

Pengembangan Otomasi Inventaris Farmasi Rumah Sakit Gigi dan Mulut dengan YOLO11s-CNN

Khairul Abdi*, Amalia, Mohammad Andri Budiman

¹Sains Data dan Kecerdasan Buatan, Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara, Medan, Indonesia

Email: ^{1,*}khairulabdi@usu.ac.id, ¹amalia@usu.ac.id, ¹mandrib@usu.ac.id

Email Penulis Korespondensi: khairulabdi@usu.ac.id*

Submitted: 24/05/2026; Accepted: 29/06/2026; Published: 30/06/2026

Abstrak—Pengelolaan inventaris farmasi di Rumah Sakit Gigi dan Mulut memerlukan proses yang cepat, teliti, dan terdokumentasi karena ketersediaan obat serta bahan kedokteran gigi berpengaruh langsung terhadap kelancaran pelayanan klinis. Penelitian ini mengembangkan prototipe otomasi inventaris berbasis citra menggunakan model deteksi objek YOLO11s yang dibangun di atas prinsip *Convolutional Neural Network*. *Dataset* terdiri atas 2.562 citra, 11 kelas produk, dan 5.486 anotasi *bounding box* yang dibagi menjadi data latih, validasi, dan uji dengan rasio 70:15:15. Tahapan penelitian meliputi akuisisi citra, anotasi, *quality control dataset*, pra-proses, augmentasi, pelatihan, evaluasi, serta perancangan integrasi hasil deteksi ke rekap inventaris. Pada data uji terbatas, model menghasilkan *precision* 99,37%, *recall* 98,86%, *mAP@0.5* 99,41%, dan *mAP@0.5:0.95* 90,03%. Hasil ini menunjukkan bahwa YOLO11s berpotensi membantu pengenalan jenis produk, lokalisasi objek, dan perhitungan jumlah item pada skenario inventaris farmasi yang tercakup dalam *dataset*. Namun, klaim kinerja masih perlu divalidasi lebih lanjut pada jumlah kelas yang lebih besar dan kondisi rak farmasi nyata yang lebih beragam.

Kata Kunci: inventaris farmasi; Rumah Sakit Gigi dan Mulut; Convolutional Neural Network; YOLO11s; deteksi objek; otomasi inventaris

Abstract—Pharmaceutical inventory management in Dental and Oral Hospitals requires a fast, accurate, and traceable process because the availability of medicines and dental materials directly affects clinical service continuity. This study developed an image-based inventory automation prototype using the YOLO11s object detection model based on Convolutional Neural Network principles. The dataset contained 2,562 images, 11 product classes, and 5,486 bounding-box annotations, split into training, validation, and test sets using a 70:15:15 ratio. The workflow included image acquisition, annotation, dataset quality control, preprocessing, augmentation, model training, evaluation, and the design of detection-result integration into an inventory summary. On the limited test set, the model achieved 99.37% precision, 98.86% recall, 99.41% *mAP@0.5*, and 90.03% *mAP@0.5:0.95*. These results indicate that YOLO11s has the potential to support product recognition, object localization, and item counting for pharmaceutical inventory scenarios covered by the dataset. Nevertheless, the performance claims require further validation using more product classes and more diverse real pharmacy-shelf conditions.

Keywords: pharmaceutical inventory; Dental and Oral Hospital; Convolutional Neural Network; YOLO11s; object detection; inventory automation

1. PENDAHULUAN

Transformasi digital pada layanan kesehatan tidak hanya terjadi pada rekam medis elektronik dan pelayanan pasien, tetapi juga pada aktivitas pendukung yang menentukan kelancaran operasional rumah sakit. Salah satu aktivitas penting adalah pengelolaan inventaris farmasi. Pada Rumah Sakit Gigi dan Mulut (RSGM), ketersediaan obat, bahan habis pakai, bahan kedokteran gigi, serta material pendukung tindakan klinis perlu dipantau secara akurat agar pelayanan tidak terganggu. Inventaris farmasi berbeda dari inventaris umum karena berkaitan dengan umur simpan, variasi kemasan, ukuran label yang kecil, dan kemiripan visual antarproduk [1], [2].

