



# Evaluasi YOLOv8 Nano Untuk Deteksi Logistik Pendaki Pada Clutter Ekstrem

Kevin Risky Abadi, Erna Daniati, Aidina Ristyawan

Fakultas Teknik dan Ilmu Komputer, Sistem Informasi, Universitas Nusantara PGRI Kediri, Kediri, Indonesia

Email: <sup>1</sup>kevinrisky18@gmail.com, <sup>2</sup>ernadaniati@unpkediri.ac.id, <sup>3</sup>aidinaristi@unpkediri.ac.id

Correspondence Author Email: ernadaniati@unpkediri.ac.id

**Abstrak-** Pemeriksaan logistik pendaki gunung secara manual saat ini dinilai tidak efisien dan sangat rentan terhadap kesalahan manusia (*human error*) akibat tingginya volume serta variasi barang bawaan yang sering menumpuk (*clutter*). Meskipun algoritma YOLOv8 sangat populer untuk deteksi objek, kinerjanya pada skenario kepadatan visual ekstrem belum teruji secara komprehensif. Oleh karena itu, penelitian ini bertujuan untuk mengevaluasi ketangguhan model YOLOv8 *Nano* dalam mengidentifikasi lima kelas logistik pendaki pada lima tingkat kepadatan, mulai dari objek tunggal hingga tumpukan ekstrem. Penelitian ini mengadopsi metodologi standar CRISP-ML(Q) dengan memanfaatkan 13.792 sampel data kustom. Fase prapemrosesan menerapkan metode *Stretch to* guna mereduksi artefak visual pada area tepi citra. Hasil eksperimen mendemonstrasikan performa yang sangat presisi, ditandai dengan nilai *Precision* sebesar 0,971, *Recall* 0,954, dan *mean Average Precision* (mAP@50) mencapai 97,8%. Arsitektur ini terbukti sanggup mendobrak limitasi penelitian terdahulu dengan keberhasilan mempertahankan stabilitas mAP@50 di angka 96,22% pada pengujian kepadatan ekstrem (lebih dari 18 objek). Implementasi sistem berbasis aplikasi web lintas perangkat juga mencatatkan waktu inferensi *real-time* yang responsif, yakni 61,48 milidetik pada peramban laptop dan 72,62 milidetik pada telepon seluler. Kesimpulannya, algoritma YOLOv8n terbukti sangat reliabel untuk mengotomatiskan pelaporan logistik lapangan. Namun, limitasi masih ditemukan berupa degradasi akurasi pada objek mikro akibat fenomena kemiripan fitur antar-kelas dan distorsi pantulan cahaya. Studi mendatang direkomendasikan untuk mengintegrasikan teknik *Slicing Aided Hyper Inference* (SAHI) guna memitigasi kegagalan tersebut.

**Kata Kunci:** Deteksi Objek, YOLOv8, Logistik Pendaki, CRISP-ML(Q), Aplikasi Web

**Abstract-** *Manual inspection of climber logistics is inefficient and susceptible to human error due to the high volume and cluttered variation of luggage. Although YOLOv8 is popular for object detection, its performance in extreme visual clutter remains unverified. Therefore, this study evaluates the robustness of the YOLOv8 Nano model in identifying five logistics classes across five density levels, from single objects to extreme clutter. This research adopted the CRISP-ML(Q) methodology utilizing 13,792 custom data samples. The preprocessing phase implemented the Stretch-to method to reduce visual artifacts on image edges. Experimental results demonstrated precise performance with a Precision of 0.971, Recall of 0.954, and mean Average Precision (mAP@50) of 97.8%. The architecture overcame previous limitations by maintaining mAP@50 stability at 96.22% during extreme clutter testing (over 18 objects). Furthermore, the cross-device web application implementation recorded responsive real-time inference times of 61.48 ms on a laptop and 72.62 ms on a smartphone. In conclusion, YOLOv8n is proven highly reliable for automating field logistics reporting. However, accuracy degradation occurred on micro-objects due to inter-class similarity and light reflection distortion. Future studies should integrate Slicing Aided Hyper Inference (SAHI) techniques to mitigate these failures.*

**Keywords:** *Object Detection, YOLOv8, Climber Logistics, CRISP-ML(Q), Web App*

## 1. PENDAHULUAN

Aktivitas penjelajahan alam bebas, khususnya pendakian gunung di Indonesia, menunjukkan tren eskalasi yang masif dalam kurun waktu terakhir, baik sebagai medium rekreasi maupun olahraga prestasi. Implikasi logis dari fenomena ini terlihat jelas pada kepadatan jalur pendakian yang meningkat drastis, terutama selama akhir pekan dan masa libur nasional. Di koridor Jawa Timur, lonjakan kuantitas pendaki pada pos-pos pendakian tertentu dilaporkan dapat menyentuh angka ratusan hingga ribuan individu dalam satu periode waktu [1]. Mobilitas massa yang tinggi ini secara otomatis berkorelasi dengan akumulasi barang bawaan logistik yang dibawa oleh setiap pendaki, mulai dari komoditas pangan instan, air minum kemasan, hingga instrumen pemenuhan nutrisi lainnya. Kesiapan dan kelengkapan logistik tersebut memegang peranan krusial sebagai faktor determinan untuk menjamin keselamatan serta meminimalisir risiko fatalitas di alam liar.

Meskipun urgensi pemeriksaan kelengkapan logistik sangat tinggi, instrumen tata kelola dan verifikasi pada pos perizinan (*basecamp*) faktanya masih bersandar pada mekanisme konvensional berbasis inspeksi visual manual. Pemeriksaan manual ini memicu problem struktural yang signifikan ketika dihadapkan pada karakteristik logistik pendaki yang heterogen. Kompleksitas visual muncul akibat variasi bentuk, dimensi, tekstur, serta kemasan objek yang sangat beragam. Kondisi ini diperparah oleh kecenderungan penataan barang yang menumpuk, saling menutupi (*overlap*), serta berada dalam lingkungan dengan tingkat kepadatan visual yang tidak beraturan (*high clutter scene*). Akibatnya, proses pemeriksaan secara manual menjadi tidak efisien, membutuhkan durasi operasional yang lama, menimbulkan antrian panjang di pos registrasi, serta sangat rentan terhadap fenomena *human error* seperti kegagalan identifikasi (*miss-detection*) maupun kesalahan penghitungan kuantitas objek.

