ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

EVALUASI KINERJA PIPELINE YOLOv8 DAN EASYOCR PADA DETEKSI KODE ASET KAMPUS

I Made Darma Putra✉, I Nyoman Eddy Indrayana, Ni Nyoman Harini Puspita
Jurusan Teknologi Informasi, Politeknik Negeri Bali, Badung, Indonesia
Email: darmaputra2017@gmail.com

DOI: https://doi.org/10.46880/jmika.Vol10No2.pp463-470

ABSTRACT

Manual asset tracking on campus is prone to inefficiencies and human error. This study evaluates the performance of a two-stage architecture that combines the You Only Look Once version 8 (YOLOv8) model for text localization and EasyOCR for optical character recognition on physical inventory labels. The scope of the study is limited to printed asset labels (printed text) tested within the Department of Information Technology. Evaluation was conducted using stress testing methods to assess the system's resilience to variations in lighting and perspective distortion. The results show that YOLOv8 is highly robust at localizing label areas, achieving a mean average precision (mAP@50-95) of 70.04%. During the character extraction stage, EasyOCR produced the lowest character error rate (CER) of 14.16% under ideal shooting conditions. However, extraction performance declined, with error rates reaching 30.82% under glare conditions and 48.58% at extreme tilt angles. This study concludes that the YOLOv8 pipeline is highly suitable for object localization in campus environments, but the OCR module requires more adaptive image preprocessing optimization to handle variations in environmental conditions. This research provides a practical contribution in the form of an adaptive computer vision architecture guideline and an objective benchmark for overcoming image degradation distortions in local-scale inventory management systems.

Keyword: Computer Vision, EasyOCR, Inventory, Image Preprocessing, YOLOv8.

ABSTRAK

Pencatatan aset secara manual di lingkungan kampus rentan terhadap inefisiensi dan kesalahan manusia. Penelitian ini mengevaluasi kinerja arsitektur two-stage yang menggabungkan model You Only Look Once versi 8 (YOLOv8) untuk lokalisasi teks dan EasyOCR untuk pengenalan karakter optik pada label inventaris fisik. Ruang lingkup penelitian dibatasi pada label aset tercetak (printed text) yang diuji di lingkungan Jurusan Teknologi Informasi. Evaluasi dilakukan menggunakan metode stress testing untuk menguji ketahanan sistem terhadap variasi pencahayaan dan distorsi sudut pandang. Hasil penelitian menunjukkan bahwa YOLOv8 mampu melokalisasi area label dengan sangat tangguh, mencapai mean Average Precision (mAP@50-95) sebesar 70.04%. Pada tahap ekstraksi karakter, EasyOCR menghasilkan Character Error Rate (CER) terendah sebesar 14.16% pada kondisi pemotretan ideal. Namun, performa ekstraksi mengalami penurunan dengan tingkat error mencapai 30.82% pada kondisi pencahayaan silau (glare) dan 48.58% pada sudut kemiringan ekstrem. Penelitian ini menyimpulkan bahwa pipeline YOLOv8 sangat ideal untuk lokalisasi objek di lingkungan kampus, namun modul OCR memerlukan optimasi prapemrosesan citra yang lebih adaptif untuk menangani variasi kondisi lingkungan. Penelitian ini memberikan kontribusi praktis berupa pedoman arsitektur Computer Vision yang adaptif serta benchmark objektif untuk mengatasi distorsi degradasi citra pada sistem manajemen inventaris berskala lokal.

Kata Kunci: Computer Vision, EasyOCR, Inventaris, Prapemrosesan Citra, YOLOv8.

PENDAHULUAN

Pengelolaan inventaris aset fisik di perguruan tinggi merupakan bagian penting dari pengelolaan kampus. Pelacakan aset untuk barang-barang seperti perangkat elektronik, komputer laboratorium, dan perabotan ruang kelas di Departemen Teknologi Informasi umumnya mengandalkan catatan berbasis spreadsheet atau aplikasi konvensional yang memerlukan entri data secara manual. Selain memakan banyak tenaga dan waktu, proses konvensional ini rentan terhadap kesalahan manusia saat membaca atau memasukkan kode aset ke dalam sistem basis data (Calista et al., 2023; Rakhma et al., 2022). Oleh karena itu, modernisasi sistem pencatatan melalui otomasi pengenalan identitas asset menjadi kebutuhan yang

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

mendesak untuk meningkatkan efisiensi operasional organisasi.

Teknologi *Computer Vision* menawarkan solusi otomatisasi tanpa kontak fisik untuk mempercepat proses identifikasi objek teks di dunia nyata. Arsitektur *You Only Look Once* versi 8 (YOLOv8) merupakan salah satu algoritma terbaik untuk lokalisasi objek berkat keseimbangan mutakhirnya antara kecepatan komputasi dan akurasi penentuan koordinat objek target. Berbagai studi kasus telah membuktikan keandalan YOLOv8 dalam mengidentifikasi area karakter dan label stiker informasi (Sun et al., 2024; Turnbull & Mannix, 2025). Setelah area label berhasil dilokalisasi melalui *bounding box*, fase berikutnya membutuhkan modul *Optical Character Recognition* (OCR) untuk mengekstraksi informasi alfanumerik. Penggunaan modul EasyOCR populer digunakan karena sifatnya yang *open-source*, ringan, dan memiliki performa optimal pada pembacaan teks lingkungan asli (*scene text*). Implementasi integrasi bertingkat (*pipeline*) antara deteksi objek YOLOv8 dan pemrosesan karakter EasyOCR terbukti memberikan hasil yang sangat baik pada domain identifikasi teks terstruktur (Ihksan et al., 2025; Nugroho et al., 2022)