Permasalahan pencatatan manual semakin penting karena farmasi rumah sakit menangani aliran barang yang langsung mendukung pelayanan klinis. Kajian *intelligent inventory management* dan *artificial intelligence* di bidang farmasi menunjukkan bahwa pendekatan cerdas dapat membantu menurunkan biaya inventaris, meningkatkan level layanan, dan memperbaiki respons rantai pasok. Namun, sebagian besar pendekatan masih bertumpu pada data transaksi dan sistem informasi, sedangkan pengenalan stok secara langsung dari citra rak atau area penyimpanan masih perlu dikembangkan, terutama pada lingkungan RSGM [3], [4], [5].

Kebutuhan terhadap pendekatan berbasis citra muncul karena kegiatan inventaris pada rak farmasi tidak selalu dapat direpresentasikan hanya melalui data transaksi. Barang dapat berpindah posisi, tertinggal di rak, tertukar lokasi, atau belum tercatat ketika proses pelayanan berlangsung cepat. Dalam situasi seperti itu, citra rak menyediakan bukti visual yang dapat ditelusuri kembali. Apabila citra tersebut dapat diproses otomatis, sistem berpeluang memberi informasi awal tentang jenis dan jumlah item sebelum petugas melakukan verifikasi akhir. Pendekatan ini relevan untuk RSGM karena variasi produk meliputi bahan kedokteran gigi, obat, alat kecil, dan bahan habis pakai yang sering memiliki bentuk kemasan tidak seragam. Oleh karena itu, penelitian tidak hanya memerlukan model yang akurat, tetapi juga alur kerja yang mampu menghubungkan hasil deteksi dengan proses inventaris yang sebenarnya.

Computer vision menawarkan solusi yang relevan karena citra digital dapat digunakan untuk mengenali jenis produk, menentukan posisi barang, dan menghitung jumlah item secara otomatis. Pada domain ritel dan farmasi, pendekatan *deep learning* telah digunakan untuk membaca kondisi rak, mengenali obat berkemasan mirip, serta mendeteksi objek pada citra medis maupun farmasi. Temuan tersebut menunjukkan bahwa pengenalan

produk farmasi berbasis citra layak dikembangkan, tetapi membutuhkan *dataset* representatif dan model yang mampu menangani variasi ukuran, orientasi, pencahayaan, serta kemiripan kemasan [6].

Convolutional Neural Network (CNN) merupakan pendekatan utama pada pengolahan citra modern karena mampu mengekstraksi fitur spasial secara bertingkat. Pada tugas deteksi objek, CNN tidak hanya mengenali kelas produk, tetapi juga memperkirakan lokasi objek dalam bentuk *bounding box*. Keluarga YOLO menjadi salah satu pendekatan populer karena memproses deteksi dalam satu tahap sehingga sesuai untuk aplikasi yang membutuhkan keseimbangan antara akurasi dan kecepatan [7], [8].

Meskipun demikian, penerapan CNN pada inventaris farmasi tidak dapat dianggap sama dengan pengenalan objek umum. Produk farmasi dapat berukuran kecil, memiliki permukaan reflektif, atau menampilkan teks yang rapat pada area label. Beberapa produk juga mempunyai warna dominan yang mirip dan hanya dibedakan oleh bentuk kemasan atau tulisan kecil. Pada citra multiobjek, sebagian kemasan dapat tertutup objek lain sehingga model harus tetap mengenali pola visual yang tidak selalu lengkap. Karena itu, keberhasilan sistem sangat dipengaruhi oleh kualitas anotasi, keseimbangan kelas, variasi citra, dan pemilihan *threshold* inferensi. Aspek-aspek tersebut menjadi perhatian utama dalam penelitian ini agar hasil yang diperoleh tidak hanya tinggi secara numerik, tetapi juga masuk akal untuk skenario kerja inventaris [9], [10].

Berdasarkan kesenjangan tersebut, penelitian ini menyajikan pengembangan otomasi inventaris farmasi RSGM berbasis CNN. Kontribusi penelitian terletak pada penyusunan *dataset* citra 11 kelas produk farmasi, proses anotasi dan validasi data, pelatihan model YOLO11s, evaluasi kuantitatif menggunakan *precision*, *recall*, *mAP@0.5*, dan *mAP@0.5:0.95*, serta rancangan rekap hasil deteksi menuju sistem inventaris. Berbeda dari penelitian yang hanya berfokus pada klasifikasi gambar tunggal, penelitian ini menekankan deteksi objek *multiitem* sehingga keluaran model dapat diterjemahkan menjadi jumlah item per kelas. Dengan demikian, penelitian diharapkan memberi kontribusi praktis bagi percepatan pencatatan inventaris dan kontribusi ilmiah pada penerapan CNN untuk manajemen inventaris farmasi rumah sakit.

