..

REFERENSI

Jurnal Online :

- [1] J. Juwari, R. P. Anugraha, M. Zanata, N. A. Leksono, M. D. Hamid, and T. Cahyono, "Quantitative risk assessment and risk reduction of worst-case accident scenario at fuel storage terminal," *J. Loss Prev. Process Ind.*, vol. 88, p. 105272, Apr. 2024, doi: 10.1016/j.jlp.2024.105272.
- [2] A. Waqar, I. Othman, N. Shafiq, and M. S. Mansoor, "Evaluating the critical safety factors causing accidents in downstream oil and gas construction projects in Malaysia," *Ain Shams Eng. J.*, vol. 15, no. 1, p. 102300, Jan. 2024, doi: 10.1016/j.asej.2023.102300.
- [5] A. S. Alroomi and S. Mohamed, "Occupational Stressors and Safety Behaviour among Oil and Gas Workers in Kuwait: The Mediating Role of Mental Health and Fatigue," *Int. J. Environ. Res. Public Health*, vol. 18, no. 21, p. 11700, Nov. 2021, doi: 10.3390/ijerph182111700.
- [6] X. Yang, Y. Yu, S. Shirowzhan, S. Sepasgozar, and H. Li, "Automated PPE-Tool pair check system for construction safety using smart IoT," *J. Build. Eng.*, vol. 32, p. 101721, Nov. 2020, doi: 10.1016/j.jobte.2020.101721.
- [7] F. Gong, X. Ji, W. Gong, X. Yuan, and C. Gong, "Deep Learning Based Protective Equipment Detection on Offshore Drilling Platform," *Symmetry (Basel)*, vol. 13, no. 6, p. 954, May 2021, doi: 10.3390/sym13060954.
- [8] X. Ji, F. Gong, X. Yuan, and N. Wang, "A high-performance framework for personal protective equipment detection on the offshore drilling platform," *Complex Intell. Syst.*, vol. 9, no. 5, pp. 5637–5652, Oct. 2023, doi: 10.1007/s40747-023-01028-0.
- [10] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Autom. Constr.*, vol. 112, p. 103085, Apr. 2020, doi: 10.1016/j.autcon.2020.103085.
- [11] M. I. B. Ahmed *et al.*, "Personal Protective Equipment Detection: A Deep-Learning-Based Sustainable Approach," *Sustainability*, vol. 15, no. 18, p. 13990, Sep. 2023, doi: 10.3390/su151813990.
- [12] H.-S. Kim, J. Seong, and H.-J. Jung, "Optimal domain adaptive object detection with self-training and adversarial-based approach for construction site monitoring," *Autom. Constr.*, vol. 158, p. 105244, Feb. 2024, doi: 10.1016/j.autcon.2023.105244.
- [13] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications," *Multimed. Tools Appl.*, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, doi: 10.1007/s11042-022-13644-y.
- [14] M. L. Ali and Z. Zhang, "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection," *Computers*, vol. 13, no. 12, p. 336, Dec. 2024, doi: 10.3390/computers13120336.
- [15] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, "A Comprehensive Survey of Image Augmentation Techniques for Deep Learning," *Pattern Recognit.*, vol. 137, p. 109347, May 2023, doi: 10.1016/j.patcog.2023.109347.
- [16] C. Santos and J. P. Papa, "Avoiding overfitting: A survey on regularization methods for convolutional neural networks," *ACM Comput. Surv.*, 2022, doi: 10.1145/3510413.
- [17] V. S. K. Delhi, R. Sankarlal, and A. Thomas, "Detection of personal protective equipment (PPE) compliance on construction site using computer vision based deep learning techniques," *Front. Built Environ.*, 2020, doi: 10.3389/fbuil.2020.00136.
- [18] Z. Wang, Y. Wu, L. Yang, A. Thirunavukarasu, C. Evison, and Y. Zhao, "Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches," *Sensors*, vol. 21, no. 10, p. 3478, May 2021, doi: 10.3390/s21103478.
- [19] H. Liang and S. Seo, "Automatic Detection of Construction Workers' Helmet Wear Based on Lightweight Deep Learning," *Appl. Sci.*, vol. 12, no. 20, p. 10369, Oct. 2022, doi: 10.3390/app122010369.