Meskipun integrasi arsitektur two-stage tersebut menawarkan potensi otomatisasi yang tinggi, performa ekstraksi teks secara riil di lapangan sering kali menghadapi kendala mekanis. Karakteristik citra yang diambil langsung di ruang kelas sering mengalami penurunan kualitas atau degradasi visual akibat faktor lingkungan fisik. Degradasi citra, seperti variasi tingkat pencahayaan yang ekstrem dan perubahan sudut pengambilan gambar, terbukti secara empiris menjadi faktor utama penentu lonjakan tingkat kesalahan pembacaan karakter pada modul OCR (Salehudin et al., 2023). Oleh karena itu, penelitian ini bertujuan untuk melakukan evaluasi kinerja secara kritis terhadap pipeline YOLOv8 dan EasyOCR pada deteksi kode aset kampus. Ruang lingkup penelitian ini secara khusus dibatasi pada pengujian label aset fisik tercetak (*printed text*) melalui skenario *stress testing* intensitas cahaya dan distorsi sudut pandang. Hasil evaluasi ini diharapkan dapat memberikan *benchmark* objektif mengenai batas ketahanan sistem OCR sebelum diterapkan pada aplikasi manajemen inventaris secara luas.

Meskipun penelitian terdahulu seperti yang dilakukan oleh (Ihksan et al., 2025) dan (Nugroho et al., 2022) telah berhasil mengintegrasikan model YOLO dan EasyOCR, mayoritas implementasi tersebut dilakukan pada area deteksi dengan pencahayaan terbuka (seperti plat nomor kendaraan) atau lingkungan industri yang terkontrol. Keterbatasan utama dari studi-studi terdahulu adalah kurangnya pengujian ketahanan (*stress testing)* terhadap anomali fisik lokal, seperti pantulan cahaya silau (*glare*) dari lampu ruangan maupun sudut kemiringan saat pemotretan manual oleh staf kampus. Oleh karena itu, research gap yang diisi oleh penelitian ini adalah pemetaan batas toleransi modul OCR saat dihadapkan pada degradasi citra tingkat ekstrem di ruang kelas. Kontribusi dari penelitian ini adalah penyediaan rekomendasi *pipeline* prapemrosesan adaptif yang ringan secara komputasi untuk meningkatkan akurasi sistem deteksi aset tanpa memerlukan infrastruktur perangkat keras berspesifikasi tinggi.

KAJIAN LITERATUR

Konsep Teoretis dan Keberhasilan Empiris

Lokalisasi Objek Berbasis YOLOv8

Secara teoretis, lokalisasi objek dalam *Computer Vision* melibatkan penentuan koordinat spasial berbentuk bounding box yang mengelilingi objek target pada sebuah citra digital. Arsitektur *You Only Look Once* versi 8 (YOLOv8) mengadopsi pendekatan *anchor-free* detection, di mana model memprediksi langsung titik pusat objek dan regresi jarak tepi tanpa bergantung pada *anchor box* prafabrikasi. Pendekatan teoretis ini secara drastis mereduksi kompleksitas komputasi dan meningkatkan kecepatan pemrosesan frame. Bukti empiris mengenai keandalan arsitektur ini ditunjukkan oleh (Sun et al., 2024) dalam modifikasi RA-YOLOv8, yang membuktikan bahwa penambahan mekanisme atensi pada YOLOv8 secara signifikan meningkatkan akurasi bounding box saat melokalisasi teks terstruktur yang padat pada stempel (*seal text detection*) di lingkungan fisik nyata. Keberhasilan empiris serupa juga dilaporkan oleh (Turnbull & Mannix, 2025) yang memanfaatkan ansambel model YOLOv8 untuk mendeteksi dan mengisolasi karakter alfanumerik individual pada naskah papirus kuno. Eksperimen tersebut membuktikan bahwa teoretis fleksibilitas fitur ekstraksi pada YOLOv8 mampu beradaptasi dengan sangat baik pada bentuk geometris teks yang kusam, pudar, dan mengalami degradasi fisik. Dalam domain aplikasi terapan, efisiensi lokalisasi *two-stage pipeline* menggunakan YOLOv8 juga telah divalidasi oleh (Ihksan et al., 2025) serta (Salsabila & Sriani, 2024) pada sistem pengenalan plat nomor kendaraan otomatis, yang mencatat bahwa stabilitas *bounding box* YOLOv8 menjadi penentu utama keberhasilan pemotongan citra (*cropping*) sebelum masuk ke tahap pembacaan teks.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Konsep Teoretis dan Implementasi Empiris Pengenalan Karakter Optik (OCR)