Selain akurasi model, penelitian pada konteks inventaris farmasi perlu memperhatikan keterhubungan antara deteksi visual dan keputusan operasional. Hasil deteksi harus dapat diterjemahkan menjadi data yang mudah diverifikasi, misalnya jumlah item per kelas, lokasi rak, dan waktu pengambilan citra. Dengan demikian, sistem tidak hanya berfungsi sebagai alat pengenalan gambar, tetapi juga sebagai pendukung proses kerja farmasi yang menuntut ketertelusuran, konsistensi pencatatan, dan pengendalian stok yang lebih objektif [2], [5], [6].

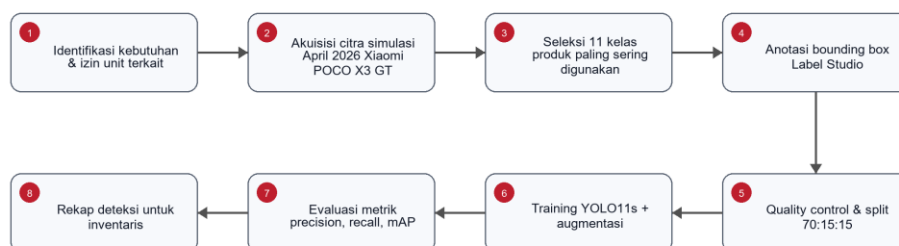
2. METODOLOGI PENELITIAN

2.1 Desain dan Alur Penelitian

Penelitian ini merupakan penelitian terapan dengan pendekatan pengembangan prototipe awal, karena tujuan utamanya adalah merancang dan mengevaluasi model deteksi objek untuk membantu proses otomasi inventaris farmasi. Model utama yang digunakan adalah YOLO11s, yaitu detektor objek satu tahap berbasis prinsip *Convolutional Neural Network* (CNN). Desain penelitian disusun secara berurutan mulai dari penentuan kebutuhan dan ruang lingkup data, pengambilan citra produk dalam kondisi simulasi, anotasi, pengendalian mutu dataset, pembagian data, pelatihan model, evaluasi, hingga rancangan integrasi hasil inferensi ke rekap inventaris [9].

Alur penelitian dibuat agar setiap tahap menghasilkan artefak yang dapat ditelusuri. Citra produk menjadi masukan awal, file label YOLO menjadi dasar pelatihan, bobot *best.pt* menjadi hasil model terbaik, sedangkan tabel rekap deteksi menjadi keluaran yang dapat digunakan dalam proses inventaris. Diagram alur desain penelitian ditampilkan pada Gambar 1, sedangkan ringkasan tahapan dan keluaran utama disajikan pada Tabel 1.

Alur Desain Penelitian Prototipe Otomasi Inventaris Farmasi



Batasan data: hanya citra produk farmasi/bahan kedokteran gigi; tidak memuat data pasien.

Gambar 1. Diagram alur desain penelitian

Tabel 1. Tahapan penelitian dan keluaran utama

Tahap	Kegiatan	Keluaran
Identifikasi kebutuhan	Mengidentifikasi masalah pencatatan inventaris farmasi dan kebutuhan deteksi produk berbasis citra.	Rumusan masalah dan kebutuhan sistem
Perizinan dan batasan data	Memastikan data hanya berupa citra produk farmasi/bahan kedokteran gigi, tidak memuat data pasien, serta telah memperoleh izin dari unit terkait.	Status izin dan batasan data nonpasien
Penentuan kelas produk	Memilih 11 kelas produk yang paling sering digunakan sebagai ruang lingkup prototipe awal.	Daftar 11 kelas objek prioritas
Akuisisi data simulasi	Mengambil citra produk dalam kondisi simulasi pada April 2026 menggunakan HP Xiaomi POCO X3 GT dengan resolusi 3472 x 4624 piksel.	Dataset citra produk farmasi
Anotasi	Memberi bounding box dan label kelas menggunakan Label Studio, kemudian mengeksport label ke format YOLO.	File label per citra
Quality control	Memeriksa pasangan citra-label, indeks kelas, posisi bounding box, duplikasi, blur, dan kualitas pencahayaan.	Dataset siap latih
Praproses dan augmentasi	Menyesuaikan ukuran input 640 piksel serta menerapkan augmentasi translate, scale, flip, mosaic, HSV, dan erasing.	Citra latih lebih bervariasi
Training, validasi, dan test	Melatih YOLO11s pada split 70:15:15, memantau validasi, dan menguji subset test sebagai evaluasi akhir.	Bobot model, kurva, confusion matrix, dan metrik evaluasi
Inferensi inventaris	Mendeteksi objek baru, menyaring confidence, menerapkan NMS, lalu merekap jumlah item per kelas.	Rekap stok visual