- [20] Y.-R. Lee, S.-H. Jung, K.-S. Kang, H.-C. Ryu, and H.-G. Ryu, "Deep learning-based framework for monitoring wearing personal protective equipment on construction sites," *J. Comput. Des. Eng.*, vol. 10, no. 2, pp. 905–917, Mar. 2023, doi: 10.1093/jcde/qwad019.
- [21] A. Barlybayev *et al.*, "Personal protective equipment detection using YOLOv8 architecture on object detection benchmark datasets: a comparative study," *Cogent Eng.*, vol. 11, no. 1, Dec. 2024, doi: 10.1080/23311916.2024.2333209.
- [22] E. Guney, H. Altin, A. Esra Asci, O. U. Bayilmis, and C. Bayilmis, "YOLO-Based Personal Protective Equipment Monitoring System for Workplace Safety," *JITSI J. Ilm. Teknol. Sist. Inf.*, vol. 5, no. 2, pp. 77–85, Jun. 2024, doi: 10.62527/jitsi.5.2.238.
- [23] Z. R. Utomo, P. W. Adi, P. S. Sasongko, and G. Rahman, "Development and Optimization of a Construction Personal Protective Equipment (PPE) Detection Model on YOLOv8 Architecture," *J. Masy. Inform.*, vol. 16, no. 1, pp. 1–14, May 2025, doi: 10.14710/jmasif.16.1.71622.
- [24] L. López, J. Suárez-Ramírez, M. Alemán-Flores, and N. Monzón, "Automated PPE compliance monitoring in industrial environments using deep learning-based detection and pose estimation," *Autom. Constr.*, vol. 176, p. 106231, Aug. 2025, doi: 10.1016/j.autcon.2025.106231.
- [25] L. Fan, J. Yu, and Z. Hu, "SSD Object Detection Algorithm Based on Feature Fusion and Channel Attention," *Int. J. Adv. Network, Monit. Control.*, vol. 7, no. 3, pp. 80–89, Jan. 2022, doi: 10.2478/ijanmc-2022-0029.
- [27] M. J. B. Brutus, A. L. Fajardo, E. P. Quilloy, L. J. R. Manuel, and A. A. Borja, "Enhancing Seed Germination Test Classification for Pole Sitao (*Vigna unguiculata* (L.) Walp.) Using SSD MobileNet and Faster R-CNN Models," *Appl. Sci.*, vol. 14, no. 13, p. 5572, Jun. 2024, doi: 10.3390/app14135572.
- [28] S. Kim, S. H. Hong, H. Kim, M. Lee, and S. Hwang, "Small object detection (SOD) system for comprehensive construction site safety monitoring," *Autom. Constr.*, vol. 156, p. 105103, Dec. 2023, doi: 10.1016/j.autcon.2023.105103.
- [29] W. Li, "Analysis of Object Detection Performance Based on Faster R-CNN," *J. Phys. Conf. Ser.*, vol. 1827, no. 1, p. 012085, Mar. 2021, doi: 10.1088/1742-6596/1827/1/012085.
- [30] U. Nepal and H. Eslamiat, "Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs," *Sensors*, vol. 22, no. 2, p. 464, Jan. 2022, doi: 10.3390/s22020464.
- [31] R. Sapkota *et al.*, "YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series," *Artif. Intell. Rev.*, vol. 58, no. 9, p. 274, Jun. 2025, doi: 10.1007/s10462-025-11253-3.
- [32] S. M. Alkentar, B. Alsahwa, A. Assalem, and D. Karakolla, "Practical comparison of the accuracy and speed of YOLO, SSD and Faster RCNN for drone detection," *J. Eng.*, vol. 27, no. 8, pp. 19–31, Aug. 2021, doi: 10.31026/j.eng.2021.08.02.
- [33] P. Vilcapoma *et al.*, "Comparison of Faster R-CNN, YOLO, and SSD for Third Molar Angle Detection in Dental Panoramic X-rays," *Sensors*, vol. 24, no. 18, p. 6053, Sep. 2024, doi: 10.3390/s24186053.
- [34] B. Mirzaei, H. Nezamabadi-pour, A. Raoof, and R. Derakhshani, "Small Object Detection and Tracking: A Comprehensive Review," *Sensors*, vol. 23, no. 15, p. 6887, Aug. 2023, doi: 10.3390/s23156887.
- [36] H. M. Ahmad and A. Rahimi, "SH17: A dataset for human safety and personal protective equipment detection in manufacturing industry," *J. Saf. Sci. Resil.*, vol. 6, no. 2, pp. 175–185, Jun. 2025, doi: 10.1016/j.jnlssr.2024.09.002.
- [37] E. Edozie, A. N. Shuaibu, U. K. John, and B. O. Sadiq, "Comprehensive review of recent developments in visual object detection based on deep learning," *Artif. Intell. Rev.*, vol. 58, no. 9, p. 277, Jun. 2025, doi: 10.1007/s10462-025-11284-w.
- [38] S. Bang, F. Baek, S. Park, W. Kim, and H. Kim, "Image augmentation to improve construction resource detection using generative adversarial networks, cut-and-paste, and image transformation techniques," *Autom. Constr.*, vol. 115, p. 103198, Jul. 2020, doi: 10.1016/j.autcon.2020.103198.
- [39] A. Aldubaikhi and S. Patel, "Advancements in Small-Object Detection (2023–2025): Approaches, Datasets, Benchmarks, Applications, and Practical Guidance," *Appl. Sci.*, vol. 15, no. 22, p. 11882, Nov. 2025, doi: 10.3390/app152211882.

Laporan Online :

- [3] International Labor Organization, "Safety and health at work." [Online]. Available: <https://www.ilo.org/topics-and-sectors/safety-and-health-work>
- [4] Satudata Kemnaker, "Kasus Kecelakaan Kerja Tahun 2024." [Online]. Available: <https://satudata.kemnaker.go.id/data/kumpulan-data/2447>

Conference Proceedings yang Terpublikasi:

- [9] P. Torres *et al.*, "A Robust Real-time Component for

Personal Protective Equipment Detection in an Industrial Setting,,” in *Proceedings of the 23rd International Conference on Enterprise Information Systems*, SciTePress, 2021, pp. 693–700. doi: 10.5220/0010452606930700.

- [26] U. Khusni, A. M. Arymurthy, and H. Susanto, “Small Object Detection Based on SSD-ResNeXt101,” in *Proceedings of the 11th International Conference on Robotics, Vision, Signal Processing and Power Applications*, Springer, 2022, pp. 1058–1064. doi: 10.1007/978-981-16-8129-5_162.
- [29] W. Li, “Analysis of Object Detection Performance Based on Faster R-CNN,” *J. Phys. Conf. Ser.*, vol. 1827, no. 1, p. 012085, Mar. 2021, doi: 10.1088/1742-6596/1827/1/012085.
- [35] R. Adidarma, J. Maverick, I. N. Alam, and L. A. Wulandhari, “Personal Protective Equipment Detection for Construction Site Safety Using YOLOv9,” in *2024 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, IEEE, Nov. 2024, pp. 467–471. doi: 10.1109/ICIMCIS63449.2024.10956767.