Optical Character Recognition (OCR) adalah proses konversi teoretis dari matriks piksel citra yang berisi representasi grafis teks menjadi kode karakter digital yang dapat diproses oleh mesin. Modul EasyOCR mengimplementasikan teori *framework deep learning* bertingkat yang mengombinasikan *Convolutional Neural Network* (CNN) berbasis arsitektur ResNet untuk ekstraksi fitur taksonomi citra, *Recurrent Neural Network* (RNN) dengan modul Long *Short-Term Memory* (LSTM) untuk analisis ketergantungan sekuensial antar karakter, dan *Connectionist Temporal Classification* (CTC) *loss* sebagai penyelarasan transkripsi akhir. Secara empiris, integrasi *pipeline* dua tahap antara pengidentifikasi objek (YOLO) dan pengekstrak karakter (EasyOCR) telah sukses diuji oleh (Nugroho et al., 2022) untuk membaca informasi teks pada label kardus logistik industri. Penelitian tersebut menunjukkan bahwa pembacaan teks alfanumerik pendek dapat berjalan otomatis secara *real-time*. Namun, bukti empiris dari (Salehudin et al., 2023) memberikan catatan kritis bahwa performa EasyOCR sangat sensitif terhadap penurunan kualitas citra masukan (*image degradation*). Eksperimen mereka mengonfirmasi bahwa *noise digital* dan distorsi resolusi pada potongan citra objek secara linier menurunkan akurasi pengenalan teks alfanumerik, sehingga membutuhkan intervensi algoritma prapemrosesan spasial sebelum citra dialirkan ke dalam jaringan LSTM EasyOCR.

Analisis Degradasi Citra dan Teori Prapemrosesan Spasial

Kondisi lingkungan fisik ruang kelas kampus memicu terjadinya degradasi citra visual akibat variasi iluminasi cahaya dan distorsi geometris saat pemotretan manual. Secara teoretis, pantulan cahaya berlebih memicu munculnya *glare* yang merusak nilai kecerahan lokal piksel, sedangkan pencahayaan redup memangkas kontras antara tinta karakter dengan latar belakang stiker label. Untuk mengatasi distorsi ini, Bilateral Filtering diterapkan dengan prinsip menghaluskan citra (*smoothing*) sekaligus mempertahankan ketajaman garis pembatas (*edge-preserving*) menggunakan kombinasi fungsi *Gaussian spasial* dan radiometrik. Di sisi lain, *Contrast Limited Adaptive Histogram Equalization* (CLAHE) digunakan untuk mengoreksi kontras secara adaptif pada konteks ubin kecil (*tiles*) tanpa mengamplifikasi noise latar belakang.. Bukti empiris keandalan kombinasi kedua algoritma ini dipaparkan oleh (Lu & Huang, 2022) yang sukses mereduksi noise ekstrem akibat kilatan cahaya serta memulihkan detail tepi objek secara presisi. Restorasi ketajaman piksel pada tepi karakter melalui prapemrosesan spasial ini menjadi intervensi yang krusial, karena penurunan kualitas kedalaman citra terbukti secara linier menurunkan akurasi translasi teks alfanumerik pada arsitektur EasyOCR. Melalui pendekatan ini, pengondisian citra hasil pemotretan liar dapat dioptimalkan sebelum dialirkan ke tahap pengenalan karakter.

Kajian Teoretis Metrik Evaluasi Karakter

Evaluasi kinerja sistem pengenalan teks tidak dapat diukur secara akurat menggunakan metrik akurasi klasifikasi biner konvensional. Teori evaluasi OCR bersandar pada konsep Levenshtein Distance, sebuah algoritma penyelarasan string yang menghitung jumlah operasi pengeditan minimum (substitusi, delesi, dan insersi) yang diperlukan untuk mengubah satu untai string prediksi menjadi *string* acuan (*ground truth*). *Metrik Character Error Rate* (CER) mengadopsi teori ini dengan membagi total operasi pengeditan tersebut dengan jumlah total karakter asli pada label.

Dalam konteks identifikasi kode alfanumerik terstruktur seperti nomor inventaris kampus penggunaan metrik *Word Error Rate* (WER) secara teoretis tidak representatif. Hal ini dikarenakan karakter pada kode inventaris tidak membentuk kata berbasis kamus linguistik, sehingga kesalahan pada satu karakter saja akan menghancurkan status validitas seluruh rangkaian kode secara tidak adil. Oleh karena itu, CER dipilih sebagai standar primer karena mampu mendeteksi kesalahan komputasi secara granular hingga tingkat karakter individual.