2.2 Lokasi, Periode, Perangkat, dan Status Izin Pengambilan Data

Pengambilan citra dilakukan dalam kondisi simulasi pada lingkungan farmasi dan area penyimpanan produk farmasi. Kondisi simulasi digunakan agar variasi susunan objek, jarak kamera, sudut pengambilan, orientasi kemasan, latar, bayangan, pantulan, dan pencahayaan dapat dikendalikan tanpa mengganggu operasional pelayanan. Area pengambilan data tidak mencakup ruang pelayanan pasien, rekam medis, resep pasien, atau dokumen klinis. Objek yang direkam hanya berupa kemasan produk farmasi dan bahan kedokteran gigi yang menjadi target inventaris.

Periode pengambilan citra dilakukan pada bulan April 2026. Pencantuman periode akuisisi diperlukan agar waktu pengambilan data dapat ditelusuri kembali dan dapat dibedakan dari tahap anotasi, pelatihan, validasi, serta pengujian model.

Citra diambil menggunakan HP Xiaomi POCO X3 GT dengan resolusi keluaran citra 3472 x 4624 piksel. Informasi perangkat dan resolusi dicantumkan untuk memperjelas kondisi akuisisi, sehingga prosedur pengambilan data dapat direplikasi pada penelitian lanjutan atau pada pengujian lapangan berikutnya.

Dataset yang digunakan hanya berisi citra produk farmasi dan bahan kedokteran gigi. Dataset tidak memuat wajah pasien, identitas pasien, rekam medis, resep, atau data klinis individual. Pengambilan data telah memperoleh izin dari unit terkait, sehingga batasan data dan ruang lingkup pengambilan citra dapat dipertanggungjawabkan. Dengan pembatasan tersebut, risiko privasi pasien dapat diminimalkan karena penelitian tidak menggunakan data subjek manusia.

2.3 Dataset, Kelas Objek, dan Anotasi

Dataset penelitian berjumlah 2.562 citra digital produk farmasi dan bahan kedokteran gigi. Dataset mencakup objek tunggal dan multiobjek agar model dapat mempelajari variasi produk pada skenario inventaris. Citra dipertahankan dalam variasi realistis, seperti perbedaan orientasi, posisi objek, ukuran tampilan, latar, bayangan, dan pencahayaan, karena sistem inventaris berbasis citra perlu bekerja pada kondisi rak atau area penyimpanan yang tidak selalu ideal.

Sebanyak 11 kelas produk dipilih sebagai ruang lingkup prototipe awal. Pemilihan kelas didasarkan pada dua pertimbangan utama. Pertama, produk tersebut termasuk produk farmasi dan bahan kedokteran gigi yang paling sering digunakan pada unit terkait sehingga relevan dengan kebutuhan inventaris harian. Kedua, penelitian masih berada pada tahap prototipe awal sehingga jumlah kelas dibatasi agar proses pengambilan citra, anotasi, pelatihan, dan evaluasi dapat dilakukan secara terkendali. Pembatasan jumlah kelas juga memungkinkan model diuji lebih dahulu pada kelompok produk prioritas sebelum diperluas ke jenis produk farmasi lainnya.