Ekskalasi disrupti teknologi di ranah kecerdasan buatan, khususnya pada klaster *computer vision* dan *deep learning*, mendatangkan pembaruan metodologis guna mengeliminasi hambatan operasional posko melalui otomasi klasifikasi



objek. Melalui rekayasa jaringan saraf konvolusional, sistem komputasi memiliki kapabilitas mengeksplorasi ekstraksi fitur topologi visual sekaligus melakukan penentuan kelas secara simultan dengan presisi tinggi. Dalam konstelasi pelacakan objek secara *real-time*, arsitektur *You Only Look Once* (YOLO) memantapkan posisinya sebagai paradigma dominan. Berdasarkan analisis historis terhadap evolusi algoritma pelokalan objek dalam dua dekade terakhir, YOLO dinilai berhasil memecahkan ambiguitas *trade-off* antara kecepatan pemrosesan data (*inference time*) dan ketajaman spasial batas objek [2]. Kedatangan iterasi mutakhir melalui variasi YOLOv8 membawa perombakan arsitektural yang fundamental pada komponen *backbone*, implementasi struktur *head* bertipe *anchor-free*, serta pembaruan formulasi *loss function*. Kerangka pembaruan ini secara teoretis mengekskalasi sensitivitas jaringan dalam menangkap objek berdimensi mikro (*small objects*) yang kerap memicu kegagalan deteksi pada model-model pendahulu. Karakteristik fungsional tersebut dinilai sangat kompatibel dan adaptif untuk diintegrasikan pada domain penapisan logistik pendaki yang didominasi oleh komoditas bergeometri tipis serta berukuran kecil.

Berdasarkan tinjauan terhadap kelima penelitian terdahulu yang secara spesifik akan diuraikan pada komparasi studi literatur di bagian metodologi, terlihat jelas adanya ruang kosong (*research gap*) dalam literatur ilmiah mengenai evaluasi ketangguhan model YOLOv8n ketika dihadapkan pada tumpukan multi-objek acak berskala masif tanpa pembatasan kuantitas. Penelitian ini hadir untuk mengisi celah tersebut dengan melakukan pengujian batas kemampuan (*stress testing*) model YOLOv8n pada *custom primary dataset* logistik pendaki yang dirancang berdasarkan lima tingkatan kepadatan visual (*clutter levels*) terstruktur, dari skenario objek tunggal hingga kondisi ekstrem di atas 18 objek per citra. Kebaruan riset ini juga terletak pada analisis komparatif performa model terhadap kendala *scale invariance* pada kelas objek mikro, serta pengujian komparasi waktu inferensi secara riil lintas infrastruktur perangkat keras (*laptop vs smartphone*) melalui *deployment* sistem berbasis *Streamlit Web Application*.

Tujuan utama yang ingin dicapai dari pelaksanaan penelitian ini adalah menganalisis secara komprehensif fluktuasi metrik akurasi yang meliputi *precision*, *recall*, *F1-score*, serta *mean Average Precision* (mAP) dari arsitektur YOLOv8n pada dataset logistik pendaki. Secara spesifik, penelitian ini bertujuan mengukur sejauh mana pengaruh variasi densitas objek dan tingkat tumpang tindih visual memicu degradasi performa model, serta membuktikan signifikansi resiliensi model pada kondisi *extreme clutter*. Melalui kontribusi ilmiah ini, diharapkan hasil eksperimen dapat memberikan pemahaman mendalam mengenai batas teoretis YOLOv8n serta menjadi rujukan teknis bagi pengembang sistem cerdas dalam merealisasikan aplikasi pelaporan logistik pendaki yang responsif, inklusif, dan aplikatif di lapangan tanpa memerlukan perangkat *edge AI* yang mahal.

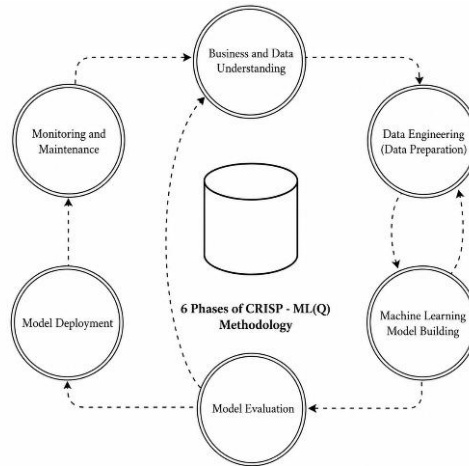
## 2. METODOLOGI PENELITIAN

### 2.1 Studi Literatur

Guna memetakan posisi ilmiah dan merumuskan arsitektur eksperimen secara objektif, penelitian ini bertumpu pada analisis komparatif terhadap lima literatur terdahulu (*state of the art*). Pertama, eksperimen yang diinisiasi oleh Visen dan Charibaldi [3] mengimplementasikan varian YOLOv8 *nano* (YOLOv8n) untuk identifikasi sampah anorganik. Riset tersebut mengidentifikasi batas kritis di mana akurasi klasifikasi merosot tajam ketika jumlah objek melebihi sepuluh unit, sehingga kuantitas uji terpaksa dibatasi guna menghindari degradasi akibat tumpang tindih. Kedua, batasan struktural pada skenario kerumunan dikonfirmasi oleh Hassan dkk. [4] yang membuktikan bahwa objek berskala mikro dan fenomena saling menutupi (*occlusion*) menjadi pemicu utama kegagalan deteksi. Ketiga, Kusumawardhani dkk. [5] membuktikan ketangguhan YOLOv8 dalam mengenali komoditas ritel pada berbagai variasi pencahayaan, namun eksperimen tersebut belum menyentuh skenario kepadatan ekstrem yang acak. Keempat, deteksi spesifik wadah plastik oleh Triyanto dkk. [6] mencatatkan nilai mAP impresif sebesar 99,5%, namun performa tersebut dievaluasi terbatas pada kondisi objek tunggal yang terisolasi (*single object scenario*). Kelima, dari aspek hilirisasi, Putra dan Saputra [7] berhasil mendemonstrasikan kodifikasi YOLOv8 berbasis *web app*, akan tetapi belum mengukur aspek portabilitas kecepatan inferensi pada perangkat keras berspesifikasi rendah (*low-end device*). Kelima literatur inilah yang menjadi landasan utama rumusan masalah dalam penelitian ini.

### 2.2 Kerangka Kerja Eksperimental

Penelitian ini mengadopsi standar tata kelola siklus hidup *machine learning* berbasis CRISP-ML(Q) (*Cross-Industry Standard Process for Machine Learning with Quality Assurance*) [8]. Pemilihan kerangka kerja ini didasarkan pada kapabilitasnya dalam menjamin validitas metodologis sejak fase identifikasi masalah hingga pemeliharaan model di lingkungan produksi. Kerangka kerja terstruktur ini juga jamak diimplementasikan dalam riset komputasi untuk menggaransi keteraturan serta konsistensi alur analisis data secara sistematis [9]. Tahapan awal difokuskan pada pemetaan kompleksitas inspeksi logistik pendaki yang rentan terhadap fenomena *overlapping* dan densitas objek tinggi. Seluruh rancangan eksperimen distrukturisasi untuk memastikan algoritma YOLOv8n tidak sekadar berfungsi sebagai pendeteksi, melainkan diuji batas reliabilitasnya (*stress testing*) melalui skenario kepadatan objek yang bertingkat. Alur linier dari kerangka kerja ini divisualisasikan pada Gambar 1.