Meskipun demikian, (Bourne et al., 2026) memberikan kajian empiris terbaru bahwa nilai CER dapat menjadi bias atau tidak terdefenisi jika terjadi kegagalan fatal pada tahap pembacaan batas atau analisis tata letak (*page parsing error*). Temuan empiris ini memvalidasi pentingnya memisahkan evaluasi kerja antara tahap lokalisasi (YOLOv8) dan tahap ekstraksi teks (EasyOCR) untuk mengidentifikasi letak *bottleneck* sistem secara objektif. Melalui pemisahan metrik ini, batas kemampuan maksimal dari arsitektur *pipeline* yang digunakan dapat dipetakan secara akurat melalui pengujian ketahanan (*robustness testing*) multidimensi pada variasi kondisi lingkungan fisik riil, sejalan dengan metodologi yang dikembangkan oleh (Mutiara Chairunnisa et al., 2025)

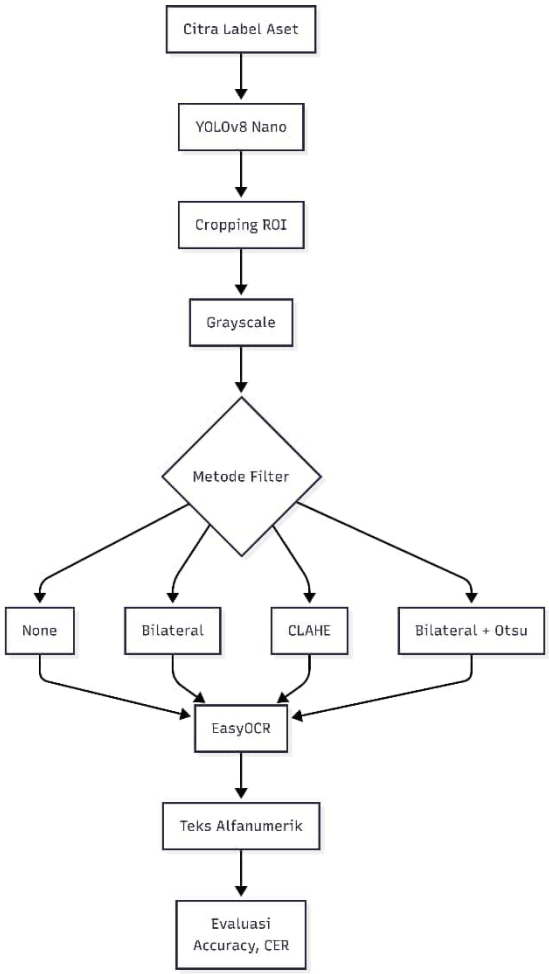
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

METODE PENELITIAN

Rancangan Sistem *(Pipeline* Arsitektur)

Penelitian ini menerapkan pendekatan eksperimental kuantitatif untuk mengevaluasi *two-stage pipeline* pada *Computer Vision*. Alur kerja sistem dirancang secara sekuensial yang terdiri dari empat tahap utama, yang seperti diilustrasikan pada Gambar 1



Gambar 1. Alur Pipeline Sistem

Dari Gambar 1 Alur *Pipeline* Sistem tersebut terdapat empat poin penting sebagai berikut:

1. Input Citra: Sistem menerima masukan berupa citra digital dari label asset kampus.
2. Lokalisasi Objek (YOLOv8): Model memproses citra untuk mendeteksi posisi stiker label aset secara akurat. Pemilihan varian YOLOv8 tipe Nano didasarkan pada pertimbangan trade-off yang optimal antara efisiensi beban komputasi dan akurasi tinggi untuk deteksi objek kelas tunggal (*single-class object*). Model yang ringan (*lightweight*) ini dipilih agar *pipeline* sistem dapat diimplementasikan secara efisien pada aplikasi berbasis web kampus tanpa membebani sumber daya server. Model menghasilkan keluaran berupa titik koordinat *bounding box*.
3. Prapemrosesan Citra (*Crop & Filter*): Area piksel di dalam bounding box dipotong (cropping) dan langsung dikonversi ke format skala keabuan (grayscale). Citra grayscale tersebut kemudian dialirkan secara independen ke dalam salah satu dari empat percabangan skenario filter (*dev mode*), yaitu: kontrol tanpa filter (*None*), Bilateral Filtering, *Contrast Limited Adaptive Histogram Equalization* (CLAHE), atau kombinasi Bilateral Filtering dengan Otsu Thresholding.
4. Ekstraksi Teks (EasyOCR): Citra hasil keluaran dari masing-masing skenario filter diproses oleh modul EasyOCR untuk mentranslasikan matriks piksel karakter alfanumerik menjadi teks digital berupa deret nomor inventaris.

Manajemen Dataset dan Augmentasi

Aset data dalam penelitian ini dibagi menjadi dua kelompok independen: dataset pelatihan (*training/validation*) dan dataset pengujian ketahanan (*stress testing*). Penggunaan ukuran dataset awal sebanyak 144 citra primer dan 40 citra uji (testing) dijustifikasi oleh tingginya homogenitas objek target. Karena ruang lingkup penelitian dibatasi secara spesifik pada varian stiker aset fisik seragam yang tercetak di Jurusan Teknologi Informasi, pendekatan *transfer learning* pada *pre-trained* model tidak membutuhkan volume *big data* untuk mencapai konvergensi. Selain itu, alokasi 40 citra uji dirancang bukan untuk generalisasi global, melainkan sebagai sampel purposif untuk mengisolasi dan menguji empat skenario degradasi lingkungan fisik secara spesifik. Pengelolaan dataset pelatihan dilakukan melalui *platform* Roboflow dengan urutan metodologi sebagai berikut:

a. Isolasi Data Awal: Sebanyak 144 citra primer yang diambil langsung di lingkungan Jurusan Teknologi Informasi (JTI), dialokasikan ke dalam folder *Training* dan *Validation* dengan rasio otomatis. Proses isolasi ini mengunci data validasi dari paparan efek digital.
b. Prapemrosesan Spasial: Seluruh citra diseragamkan ukurannya (*resize*) menjadi resolusi 640 x 640 piksel sesuai arsitektur *input* standar YOLOv8
c. Amplifikasi Data (Augmentasi): Proses augmentasi diterapkan secara ekslusif hanya pada folder *Training* dengan perkalian 3x berbasis parameter fisik: rotasi geometris (-15° hingga +15°), kecerahan (*brightness* -20% hingga +20%), efek kabur (*blur* 1px), dan noise matriks (1.2%). Proses ini menghasilkan total 348 citra siap latih.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Skenario Pengujian *Hyperparameter*