Anotasi dilakukan menggunakan Label Studio. Setiap objek diberi bounding box yang mengelilingi kemasan produk secara ketat, kemudian diekspor ke format YOLO dengan struktur class_id, x_center, y_center, width, dan height. Berdasarkan rekap anotasi, terdapat 5.486 label pada 11 kelas objek. Dataset dibagi pada level citra dengan rasio 70:15:15, yaitu sekitar 1.794 citra untuk data latih, 384 citra untuk data validasi, dan 384 citra

untuk data uji. Berdasarkan rekap anotasi, pembagian subset terdiri atas 3.840 anotasi pada data latih, 823 anotasi pada data validasi, dan 823 anotasi pada data uji. Distribusi tiap kelas ditampilkan pada Tabel 2.

Sebelum pelatihan, dataset melalui quality control untuk memastikan konsistensi pasangan file citra dan label, validitas indeks kelas, posisi bounding box, serta kesesuaian nama kelas pada data.yaml. Pemeriksaan ini penting karena kesalahan kecil pada label dapat menyebabkan model belajar pola yang keliru. Citra yang terlalu gelap, blur ekstrem, atau tidak informatif ditinjau kembali, sedangkan citra dengan variasi yang masih realistis tetap dipertahankan agar model lebih tahan terhadap kondisi lapangan.

Tabel 2. Distribusi anotasi latih, validasi, dan uji, AP@0.5 dan AP@0.5:0.95 per kelas

No	Kelas	Total	Latih	Validasi	Uji	AP@0.5	AP@0.5:0.95
1	cocoabutter_for_gcfuji_ionomerfilling	605	423	91	91	0.995	0,889
2	e-flex_blue-057	453	317	68	68	0.972	0,924
3	edenta_fgdiambdurs_studentkit	437	306	65	66	0.995	0,930
4	flexiband_matrixbands_dentamerica_no-1	401	281	60	60	0.995	0,931
5	formocresol	520	364	78	78	0.995	0,899
6	gc_goldlabel_hsposterior_extra	573	401	86	86	0.995	0,882
7	guttaperchapoints_120pcs	398	279	60	59	0.995	0,883
8	indodent_absorbent_paperpoints_size45-80	560	392	84	84	0.995	0,899
9	indodent_guttaperchapoints_size45-80	580	406	87	87	0.994	0,909
10	precaine_stroberrry	513	359	77	77	0.995	0,864
11	shileld_dentin_desensitizer	446	312	67	67	0.995	0,893
	Total	5.486	3.840	823	823	-	-

2.4 Konfigurasi Model dan Evaluasi

Pelatihan model dilakukan menggunakan YOLO11s dengan bobot awal yolo11s.pt. Pada arsitektur ini, backbone CNN mengekstraksi feature map melalui operasi konvolusi bertingkat, neck menggabungkan informasi multiskala, dan detection head menghasilkan koordinat bounding box, label kelas, serta confidence score. Model dilatih secara supervised learning karena setiap citra telah memiliki anotasi lokasi dan kelas target. Prinsip operasi konvolusi dan rancangan jaringan sederhana-efektif menjadi dasar pembelajaran fitur visual pada arsitektur CNN [7], [8], [9].

Konfigurasi utama pelatihan meliputi 80 epoch, batch size 8, ukuran input 640 x 640 piksel, optimizer otomatis, learning rate awal 0,01, momentum 0,937, weight decay 0,0005, workers 2, cache aktif, dan device GPU 0. Pelatihan dijalankan pada komputer dengan RAM 16 GB, CPU Intel Core i3-9100F, dan GPU NVIDIA RTX 2080 8 GB. Waktu training tercatat sekitar 3.588,63 detik atau 59,81 menit.

Augmentasi digunakan untuk memperluas variasi visual data latih [11]. Parameter yang digunakan meliputi translasi 0,1, skala 0,5, flipping horizontal 0,5, mosaic 1,0, pengubahan HSV, dan erasing 0,4. Rotasi eksplisit bernilai 0,0 sehingga variasi sudut terutama berasal dari data asli dan transformasi lain dalam pipeline. Evaluasi dilakukan menggunakan Intersection over Union (IoU), precision, recall, F1-score, mAP@0.5, dan mAP@0.5:0.95. Metrik precision dan recall digunakan untuk menilai kesalahan false positive dan false negative, sedangkan mAP digunakan untuk menilai kualitas deteksi dan lokalisasi objek pada berbagai ambang IoU.