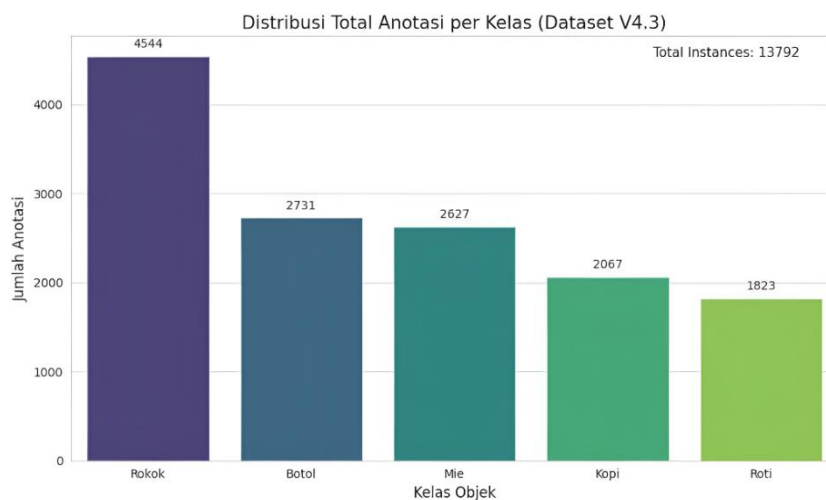
**Gambar 1.** Diagram Metode Penelitian CRISP-ML(Q)

### 2.3 Rekayasa Data dan Konfigurasi Dataset

Kualitas representasi data visual merupakan parameter mutlak dalam arsitektur deteksi objek. Oleh karena itu, penelitian ini menggunakan *custom primary dataset* yang dipublikasikan secara terbuka melalui platform Roboflow Universe [10]. *Dataset* ini merangkum lima *class* logistik pendaki: Botol, Kopi, Mie, Rokok, dan Roti, yang secara keseluruhan diekstraksi ke dalam 13.792 anotasi *bounding box*. Skema ketersediaan ukuran data serta pemerataan distribusi data asli ini menjadi faktor determinan dalam mengukuhkan stabilitas performa serta akurasi model deteksi [11]. Skema *data splitting* diatur secara proporsional menjadi 80% data latih, 10% validasi, dan 10% data uji. Penanganan *class imbalance* diintervensi melalui pendekatan *manual oversampling* pada kelas "Rokok" (mencapai 4.544 anotasi) untuk mengkompensasi dimensi fisiknya yang berukuran mikro dan rentan terhadap anomali *scale invariance* [12].

**Tabel 1.** Distribusi Anotasi Dataset Logistik Pendaki

Kelas Objek	Train	Valid	Test
Botol	2.214	296	221
Kopi	1.653	214	200
Mie	2.111	268	248
Rokok	3.764	435	345
Roti	1.469	195	159
Total	11.211	1.408	1.173



**Gambar 2.** Grafik Distribusi Total Anotasi per Kelas (Dataset V4.3)

Pada tahap prapemrosesan, normalisasi dimensi citra dikunci pada resolusi 640x640 piksel menggunakan algoritma *Stretch to*. Pendekatan *Fit (Reflect Edges)* secara empiris ditolak karena efek cermin yang dihasilkannya memicu distorsi *edge artifacts* dan menduplikasi fitur latar belakang semu, yang secara langsung mengekskalasi *false positive*. Pemilihan strategi prapemrosesan yang tepat ini sangat krusial karena optimasi pada fase awal terbukti memberikan dampak linear yang signifikan pada performa akhir model [13]. Guna menjaga objektivitas uji stres pada kelas kepadatan objek, seluruh konfigurasi augmentasi sintesis eksternal dinonaktifkan. Hal ini didasarkan pada prasyarat pemeliharaan kemurnian fitur visual asli, di mana ekspansi data latih murni dialihkan untuk mengandalkan *Internal Augmentation* bawaan YOLOv8, yakni teknik *Mosaic* dan *Mixup*, guna memperkuat aspek *robustness* model dalam mendiferensiasi objek pada skenario *occlusion* ekstrem [14].

## 2.4 Arsitektur dan Skenario Pelatihan Model

Eksperimen komputasi dieksekusi menggunakan model YOLOv8 *Nano* (YOLOv8n) yang mengonstruksi 225 *layers* dan 3.011.823 *parameters* [3]. Pemilihan varian komputasi ringan ini didasarkan pada efisiensi ekstraksi fitur *real-time* yang dioptimalkan lewat pembaruan arsitektural komprehensif dari generasi-generasi pendahulunya [15]. Dasar matematis dan fungsional dari pemrosesan multilapis pada jaringan saraf dalam ini secara teoretis bersandar pada kaidah generalisasi bobot dan bias dalam komputasi *deep learning* baku [16]. Proses *training* dijalankan di atas ekosistem Kaggle dengan akselerasi perangkat keras GPU NVIDIA T4. Untuk mengoptimalkan pencarian fitur (*feature extraction*) pada fase awal, diimplementasikan teknik *transfer learning* dengan memuat matriks bobot prakonvergensi *yolov8n.pt* [3]. Inisialisasi *hyperparameter* diatur dengan presisi tinggi untuk mencegah *overfitting* sekaligus memastikan ekuilibrium gradien. Pelatihan dilakukan selama 50 *epoch* dengan *batch size* sebesar 16. Optimisasi fungsi kerugian (*loss function*) dikendalikan oleh *optimizer* AdamW, di mana *learning rate* awal ditetapkan pada nilai spesifik 0,001111 untuk menjamin stabilitas konvergensi arsitektur berskala *nano* [3].

## 2.5 Skenario Pengujian Clutter Levels dan Metrik Evaluasi

Berbeda dengan riset konvensional yang membatasi volume deteksi, pengujian dalam penelitian ini diklasifikasikan ke dalam lima kelas kepadatan visual (*Clutter Levels*). Skenario ini dirancang untuk mendeteksi ambang batas degradasi (*performance breakdown*) model saat memproses kepadatan visual ekstrem.

**Table 2.** Klasifikasi Tingkat Kepadatan Objek (*Clutter Levels*)

No	Tingkat Kepadatan	Rentang Jumlah Objek
1	<i>Single</i>	1 Objek
2	<i>Low Multi</i>	2- 5 Objek
3	<i>Medium Clutter</i>	6 – 10 Objek
4	<i>High Dense</i>	11 – 17 Objek
5	<i>Extreme Clutter</i>	≥ 18 Objek

Validasi kinerja deteksi dikuantifikasi menggunakan metrik standar *computer vision* yang berbasis pada irisan matriks prediksi (*True Positive*, *False Positive*, dan *False Negative*). Persamaan matematis yang merepresentasikan tingkat presisi model dan kapabilitas pencarian objek dijabarkan berturut-turut pada rumus (1) hingga (4):

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (4)$$

Pengukuran *mean Average Precision* (mAP) difokuskan pada *threshold* absolut mAP@50 serta kondisi ketat mAP@50-95. Selain parameter akurasi spasial, latensi komputasi juga dianalisis melalui kalkulasi waktu inferensi (*inference time*) dalam rentang milidetik (ms) untuk memvalidasi kelayakan operasional secara riil [17].