Untuk mendapatkan model lokalisasi yang paling presisi, fase pelatihan model YOLOv8 dilakukan melalui eksperimen komparatif menggunakan tiga variasi konfigurasi *Hyperparameter*. Batas atas pelatihan sebesar 100 *epoch* dengan fitur penghentian dini (*early stopping*) aktif pada nilai "patience=25". Tiga scenario yang diuji meliputi:

a. Skenario A: Ukuran *batch* 16 menggunakan *optimizer* bawaan (*auto* /AdamW).
b. Skenario B: Ukuran *batch* 32 menggunakan *optimizer* bawaan (*auto* /AdamW).
c. Skenario C: Ukuran *batch* 16 menggunakan optimizer SGD.

Skenario Pengujian Ketahanan (*Stress Testing*) OCR

Fase pengujian modul EasyOCR dilakukan menggunakan 40 citra independen yang sepenuhnya baru dan bersih dari manipulasi augmentasi digital. Untuk menguji batas ketahanan operasional sistem terhadap gangguan lingkungan fisik ruang kelas, dataset uji didistribusikan secara merata ke dalam empat skenario pemotretan. Representasi visual sampel citra dari masing-masing skenario pengujian diilustrasikan pada Gambar 2, dengan rincian sebagai berikut:

1. Kondisi Normal (10 citra): Pemotretan tegak lurus pada intensitas cahaya ruangan normal
2. Kondisi Redup (10 citra): Pemotretan pada kondisi minim pencahayaan.
3. Kondisi Silau (10 citra): Pemotretan menggunakan lampu kilat kamera (*flash*) yang memicu pantulan cahaya ekstrem (*glare*) pada permukaan label.
4. Distorsi Kemiringan (10 citra): Pemotretan dengan sudut kemiringan ekstrem untuk memicu distorsi perspektif.



Gambar 2. Sampel Citra Uji Fisik

Setiap skenario diuji menggunakan *dev mode* pada aplikasi manajemen inventaris InvenTrack untuk membandingkan kinerja empat mode prapemrosesan: Mode kontrol (tanpa filter), Bilateral Filter, *Contrast Limited Adaptive Histogram Equalization* (CLAHE), dan kombinasi Bilateral Filter dengan Otsu *Thresholding*.

Metrik Evaluasi Performa

Evaluasi kinerja kuantitatif dipisahkan secara objektif pada masing-masing modul. Performa lokalisasi YOLOv8 diukur menggunakan metrik *mean Average Precision* pada ambang batas *Intersection over Union* 50% (mAP@50) serta mAP@50-95 untuk menilai ketatnya ukuran *bounding box*. Sementara itu, performa ekstraksi teks EasyOCR diukur menggunakan metrik *Character Error Rate* (CER) berbasis jarak Levenshtein, yang dikalkulasi setelah melalui tahapan normalisasi teks alfanumerik. Persamaan matematis untuk CER didefinisikan sebagai berikut:

$$CER = \frac{S + D + I}{N}$$

Di mana variabel S adalah jumlah substitusi karakter (salah baca), D adalah jumlah delesi (karakter hilang), I adalah jumlah insersi (karakter tambahan di luar konteks), dan N merepresentasikan total panjang karakter asli pada label acuan (*ground truth*).

HASIL DAN PEMBAHASAN

Evaluasi Performa Lokalisasi YOLOv8

Tahap pertama dari *pipeline* pengenalan kode aset melibatkan pelatihan model YOLOv8 tipe Nano untuk mengisolasi area stiker label inventaris dari latar belakang citra ruang kelas. Pengujian dilakukan melalui tiga skenario variasi *hyperparameter* (Skenario A, B, dan C) untuk memetakan pengaruh ukuran *batch* dan jenis *optimizer* terhadap kerapatan *bounding box*. Hasil evaluasi kuantitatif dari fase pelatihan ini dirangkum dalam Tabel 1.