Tabel 3. Konfigurasi utama pelatihan YOLO11s

Parameter	Nilai	Parameter	Nilai
Model awal	yolo11s.pt	Epoch	80
Batch size	8	Image size	640
Learning rate awal	0,01	Momentum	0,937
Weight decay	0,0005	Patience	30
Mosaic	1,0	Flip horizontal	0,5
Translate	0,1	Scale	0,5

Parameter	Nilai	Parameter	Nilai
GPU	NVIDIA RTX 2080 8 GB	Waktu pelatihan	59,81 menit

3. HASIL DAN PEMBAHASAN

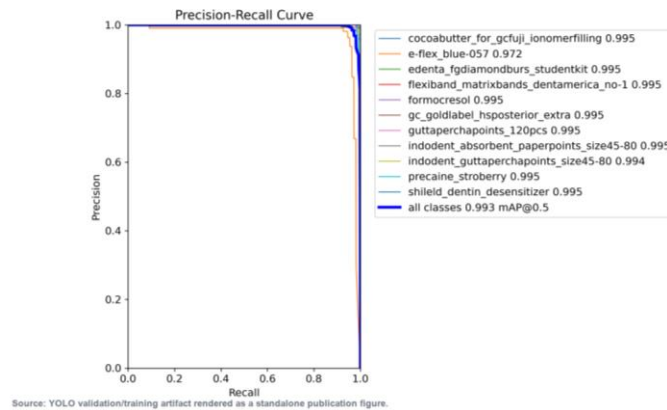
3.1 Evaluasi Global Model

Hasil pelatihan memperlihatkan bahwa *box loss*, *classification loss*, dan *distribution focal loss* menurun secara konsisten sampai *epoch* ke-80. Pada saat yang sama, *precision*, *recall*, mAP@0.5, dan mAP@0.5:0.95 meningkat menuju kondisi konvergen. Pola ini menunjukkan bahwa model mempelajari representasi visual secara stabil dan tidak memperlihatkan gejala divergensi yang dominan [12].

Ringkasan metrik pada Tabel 4 membedakan hasil validasi dan hasil uji. Perbedaan ini penting karena tesis terbaru menggunakan *subset uji* sebagai dasar performa akhir. Pada data uji, model memperoleh *precision* 99,37%, *recall* 98,86%, mAP@0.5 99,41%, dan mAP@0.5:0.95 90,03%. Nilai ini menunjukkan bahwa model tidak hanya baik pada data validasi, tetapi juga mampu mempertahankan kemampuan generalisasi pada data yang tidak digunakan selama pelatihan maupun pemilihan bobot.

Precision-Recall Curve

The precision-recall relationship indicates strong separability and a favorable operating region for threshold selection.



Gambar 2. Kurva Precision-Recall per kelas pada data validasi.

Perbedaan antara mAP@0.5 dan mAP@0.5:0.95 memberikan informasi penting bagi implementasi. Pada ambang IoU 0,5, model dianggap benar apabila kotak prediksi telah bertumpang tindih memadai dengan *ground truth*. Pada evaluasi 0,5:0,95, model harus mempertahankan kualitas lokalisasi pada ambang yang lebih ketat [13]. Karena itu, nilai mAP@0.5:0.95 yang lebih rendah dibanding mAP@0.5 menunjukkan bahwa masih ada ruang perbaikan pada presisi batas kotak, terutama untuk objek kecil, objek miring, dan kemasan yang sebagian tertutup.

Perkembangan metrik selama pelatihan juga menunjukkan bahwa model tidak hanya menghafal data latih. *Precision* meningkat sampai 0,9941 pada *epoch* ke-80, *recall* bertahan tinggi dan berakhir pada 0,9875, sedangkan mAP@0.5:0.95 meningkat dari sekitar 0,7976 pada *epoch* pertama menjadi 0,8828 pada *epoch* terakhir. Kenaikan bertahap ini menunjukkan bahwa model terus memperbaiki kualitas lokalisasi meskipun kemampuan deteksi dasar telah tinggi sejak awal.

Artefak *training* yang dihasilkan juga memperkuat reproduibilitas eksperimen. *File* best.pt berukuran sekitar 18,29 MB menjadi bobot utama untuk inferensi, sedangkan results.csv dan args.yaml menyimpan riwayat metrik serta konfigurasi yang digunakan. Grafik results.png, kurva *precision-recall*, kurva *F1-confidence*, *confusion matrix*, dan labels.jpg memberikan bukti tambahan bahwa eksperimen tidak hanya menghasilkan nilai akhir, tetapi juga menyimpan jejak proses pembelajaran. Dalam penelitian berbasis *deep learning*, dokumentasi artefak seperti ini penting karena memungkinkan peneliti lain meninjau konfigurasi, membaca pola konvergensi, dan mengulangi *training* pada *dataset* yang sama.