## 2.6 Implementasi Sistem dan Penjaminan Mutu (*Quality Assurance*)

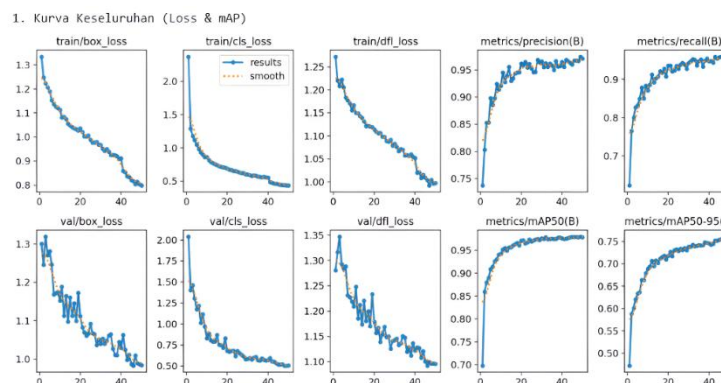
Sebagai *Proof of Concept* (PoC) hilirisasi teknologi, *best weights* yang dihasilkan pada *epoch* ke-50 dikapsulasi ke dalam antarmuka *Web Application* menggunakan *framework* Streamlit. Integrasi arsitektur kecerdasan buatan ke dalam platform berbasis web ini ditujukan untuk mentransformasi model matematis menjadi sistem otomatisasi praktis yang mampu mendukung efisiensi proses komputasi di lapangan [7]. Secara fungsional, purwarupa ini beroperasi sebagai sistem penunjang keputusan taktis untuk mempermudah tata kelola klasifikasi objek melalui pendekatan *machine learning* terapan [18]. Sistem ini dirancang adaptif melintasi platform (*laptop* dan *smartphone*) dengan menyediakan kapabilitas input ganda (unggah berkas dan *live camera*), pengaturan *confidence threshold* dinamis, pemantauan latensi inferensi *real-time*, serta ekstraksi log data berformat CSV untuk otomatisasi manajemen logistik di *basecamp*. Mekanisme pasca-implementasi diatur melalui protokol *Performance Monitoring* sebagai instrumen penjaminan mutu (*Quality Assurance*) untuk menjaga stabilitas kinerja model dari dinamika data lapangan yang fluktuatif [19]. *Baseline* minimum diinstruksikan pada mAP50 sebesar 87%. Apabila terjadi fluktuasi akurasi lapangan yang memicu penurunan performa melebihi 10% dari tolok ukur awal, sistem akan mengidentifikasi fenomena tersebut sebagai *model drift* [20]. Mitigasi teknis yang disiapkan mencakup skema pemeliharaan adaptif berupa pelatihan ulang (*retraining*) menggunakan tambahan *dataset* lapangan baru, serta penerapan teknik *quantization* untuk menekan beban memori tanpa mengorbankan waktu inferensi batas aman 100 ms pada perangkat berspesifikasi rendah [19].

## 3. HASIL DAN PEMBAHASAN

Bagian ini menguraikan secara komprehensif temuan empiris dari eksperimen pelabelan dan deteksi multi-objek logistik pendaki berbasis YOLOv8 *Nano*. Analisis dibagi menjadi evaluasi metrik kinerja model pada tahap pelatihan, pengujian batas kemampuan (*stress testing*) terhadap variasi tingkat kepadatan visual, hingga validasi waktu inferensi pada implementasi arsitektur perangkat keras yang berbeda.

### 3.1 Evaluasi Kinerja Pelatihan dan Penentuan *Confidence Threshold*

Proses komputasi model dieksekusi di dalam ekosistem Kaggle dengan dukungan akselerator grafis NVIDIA Tesla T4. Selama fase pelatihan yang berlangsung sebanyak 50 *epoch* dengan resolusi spasial 640x640 piksel, arsitektur jaringan menunjukkan grafik konvergensi yang sangat stabil. Tidak ditemukan indikasi anomali seperti *overfitting* maupun *underfitting*. Hal ini dibuktikan melalui penurunan kurva *loss function* (meliputi *box loss*, *classification loss*, dan *dfl loss*) yang terjadi secara konsisten dan linear pada data latih maupun data validasi, sebagaimana divisualisasikan pada Gambar 3 berikut.

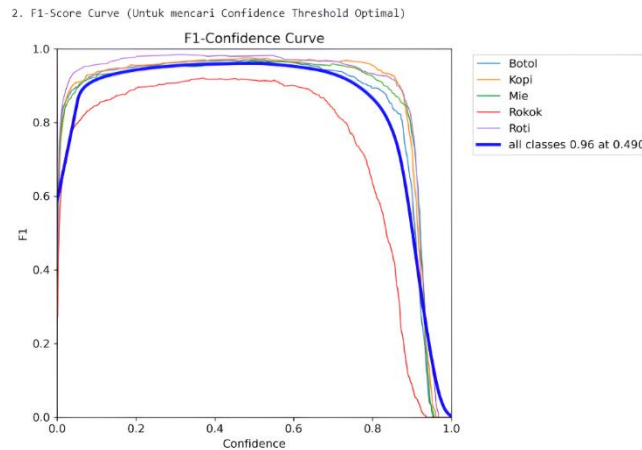


**Gambar 3.** Kurva Keseluruhan *Loss* dan mAP

Berdasarkan ekstraksi data pada iterasi terakhir, model klasifikasi ini membuktikan kapabilitas ekstraksi fitur yang sangat presisi dengan capaian *mean Average Precision* (mAP@50) di angka 97,8% dan mAP@50-95 sebesar 76,1%. Sementara itu, metrik *Precision* menyentuh angka 97,1% yang diiringi oleh tingkat *Recall* sebesar 95,4%. Jika dikomparasikan secara langsung dengan penelitian *state-of-the-art* sebelumnya yang dilakukan oleh Visen dan Charibaldi pada kelas objek anorganik, temuan ini menunjukkan eskalasi performa yang sangat massif [3]. Pada riset acuan tersebut, varian YOLOv8 *Nano* hanya sanggup memproduksi mAP@50 maksimal sebesar 87,1% dan mAP@50-95 sebesar 72,1%. Selain itu, tingkat presisi pada riset acuan hanya mencapai 86,2% dan *recall* sebesar 79,1% pada iterasi ke-50. Lonjakan akurasi

pada penelitian ini sangat dipengaruhi oleh optimalisasi kuantitas *dataset* primer yang berjumlah masif (total 13.792 *instances*) dan skema anotasi yang ketat [3].

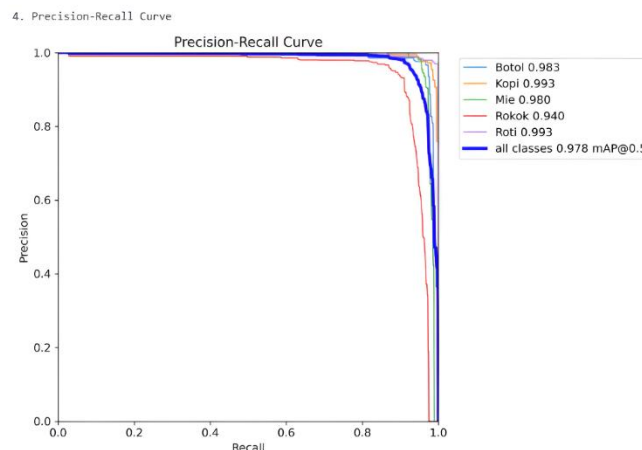
Guna menjamin keseimbangan antara sensitivitas model dalam menemukan objek (*recall*) dan ketepatan prediksi (*precision*), dilakukan kalibrasi ambang batas melalui analisis kurva *F1-Confidence* yang ditunjukkan secara detail pada Gambar 4.