Tabel 1. Hasil Evaluasi Kuantitatif Pelatihan Model YOLOv8 Nano

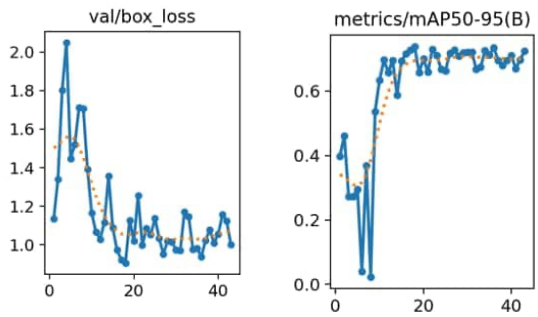
Skenario	Konfigurasi *Hyperparameter*	Precision (%)	Recall (%)
Skenario A	*Batch* 16, *Optimizer Auto* (AdamW)	94,50	92,10
Skenario B	*Batch* 32, *Optimizer Auto* (AdamW)	91,20	88,50
Skenario C	*Batch 16, Optimizer* SGD	78,40	76,00

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Berdasarkan data pada Tabel 1, Skenario A yang memanfaatkan ukuran *batch* 16 dan *optimizer* AdamW keluar sebagai model terbaik dengan nilai mAP@50-95 tertinggi sebesar 70,04%. Fenomena ini secara teoretis dipicu oleh sifat adaptif dari *optimizer* AdamW dalam memperbarui bobot secara mandiri melalui kalkulasi momentum kuadratik, yang terbukti jauh lebih stabil dibandingkan dengan *Stochastic Gradient Descent* (SGD) pada Skenario C (mAP@50-95 sebesar 52,15%) ketika memproses dataset primer yang berukuran spesifik.

Stabilitas fase pelatihan pada Skenario A ini dapat divalidasi secara visual melalui kurva konvergensi metrik evaluasi yang diilustrasikan pada Gambar 3. Grafik *validation box loss* (Gambar 3A) menunjukkan tren penurunan eksponensial yang mulus tanpa adanya indikasi lonjakan kembali (*loss spike*) hingga akhir *epoch*. Penurunan nilai loss ini berjalan beriringan dengan lonjakan kurva *mean Average Precision* (mAP@50-95) yang merangkak naik dan mencapai titik saturasi stabil di kisaran nilai 0,70 (Gambar 3b). Fenomena ini membuktikan secara empiris bahwa model YOLOv8 Nano pada Skenario A berhasil melewati fase konvergensi dengan sangat sehat tanpa mengalami overfitting maupun underfitting.



a *b*

Gambar 1. Kurva Kinerja Pelatihan Skenario A: (a) Grafik Validation Box Loss, (b) Grafik Metrik Map@50-95.

Lebih lanjut, penurunan performa pada Skenario B (mAP@50-95 sebesar 65,20%) membuktikan bahwa peningkatan ukuran *batch* menjadi 32 berpotensi memangkas varians gradien pada fase penyusutan fungsi kerugian (*loss convergence*), sehingga model rentan terjebak dalam kondisi *local minima*. Oleh karena itu, bobot model terbaik (best.pt) dari Skenario A secara resmi diintegrasikan ke dalam arsitektur *backend* aplikasi Manajemen Inventaris InvenTrack sebagai komponen utama pemotongan citra (*cropping*) otomatis sebelum dialirkan ke tahap prapemrosesan filter spasial.

Evaluasi Kinerja Ekstraksi Teks EasyOCR

Fase pengujian ketahanan (*stress testing*) diterapkan secara independen menggunakan 40 citra label inventaris baru untuk mengukur tingkat ketahanan operasional modul EasyOCR. Pengujian ini mensimulasikan gangguan lingkungan fisik nyata yang dibagi secara merata ke dalam empat skenario lingkungan (Normal, Redup, Silau, dan Miring) dengan mengevaluasi pengaruh empat metode prapemrosesan spasial. Nilai kesalahan pembacaan karakter dihitung secara objektif menggunakan metrik *Character Error Rate* (CER) setelah *string* hasil pembacaan melalui proses normalisasi alfanumerik. Hasil komparasi CER secara menyeluruh dipaparkan pada Tabel 2.

Tabel 2. Rata-Rata Character Error Rate (CER) EasyOCR pada Skenario Uji Fisik (%)

Skenario Pemotretan Lingkungan Fisik	Mode: *None*	Mode: CLAHE	Mode: Bilateral	Mode: Bilateral. + Otsu
Kondisi Normal	14,16	17,84	53,46	91,06
Kondisi Redup	57,10	30,08	89,18	84,84
Kondisi Silau	30,82	26,96	48,64	97,65
Distorsi Kemiringan	48,58	*N/A*	*N/A*	*N/A*

Pengujian filter prapemrosesan tidak diaplikasikan (*N/A*) pada skenario kemiringan ekstrem karena manipulasi kontras warna tidak dapat merekonstruksi distorsi geometris.

Pembahasan Skenario Kondisi Normal

Pada skenario kondisi normal dengan intensitas cahaya ruangan kelas yang ideal, mode tanpa prapemrosesan (*None*) mencatatkan akurasi pengenalan terbaik dengan nilai rata-rata CER terendah yaitu 14,16%. Sebaliknya, penerapan kombinasi filter Bilateral dan Otsu Thresholding justru memicu anomali degradasi performa yang masif dengan nilai CER melonjak drastis hingga 91,06%. Kegagalan fatal ini disebabkan oleh karakteristik font cetakan stiker label aset JTI yang memiliki ketebalan garis piksel yang tipis. Penerapan filter Bilateral yang melunakkan intensitas warna (*smoothing*) dikombinasikan dengan ambang batas binerisasi Otsu secara matematis melenyapkan keberadaan fitur garis tipis tersebut (*white-out effect*), mengubah area stiker menjadi blok putih kosong sehingga modul EasyOCR tidak mampu mendeteksi keberadaan teks.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Pembahasan Skenario Kondisi Redup

Evaluasi pada kondisi redup mengungkap kerentanan masif arsitektur EasyOCR jika bekerja tanpa intervensi filter, di mana nilai kesalahan pada mode None membengkak hingga 57,10%. Degradasi pencahayaan melenyapkan kontras alami antara tinta karakter alfanumerik dengan latar belakang stiker, memicu terjadinya galat delesi (karakter terlewat).