Tabel 4. Ringkasan metrik evaluasi global

Metrik	Validasi	Uji	Interpretasi
Jumlah citra evaluasi	384	384	Subset validasi dan uji terpisah dari data latih
Precision	99,41%	99,37%	Prediksi positif model sangat sedikit menghasilkan false positive

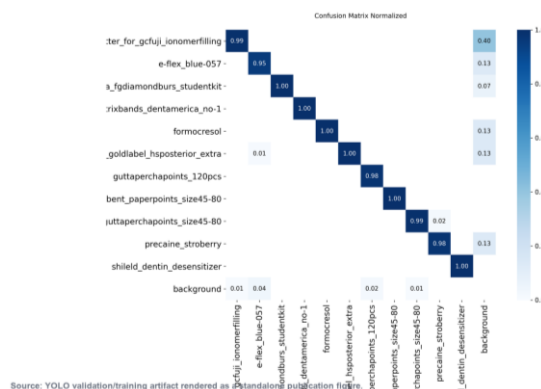
Metrik	Validasi	Uji	Interpretasi
Recall	98,75%	98,86%	Sebagian besar objek aktual pada data validasi berhasil ditemukan
mAP@0.5	99,28%	99,41%	Deteksi objek sangat kuat pada ambang IoU 0,5
mAP@0.5:0.95	88,28%	90,03%	Lokalisasi tetap baik pada evaluasi IoU yang lebih ketat
Epoch terbaik	80	-	Kinerja terbaik diperoleh pada akhir pelatihan
Waktu pelatihan	59,81 menit	-	Training realistis untuk perangkat komputasi menengah

3.2 Pembahasan Performa per Kelas

Performa per kelas pada data uji menunjukkan bahwa hampir seluruh produk memperoleh AP@0.5 pada rentang 0,994-0,995. Capaian ini menandakan bahwa karakter visual sebagian besar kelas dapat dipelajari oleh CNN secara konsisten. Produk berbentuk kotak dengan warna dominan dan batas tegas, seperti formocresol, gc_goldlabel_hsposterior_extra, dan shileld_dentin_desensitizer, cenderung lebih mudah dipisahkan dari latar maupun kelas lain.

Normalized Confusion Matrix

The normalized matrix emphasizes per-class accuracy and helps identify residual confusion in relative terms.



Gambar 3. Confusion matrix ternormalisasi hasil validasi model

Kelas e-flex_blue-057 memperoleh AP@0.5 sebesar 0,972, masih sangat baik tetapi lebih rendah dibanding kelas lain. Berdasarkan karakteristik visualnya, kelas ini berupa *blister* tipis memanjang dengan teks rapat sehingga lebih sensitif terhadap orientasi, pantulan, ukuran objek, dan jarak kamera. Pada implementasi lapangan, kelas seperti ini membutuhkan pencahayaan lebih merata dan *framing* yang cukup dekat agar label, tepi, dan bentuk *blister* tetap terbaca.

Confusion matrix ternormalisasi pada Gambar 3 memperlihatkan bahwa nilai diagonal mendominasi hampir seluruh kelas. Kesalahan yang paling sering muncul berupa objek yang diprediksi sebagai background, yaitu false negative ketika model tidak cukup yakin terhadap objek. Sebagian kecil kekeliruan antar-kelas juga muncul, misalnya beberapa deteksi indodont_guttaperchapoins_size45-80 terbaca sebagai precaine_stroberry. Meskipun demikian, recall ternormalisasi per kelas berada pada kisaran sekitar 0,95 sampai 1,00, sehingga model tidak hanya kuat secara rata-rata tetapi juga relatif stabil pada sebagian besar kelas.