**Gambar 4.** *F1-Confidence Curve*

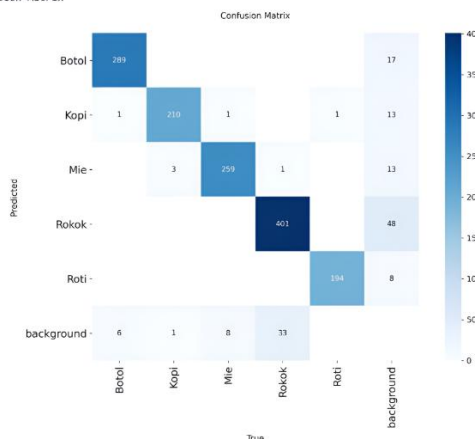
Plot kurva F1 mendemonstrasikan bahwa titik ekuilibrium (nilai harmonik maksimal) sebesar 0,96 tercapai pada rentang *confidence threshold* 0,490 untuk keseluruhan kelas. Angka spesifik 0,49 ini kemudian dikunci sebagai parameter definitif pada sistem *deployment* aplikasi. Pemilihan ambang batas ini berfungsi sebagai filter rasional untuk menekan kemunculan *false positive* (objek latar belakang yang dikira barang logistik) tanpa harus mengorbankan objek berdimensi mikro yang memiliki bobot piksel rendah.

Analisis partisi kelas melalui *Confusion Matrix* mengonfirmasi bahwa entitas "Botol", "Kopi", "Mie", dan "Roti" mendominasi puncak akurasi dengan nilai mAP50 secara berurutan sebesar 98,3%, 99,3%, 98,0%, dan 99,3%. Di sisi lain, objek "Rokok" merepresentasikan kelas dengan tantangan visual tertinggi. Meskipun mAP50 pada kelas ini berhasil didongkrak hingga 94,0%, penurunan nilai mAP50-95 menjadi 60,1% merefleksikan kesulitan intrinsik algoritma dalam mempertahankan lokalisasi kotak pembatas (*bounding box*) pada objek silinder berskala mikro, di mana matriks mencatat terjadinya 33 kasus *False Negative* dan 48 kasus *False Positive* [21]. Distribusi keseimbangan presisi dan sensitivitas ini dapat diamati pada kurva Precision-Recall di Gambar 5, sementara rincian prediksi antar-kelas dijabarkan melalui *Confusion Matrix* pada Gambar 6.



**Gambar 5.** *Precision-Recall Curve*

3. Confusion Matrix



**Gambar 6.** Confusion Matrix

### 3.2 Validasi Deteksi Berdasarkan Tingkat Kepadatan (*Clutter Levels*)

Karakteristik unik dari pengujian operasional ini terletak pada implementasi skenario *stress testing* yang secara sengaja mendobrak batasan deteksi objek konvensional. Data pengujian disegmentasi ke dalam lima kelas kepadatan (*Clutter Levels*) untuk mengamati fluktuasi akurasi saat model merespons tumpukan visual logistik. Hasil evaluasi kuantitatif untuk setiap tingkatan kepadatan tersebut disajikan secara rinci pada Tabel 3.

**Tabel 3.** Hasil Evaluasi Per Tingkat Kepadatan Logistik

No	Tingkat Kepadatan	Jumlah Gambar	mAP 50	Map 50-95	Rata-rata Waktu Inferensi (ms)
1	<i>Single</i>	80	95.94	83.33	6.29
2	<i>Low Multi</i>	26	77.06	58.17	6.83
3	<i>Medium Clutter</i>	17	99.10	77.63	8.26
4	<i>High Dense</i>	30	98.33	76.80	5.89
5	<i>Extreme Clutter</i>	22	96.22	73.50	6.51

Data pada Tabel 3 memperlihatkan pola resiliensi jaringan saraf yang luar biasa. Pada fase densitas puncak (*High Dense* dan *Extreme Clutter*), arsitektur model tidak mengalami kelumpuhan deteksi, dengan nilai mAP50 yang konsisten bertahan di margin 98,33% dan 96,22%. Arsitektur ini sukses memetakan lebih dari ≥18 objek yang saling tumpang tindih secara simultan dalam satu bingkai citra. Pencapaian ini secara empiris membantah postulat dari penelitian sebelumnya yang menyatakan bahwa kemampuan deteksi akan mengalami penurunan akurasi klasifikasi secara drastis ketika jumlah objek dalam sebuah *frame* melebihi batas sepuluh unit [3].

Terkait fluktuasi penurunan nilai mAP50 pada kategori *Low Multi* (77,06%), fenomena ini teridentifikasi sebagai anomali statistik yang dipicu oleh keterbatasan distribusi sampel acak pada set data uji kategori tersebut, dan bukan merepresentasikan kecacatan arsitektur dalam mengenali objek dalam jumlah kecil.

### 3.3 Komparasi Latensi Inferensi Lintas Infrastruktur

Analisis kelayakan operasional dinilai melalui pengukuran *Inference Time* lintas arsitektur perangkat keras. Uji coba dilakukan pada tiga ekosistem komputasi: ekosistem awan (GPU NVIDIA Tesla T4), ekosistem *desktop* (*browser* Laptop Infinix Inbook X1 - CPU x86-64), dan ekosistem *mobile* (Apple iPhone 13 - A15 Bionic), Rincian komparasi waktu inferensi dari ketiga infrastruktur perangkat keras tersebut diuraikan pada Tabel 4.

**Tabel 4.** Perbandingan Waktu Inferensi Lintas Perangkat

Perangkat	<i>Single</i>	<i>Low Multi</i>	<i>Medium Clutter</i>	<i>High Dense</i>	<i>Extreme Clutter</i>	<i>Rata-rata Akhir</i>
GPU NVIDIA Tesla T4	6.29	6.83	8.26	5.89	6.51	6,76 ms
Laptop Infinix Inbook X1	59,52	88,03	60,13	61,79	56,92	61,48 ms



Iphone 13 (A15 Bionic)	76,83	73,32	85,24	77,97	59,40	72,62 ms
------------------------	-------	-------	-------	-------	-------	----------

Berdasarkan komparasi di atas, akselerator GPU T4 mendominasi kecepatan absolut di angka 6,76 ms. Namun, temuan krusial dari pengujian ini justru terletak pada ketangguhan model saat dieksekusi secara lokal (*on-device*). Meskipun tidak memiliki intervensi GPU eksternal, latensi pada laptop (61,48 ms) dan *smartphone* (72,62 ms) terbukti sangat stabil dan berada di bawah parameter batas maksimal interaksi *real-time* manusia (100 ms). Hal ini membuktikan bahwa enkapsulasi YOLOv8 Nano ke dalam Web App memiliki portabilitas lintas platform yang tinggi tanpa mengorbankan utilitas operasional di lapangan.