Namun demikian, implementasi filter *Contrast Limited Adaptive Histogram Equalization* (CLAHE) terbukti secara efektif mampu bertindak sebagai solusi pemulihan citra adaptif dengan memangkas tingkat kesalahan hingga ke level CER 30,08%. Keberhasilan empiris ini sejalan dengan teori manipulasi histogram lokal, di mana CLAHE membagi potongan gambar stiker menjadi konteks ubin kecil (tiles) dan mendistribusikan kembali kecerahan piksel secara lokal tanpa mengeksploitasi *noise*, sehingga mampu memulihkan kedalaman fitur garis tepi huruf yang tenggelam di balik bayangan kusam.

Pembahasan Skenario Kondisi Silau (*Glare)*

Skenario paparan lampu kilat kamera menciptakan bercak pantulan cahaya putih yang merusak nilai kecerahan lokal piksel stiker. Kondisi ini menurunkan ketajaman model tanpa filter (*None*) menjadi CER 30,82% akibat terjadinya galat substitusi (contohnya angka 0 dibaca sebagai huruf O atau angka 3 dikenali sebagai 8).

Filter CLAHE kembali membuktikan keunggulannya dengan mencatatkan performa terbaik pada level CER 26,96%. Berbeda dengan binerisasi Otsu yang lumpuh total (CER 97,65%) karena mengubah kilatan glare menjadi blok putih mutlak yang menghapus teks, sifat adaptif CLAHE sukses meredam intensitas over-exposure dan mengekstraksi sisa-sisa kontras hitam dari karakter alfanumerik yang tersembunyi di bawah pantulan cahaya, memperkuat temuan empiris terdahulu dari (Lu & Huang, 2022).

Pembahasan Skenario Distorsi Kemiringan (*Tilt*)

Pengujian sudut pemotretan ekstrem menghasilkan nilai CER kritis sebesar 48,58% pada mode None. Analisis log data mentah mengonfirmasi terjadinya fenomena *Spatial Misalignment* (Ketidaksejajaran Spasial) dan *Reading Order Error*, di mana EasyOCR melakukan kesalahan pengurutan token teks (misalnya teks NUP atau tahun terdeteksi lebih awal dibandingkan kode lokasi karena posisi koordinat sumbu Y yang bergeser akibat efek perspektif). Selain itu, penyempitan jarak geometris menyebabkan antar-karakter alfanumerik terlihat berdempetan, memicu bias interpretasi arsitektur LSTM EasyOCR. Hal ini memvalidasi teori dari (Bourne et al., 2026) mengenai pentingnya memisahkan metrik lokalisasi dan ekstraksi guna mengisolasi letak *bottleneck* sistem secara objektif pada kondisi pemotretan liar.

Komparasi Kinerja terhadap Studi Terdahulu

Sebagai perbandingan dengan studi *Computer Vision* sebelumnya, evaluasi ini sejalan dengan temuan (Salehudin et al., 2023) yang menegaskan bahwa modul EasyOCR sangat rentan terhadap degradasi citra, di mana kesalahan pembacaan melonjak signifikan tanpa adanya prapemrosesan. Namun demikian, penelitian ini menemukan hasil yang bertentangan dengan implementasi (Lu & Huang, 2022) yang sangat mengandalkan Bilateral Filtering untuk restorasi citra industri. Pada konteks deteksi teks tercetak berukuran tipis di label aset, penelitian ini membuktikan bahwa kombinasi algoritma Bilateral & Otsu Thresholding justru merusak integritas matriks piksel huruf (*white-out effect*). Evaluasi komparatif dalam studi ini menegaskan bahwa penggunaan filter adaptif (CLAHE) jauh lebih superior untuk restorasi kontras teks di bawah pantulan cahaya ekstrem dibandingkan metode *smoothing* statis. Di sisi lain, mAP@50-95 sebesar 70,04% yang dicetak oleh YOLOv8 Nano pada lokalisasi *bounding box* membuktikan bahwa arsitektur kelas ringan (*lightweight*) mampu menyamai reliabilitas pelacakan objek dari model-model pendahulunya yang lebih berat, memberikan daya saing komputasi yang tinggi untuk pengembangan sistem inventaris terpadu di masa depan.