### 3.4 Visualisasi Kualitatif Deteksi Spasial

Untuk memvalidasi interpretasi data kuantitatif, inspeksi kualitatif dilakukan terhadap citra acak dari luar data pelatihan yang merepresentasikan kelima level kepadatan visual. Evaluasi visual ini mengadaptasi metodologi observasi objek berjenjang untuk membuktikan kemampuan pelacakan algoritma dalam merekonstruksi letak dan kelas objek logistik secara aktual. Ringkasan visualisasi hasil deteksi beserta tingkat akurasinya dipaparkan pada Tabel 5.

**Tabel 5.** Visualisasi Hasil Deteksi Berdasarkan Tingkat Kepadatan

No	Tingkat Kepadatan	Hasil Deteksi (Visual)	Jumlah Prediksi & Hasil	Aktual	Kesimpulan
1	Single		1 objek (Benar)	1 Objek	Berhasil karena dapat mendeteksi 1 objek <i>class</i> Roti
2	Low Multi		3 Objek (Benar)	3 Objek	Berhasil karena dapat mendeteksi 3 objek <i>class</i> Botol, Kopi dan Roti

3	<i>Medium Clutter</i>		8 Objek (Benar)	8 Objek	Berhasil karena dapat mendeteksi 5 objek <i>class</i> Botol (2), Mie (2), Rokok (2), Kopi dan Roti
4	<i>High Dense</i>		13 Objek (Benar)	13 Objek	Berhasil karena dapat mendeteksi 5 objek <i>class</i> Botol (3), Mie (3), Rokok (3), Kopi (2) dan Roti (2)
5	<i>Extreme Clutter</i>		18 Objek (Benar)	18 Objek	Berhasil karena dapat mendeteksi 5 objek <i>class</i> Rokok (5), Mie (4), Kopi (3), Roti (3) dan Botol (3)

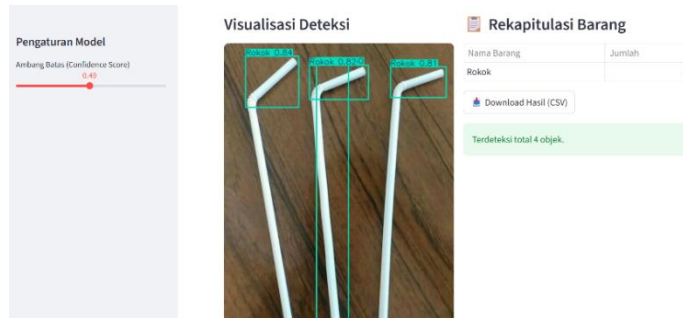
### 3.5 Analisis Akurasi dan Optimalisasi Ekstraksi Fitur Multi-Objek

Pencapaian metrik akurasi  $mAP@50$  absolut sebesar 97,8% dalam eksperimen ini membuktikan bahwa keterbatasan arsitektur model berskala *nano* dapat dimitigasi secara efektif melalui suplai data latih yang komprehensif. Keberhasilan model YOLOv8 *Nano* dalam melokalisasi lebih dari  $\geq 18$  objek bertumpuk secara simultan sangat didorong oleh volume masif *dataset* kustom yang dibangun, yakni mencapai 11.211 *instances* pelatihan dari total keseluruhan 13.792 *instances*. Volume data ini memberikan divergensi fitur spasial yang sangat kaya bagi jaringan saraf untuk mempelajari berbagai probabilitas tumpang tindih (*occlusion*).

Sebagai komparasi kritis, penelitian terdahulu oleh Visen dan Charibaldi yang mengklaim adanya degradasi performa pada deteksi di atas sepuluh objek hanya didukung oleh 1.967 sampel data pelatihan [3]. Hal ini secara logis menjelaskan mengapa model pada riset acuan tersebut mengalami *underfitting* visual ketika dihadapkan pada skenario kerumunan objek. Ketersediaan sampel anotasi primer yang masif dalam penelitian ini membuktikan bahwa varian YOLOv8 dengan parameter terkecil sekalipun memiliki kapasitas yang tangguh untuk memproses skenario kepadatan ekstrem (*Extreme Clutter*), selama jaringan tersebut dilatih dengan varian posisi spasial dan sudut pandang multiobjek yang memadai.

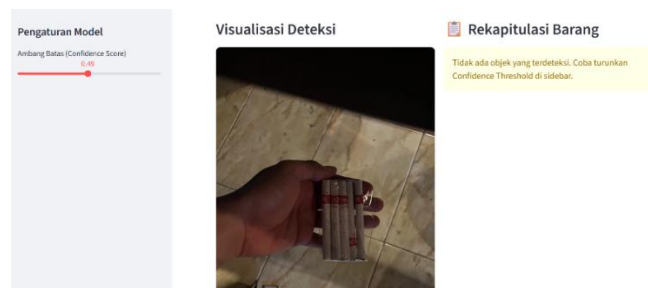
### 3.6 Analisis Anomali Klasifikasi dan Fenomena *Inter-Class Similarity*

Walaupun performa umum menunjukkan tren positif, evaluasi mendalam menemukan anomali spesifik pada kelas logistik berdimensi mikro. Kelas objek "Rokok" teridentifikasi sebagai entitas dengan kompleksitas visual tertinggi, yang direpresentasikan oleh nilai  $mAP@50-95$  sebesar 60,1%. Investigasi kualitatif terhadap kegagalan inferensi (*miss-detection*) mengungkap terjadinya fenomena *Inter-Class Similarity* yang kuat. Arsitektur jaringan tercatat mengklasifikasikan objek non-logistik, yakni sedotan putih, sebagai objek "Rokok" dengan skor probabilitas (*confidence score*) mencapai 0,8, seperti yang didemonstrasikan pada Gambar 7.

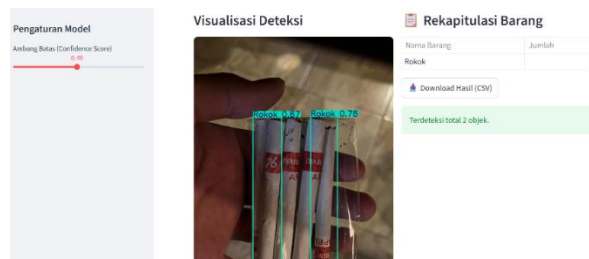


**Gambar 7.** Kasus *False Positive*: Sedotan Putih Terdeteksi Sebagai Rokok

Kegagalan pemisahan batas kelas (*decision boundary*) ini dilatarbelakangi oleh kemiripan fitur geometri intrinsik antara kedua objek tersebut, yang sama-sama memiliki rasio silinder memanjang (*cylindrical aspect ratio*) dan dominasi distribusi piksel berwarna putih. Selain itu, hambatan deteksi yang krusial juga ditemukan pada objek yang terbungkus oleh material plastik transparan. Efek *specular reflection* (pantulan spekular cahaya) pada permukaan plastik mendistorsi ekstraksi gradien tekstur objek asli. Ketika citra diambil dari jarak pandang yang jauh, efek pantulan ini bertindak sebagai *noise* dominan yang mengaburkan fitur lokal, sehingga memicu *False Negative*. Degradasi deteksi akibat pantulan cahaya dari jarak jauh ini divisualisasikan pada Gambar 8, sedangkan keberhasilan pemulihan resolusi ekstraksi fitur melalui penyesuaian jarak dekat ditunjukkan pada Gambar 9.



**Gambar 8.** Efek *Specular Reflection* Memicu Kegagalan Deteksi Jarak Jauh



**Gambar 9.** Resolusi Ekstraksi Fitur Berhasil pada Jarak Dekat

Kendala distorsi pantulan ini terbukti dapat diselesaikan secara fungsional melalui penyesuaian jarak operasional akuisisi citra. Pengambilan gambar dari jarak dekat (*close-up*) mampu mereduksi dominasi *noise* spekular dan mengembalikan ketajaman *edge detection* (pendeteksian tepi) pada objek di balik plastik transparan.

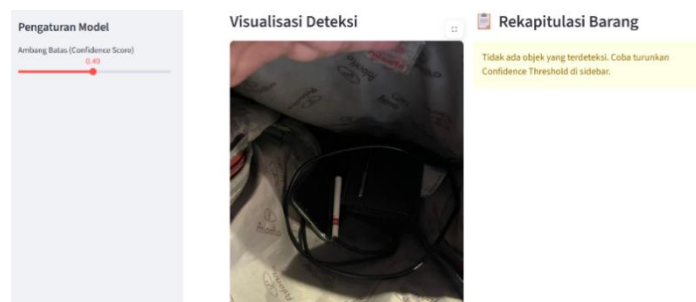
### 3.7 Evaluasi Kompatibilitas Infrastruktur Lintas Perangkat Keras

Pengujian *deployment* sistem berbasis *Web Application* memvalidasi tingkat portabilitas arsitektur YOLOv8 *Nano* pada ekosistem komputasi klien (*client-side execution*). Data eksperimental menunjukkan latensi inferensi rata-rata sebesar 61,48 ms pada peramban laptop berbasis arsitektur CPU x86-64 (Infinix Inbook X1) dan 72,62 ms pada peramban seluler berbasis arsitektur ARM (iPhone 13 - A15 Bionic).

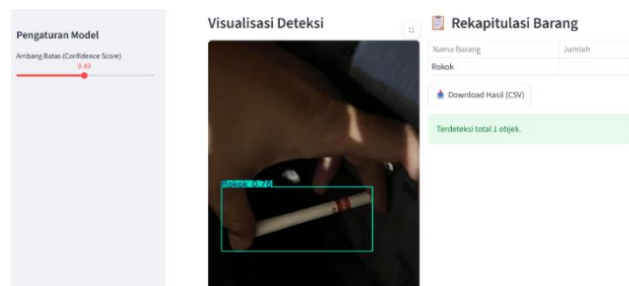
Perbedaan tipis latensi ini menyoroti fenomena utilisasi perangkat keras tingkat peramban. Lingkungan *browser* pada ekosistem PC memiliki keuntungan dari integrasi instruksi vektor tingkat lanjut (*Advanced Vector Extensions / AVX2*) yang mempercepat operasi perkalian matriks secara substansial. Sebaliknya, ekosistem *browser* iOS memiliki restriksi keamanan dari sistem operasi yang membatasi akses optimalisasi penuh terhadap akselerator spesifik seperti *Neural Engine*, sehingga bergantung pada komputasi CPU standar. Meskipun demikian, capaian latensi di bawah ekuilibrium standar interaksi responsif (100 ms) pada kedua perangkat tersebut membuktikan bahwa solusi deteksi logistik berbasis web ini sangat layak (*viable*) untuk beroperasi di lapangan secara mandiri, inklusif, dan bebas dari kewajiban instalasi aplikasi *native*.

### 3.8 Pengaruh Iluminasi Lingkungan Terhadap Stabilitas Deteksi

Kondisi operasional posko pendakian (basecamp) yang fluktuatif menuntut model *computer vision* untuk memiliki daya tahan terhadap variasi iluminasi. Pengujian pada lingkungan dengan intensitas cahaya rendah (*low-light conditions*) membuktikan bahwa jarak spasial antara lensa kamera dan objek logistik bertindak sebagai variabel independen yang sangat deterministik. Kegagalan aktivasi deteksi pada kondisi redup dari jarak jauh dapat diamati secara visual pada Gambar 10, sementara pemulihan stabilitas deteksi pada jarak akuisisi ideal ditunjukkan pada Gambar 11.



**Gambar 10.** Degradasi Performa Inferensi pada Kondisi Iluminasi Redup dan Jarak Jauh



**Gambar 11.** Stabilitas Deteksi pada Kondisi Redup dengan Jarak Akuisisi Ideal

Pada kondisi redup dengan jarak pengambilan gambar yang jauh, citra mengalami kehilangan informasi spasial (*spatial resolution loss*) yang memicu kegagalan deteksi. Namun, model terbukti tetap mampu melokalisasi objek dengan probabilitas 0,76 asalkan diakuisisi dari jarak rapat. Oleh karena itu, dirumuskan Standar Operasional Prosedur (SOP) taktis: akuisisi citra wajib dilakukan pada jarak ideal kurang dari 1 meter untuk menjaga integritas deteksi.

Secara konklusif, integrasi model algoritma YOLOv8 *Nano* yang dilatih dengan 13.792 anomali kelas logistik ini membuktikan bahwa hambatan skenario densitas ekstrem dapat dipecahkan tanpa memerlukan beban komputasi besar. Kesuksesan arsitektur memetakan anomali visual tanpa restriksi jumlah objek mendeskripsikan kesiapan sistem otomasi klasifikasi logistik ini untuk diterapkan dalam skala industri maupun manajerial ekosistem pariwisata pendakian yang sesungguhnya.

## 4. KESIMPULAN

Penelitian ini menyimpulkan bahwa model YOLOv8 Nano sangat tangguh mengklasifikasi multi-objek logistik pendaki pada berbagai tingkat kepadatan visual. Kinerja model mencatatkan nilai *Precision* 0,971, *Recall* 0,954, dan *mAP@50* sebesar 97,8% yang didukung oleh 13.792 *dataset* kustom dan algoritma prapemrosesan *Stretch to*. Ekstraksi fitur ini berhasil mematahkan limitasi penelitian terdahulu dengan mempertahankan stabilitas akurasi di atas 95% pada kondisi *High Dense* hingga *Extreme Clutter* (96,22%). Waktu inferensi lintas platform juga berjalan responsif, yakni 61,48 ms (desktop) dan 72,62 ms (seluler), sehingga layak diimplementasikan sebagai aplikasi web pelaporan mandiri. Namun, sistem memiliki batasan berupa degradasi akurasi saat mendeteksi objek mikro (rokok) pada kondisi saling menutupi, dengan nilai *mAP@50-95* merosot menjadi 0,601 akibat fenomena kemiripan kelas (*inter-class similarity*) dan distorsi pantulan cahaya (*glare*) plastik. Penelitian lanjutan disarankan untuk mengintegrasikan *teknik Slicing Aided Hyper Inference* (SAHI) guna meningkatkan lokalisasi spasial objek kecil, serta menambahkan modul panduan waktu nyata (*real-time guidance*) pada aplikasi untuk memastikan jarak kamera tetap berada di bawah satu meter pada kondisi pencahayaan yang stabil.