KESIMPULAN

Evaluasi arsitektur two-stage pipeline untuk pengenalan kode inventaris membuktikan bahwa model YOLOv8 tipe Nano memiliki tingkat keandalan yang sangat tangguh dalam melokalisasi area stiker aset secara spesifik. Pengujian *hyperparameter* menunjukkan bahwa penggunaan optimizer AdamW dengan ukuran batch 16 menghasilkan akurasi lokalisasi paling stabil dengan nilai mAP@50-95 mencapai 70,04%. Namun, fase ekstraksi teks alfanumerik oleh arsitektur EasyOCR terbukti sangat fluktuatif dan bergantung pada intervensi prapemrosesan spasial saat dihadapkan pada gangguan lingkungan fisik riil.

Pengujian ketahanan (*stress testing*) menyimpulkan bahwa citra tanpa filter (*baseline*) memberikan kinerja terbaik pada kondisi pencahayaan ideal dengan tingkat *Character Error Rate* (CER) sebesar 14,16%. Pada skenario degradasi lingkungan

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

berupa minim cahaya (redup) dan pantulan cahaya kilat (*glare*), algoritma adaptif CLAHE terbukti menjadi prapemrosesan paling krusial yang sukses mereduksi tingkat kesalahan berturut-turut menjadi 30,08% dan 26,96%, sekaligus mematahkan kegagalan fatal (*white-out effect*) yang dihasilkan oleh algoritma binerisasi Otsu. Di sisi lain, distorsi perspektif akibat kemiringan sudut pandang terbukti menjadi parameter paling merusak tanpa bisa dipulihkan oleh manipulasi warna, dengan nilai CER mencapai 48,58% akibat terjadinya ketidaksejajaran spasial antar-karakter.

Secara keseluruhan, integrasi YOLOv8 dan EasyOCR ini sangat layak untuk menjadi fondasi otomatisasi pencatatan fisik pada aplikasi manajemen inventaris. Sebagai rekomendasi pengembangan lebih lanjut, sistem disarankan untuk mengimplementasikan mekanisme deteksi iluminasi cerdas guna mengaktifkan filter CLAHE secara dinamis sesuai kondisi cahaya ruangan, serta mengadopsi algoritma transformasi perspektif (*perspective un-warping*) untuk mengoreksi distorsi sudut sebelum citra dialirkan ke dalam modul ekstraksi teks.

DAFTAR PUSTAKA

Bourne, J., Simbeye, M., & Nockels, J. (2026). *The Character Error Vector: Decomposable errors for page-level OCR evaluation*. http://arxiv.org/abs/2604.06160

Calista, S., Husaein, A., & Ilmu Komputer, F. (2023). Perancangan Sistem Informasi Inventory Barang Berbasis Web pada Toko Laris Furniture Jambi. *Jurnal Manajemen Teknologi Dan Sistem Informasi (JMS)*, *3*(2). https://doi.org/10.33998/jms.v3i1

Ihksan, M., Fauzi, D., Isnan, M., & Sari, I. (2025). Arsitektur Sistem Automatic Number Plate Recognition Berbasis Web dengan Pipeline Deteksi YOLOv8 dan Pengenalan Karakter EasyOCR. *206 JURTI*, *9*(2).

Lu, P., & Huang, Q. (2022). Robotic Weld Image Enhancement Based on Improved Bilateral Filtering and CLAHE Algorithm. *Electronics (Switzerland)*, *11*(21). https://doi.org/10.3390/electronics11213629

Mutiara Chairunnisa, C., Husni, N. L., & Kusumanto, R. D. (2025). Robustness Testing of TrOCR for Multi-Condition Food Ingredient Labels Detected By YOLO 1. *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)*, *8*(2), 538–543. https://doi.org/10.24014/ijaidm.v8i2.37301

Nugroho, S. A., Kholis, N., Endryansyah, E., & Baskoro, F. (2022). Rancang Bangun Sistem Deteksi Label Kardus Berbasis Model Kecerdasan Buatan YOLO dan EasyOCR serta ESP32-CAM. *Jurnal Teknik Elektro, 11*(2), 190-200.

Rakhma, S. A., Tullah, R., & Maisaroh, S. (2022). Sistem Monitoring Data Aset dan Inventaris IT Berbasis Web pada PT. Pan Brothers Tbk. *Jurnal Topik Global, 1*(2).

Salehudin, M. A. M., Basah, S. N., Yazid, H., Basaruddin, K. S., Safar, M. J. A., Som, M. H. M., & Sidek, K. A. (2023). Analysis of Optical Character Recognition using EasyOCR under Image Degradation. *Journal of Physics: Conference Series*, *2641*(1). https://doi.org/10.1088/1742-6596/2641/1/012001

Salsabila, N., & Sriani, S. (2024). Enhancing Automated Vehicle License Plate Recognition with YOLOv8 and EasyOCR. *Journal of Information Systems and Informatics*, *6*(3), 1577–1597. https://doi.org/10.51519/journalisi.v6i3.848

Sun, H., Tan, C., Pang, S., Wang, H., & Huang, B. (2024). RA-YOLOv8: An Improved YOLOv8 Seal Text Detection Method. *Electronics (Switzerland)*, *13*(15). https://doi.org/10.3390/electronics13153001

Turnbull, R., & Mannix, E. (2025). Detecting and recognizing characters in Greek papyri with YOLOv8, DeiT and SimCLR. *International Journal on Document Analysis and Recognition (IJDAR)*, *28*(2), 277–285. https://doi.org/10.1007/s10032-024-00504-8

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.