## REFERENCES

- [1] Fatihah Ibnu Fiqri, “Lonjakan Pendaki Gunung Lawu saat Libur Panjang, Perhutani Waspada Open Trip Nakal,” *beritajatim*. Accessed: Oct. 03, 2025. [Online]. Available: <https://beritajatim.com/lonjakan-pendaki-gunung-lawu-saat-libur-panjang-perhutani-waspada-open-trip-nakal>
- [2] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, “Object Detection in 20 Years: A Survey,” *Proceedings of the IEEE*, vol. 111, no. 3, pp. 257–276, Mar. 2023, doi: 10.1109/JPROC.2023.3238524.
- [3] Visen and N. Charibaldi, “Penerapan Object Detection Menggunakan Deep Learning Yolov8 Untuk Mengidentifikasi Sampah Anorganik (Maksimal Sepuluh Objek) Dalam Satu Citra,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 1, pp. 195–202, Feb. 2025, doi: 10.25126/jtiik.2025129012.
- [4] M. Hassan, F. Hussain, S. D. Khan, M. Ullah, M. Yamin, and H. Ullah, “CROWD COUNTING USING DEEP LEARNING BASED HEAD DETECTION,” in *IS and T International Symposium on Electronic Imaging Science and Technology*, Society for Imaging Science and Technology, 2023. doi: 10.2352/EI.2023.35.9.IPAS-293.
- [5] L. A. Kusumawardhani, D. K. R. Yasir, I. B. I. Kusuma, S. adhi(s) Wibowo, R. Rahmania, and A. R. Hidayat, “ANALISIS KINERJA DETEKSI GERAKAN DAN PENGENALAN OBJEK PRODUK RITEL BERBASIS YOLOV8,” *Jurnal Elektro dan Telekomunikasi Terapan*, vol. 11, no. 1, pp. 43–53, Jun. 2024, doi: 10.25124/jett.v11i1.7482.
- [6] D. Triyanto, M. Zidan, M. Wahyudi, L. Pujiastuti, U. Bina Sarana Informatika, and S. Antar Bangsa, “Pengembangan Sistem Deteksi Objek Botol Real-Time dengan YOLOv8 untuk Aplikasi Vision,” *Journal Computer Science*, vol. 3, no. 1, 2024.
- [7] A. D. Putra and G. E. Saputra, “Implementation of YOLOv8 Algorithm for Web-Based Detection of Coffee Fruit Ripeness,” *Journal of Artificial Intelligence and Software Engineering*, vol. 5, no. 2, pp. 432–445, 2025, doi: 10.30811/jaise.v5i2.6730.
- [8] S. Studer *et al.*, “Towards CRISP-ML(Q): A Machine Learning Process Model with Quality Assurance Methodology,” Mar. 03, 2021. doi: 10.20944/preprints202103.0135.v1.
- [9] E. Daniati, R. Firliana, A. Imron, and M. F. Aditiya Mufid, “ANALISIS DATA PENJUALAN PADA UMKM KONVEKSI MENGGUNAKAN METODE K-MEANS CLUSTERING DENGAN MENERAPKAN CRISP-DM,” *Prosiding Seminar Nasional Teknologi dan Sistem Informasi*, vol. 5, no. 1, pp. 257–265, Nov. 2025, doi: 10.33005/sitasi.v5i1.2542.
- [10] K. R. Abadi, “Deteksi Logistik Pendaki Dataset,” May 2026, *Roboflow*. [Online]. Available: <https://universe.roboflow.com/kevins-workspace-mdfob/deteksi-logistik-pendaki>
- [11] Sucipto, D. Dwi Prasetya, and T. Widiyaningtyas, “An Evaluation of the Impact of Dataset Size on Classification Performance in the Cognitive Bloom’s Taxonomy,” in *2024 Beyond Technology Summit on Informatics International Conference (BTS-I2C)*, IEEE, Dec. 2024, pp. 131–136. doi: 10.1109/BTS-I2C63534.2024.10941910.
- [12] A. Nugroho, M. A. Soeleman, R. A. Pramunendar, A. Affandy, and A. Nurhindarto, “Peningkatan Performa Ensemble Learning pada Segmentasi Semantik Gambar dengan Teknik Oversampling untuk Class Imbalance,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 4, pp. 899–908, Aug. 2023, doi: 10.25126/jtiik.2024106831.

- [13] A. Ristyawan, A. Nugroho, and T. K. Amarya, “Optimasi Preprocessing Model Random Forest Untuk Prediksi Stroke,” vol. 12, no. 1, pp. 29–44, 2025.
- [14] A. H. Pradhana and E. Daniati, “BENCHMARKING VISION TRANSFORMER KLASIFIKASI VISUAL MASAKAN PADANG DENGAN ROBUSTNESS MELALUI AUGMENTASI DATA,” *Prosiding Seminar Nasional Teknologi dan Sistem Informasi*, vol. 5, no. 1, pp. 152–164, Nov. 2025, doi: 10.33005/sitasi.v5i1.2527.
- [15] J. Terven and D. Cordova-Esparza, “A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS,” *Mach. Learn. Knowl. Extr.*, vol. 5, no. 4, pp. 1680–1716, Feb. 2024, doi: 10.3390/make5040083.
- [16] Ian. Goodfellow, Yoshua. Bengio, and Aaron. Courville, *Deep learning*. The MIT Press, 2017.
- [17] R. Padilla, S. L. Netto, and E. A. B. da Silva, “A Survey on Performance Metrics for Object-Detection Algorithms,” in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, IEEE, Jul. 2020, pp. 237–242. doi: 10.1109/IWSSIP48289.2020.9145130.
- [18] E. Daniati, R. Firliana, and A. S. Wardani, *Dukungan Keputusan dengan Pendekatan Machine Learning*. Penerbit Universitas Nusantara PGRI Kediri, 2024. doi: <https://doi.org/10.29407/punp.vi.68>.
- [19] D. Kreuzberger, N. Kühn, and S. Hirschl, “Machine Learning Operations (MLOps): Overview, Definition, and Architecture,” *IEEE Access*, vol. 11, pp. 31866–31879, 2023, doi: 10.1109/ACCESS.2023.3262138.
- [20] A. Paleyes, R.-G. Urma, and N. D. Lawrence, “Challenges in Deploying Machine Learning: A Survey of Case Studies,” *ACM Comput. Surv.*, vol. 55, no. 6, pp. 1–29, Jul. 2023, doi: 10.1145/3533378.
- [21] M. Kasper-Eulaers, N. Hahn, S. Berger, T. Sebulonsen, Ø. Myrland, and P. E. Kummervold, “Short Communication: Detecting Heavy Goods Vehicles in Rest Areas in Winter Conditions Using YOLOv5,” *Algorithms*, vol. 14, no. 4, p. 114, Mar. 2021, doi: 10.3390/a14040114.



# Bulletin of Information Technology (BIT)

Vol 7, No 2, Juni 2026, Hal. 289 – 301

ISSN 2722-0524 (media online)

DOI [10.47065/bit.v5i2.2786](https://doi.org/10.47065/bit.v5i2.2786)

<https://journal.fkpt.org/index.php/BIT>