Dari sisi sebaran data, jumlah anotasi tiap kelas berada pada rentang yang cukup seimbang. Kelas dengan anotasi terbanyak adalah cocoabutter_for_gcfuji_ionomerfilling sebanyak 605 label, sedangkan kelas dengan anotasi paling sedikit adalah guttaperchapoins_120pcs sebanyak 398 label. Rentang ini tidak terlalu ekstrem sehingga model memiliki kesempatan belajar yang relatif proporsional pada seluruh kelas. Keseimbangan seperti ini penting karena sistem inventaris tidak boleh hanya akurat pada kelas mayoritas. Kegagalan pada satu kelas tertentu tetap dapat menimbulkan kesalahan stok, walaupun nilai rata-rata metrik tampak sangat tinggi.

Karakteristik objek juga memengaruhi jenis fitur yang paling berguna bagi CNN. Produk dengan kotak dan batas tegas cenderung memberi sinyal bentuk yang stabil. Produk dengan botol silindris memerlukan pembacaan kombinasi bentuk botol, warna label, dan kontur melengkung. Produk berwarna putih atau terang lebih bergantung pada ketajaman tepi dan kontras dengan latar. Sementara itu, produk dengan teks kecil tidak selalu dikenali karena teks terbaca secara eksplisit, melainkan karena model mempelajari pola global kemasan, susunan warna, dan struktur visual. Hal ini menjelaskan mengapa model tetap mampu bekerja baik walaupun penelitian belum menggunakan OCR.

3.3 Implikasi Operasional Inventaris

Hasil deteksi yang tinggi dapat diterjemahkan menjadi alur inventaris yang lebih operasional. Model menerima citra rak atau area penyimpanan, menghasilkan bounding box, label kelas, dan confidence score, kemudian sistem merekap jumlah objek per kelas. Rekap tersebut dapat disimpan sebagai log deteksi dan digunakan sebagai dasar verifikasi stok. Dengan cara ini, keluaran model tidak berhenti sebagai visualisasi kotak pada gambar, tetapi menjadi data inventaris yang dapat membantu stock opname, penerimaan barang, penyusunan rak, audit, dan pencocokan antara stok fisik dan catatan digital. Contoh hasil prediksi pada Gambar 4 memperlihatkan bahwa keluaran model dapat dibaca langsung sebagai informasi kelas, lokasi, dan jumlah objek [5], [6].

Validation Prediction Example 1

Standalone qualitative result showing predicted boxes and labels on a validation mini-batch sample.



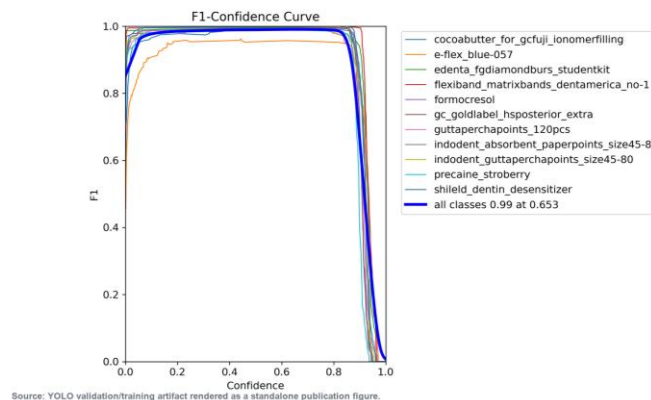
Gambar 4. Contoh hasil prediksi bounding box pada data validasi

Berdasarkan kurva F1-confidence pada Gambar 5, titik kerja agregat yang baik berada pada confidence sekitar 0,653 dengan F1 sekitar 0,99. Untuk implementasi awal, rentang threshold 0,60-0,70 dapat dipertimbangkan karena menjaga keseimbangan antara precision dan recall. Jika sistem memprioritaskan penghindaran overcount stok, threshold dapat dinaikkan agar prediksi menjadi lebih konservatif. Jika sistem memprioritaskan penghindaran objek terlewat, threshold dapat diturunkan secara hati-hati sambil memantau false positive.

Pemilihan *threshold* juga perlu dikaitkan dengan karakter kelas. Untuk kelas dengan AP sangat tinggi dan visual yang tegas, *threshold* menengah cenderung cukup aman karena model sudah stabil. Untuk kelas yang lebih menantang, seperti e-flex_blue-057, *threshold* terlalu tinggi dapat meningkatkan risiko objek valid tidak terhitung. Sebaliknya, *threshold* terlalu rendah dapat memunculkan deteksi yang kurang yakin pada latar atau kemasan mirip. Karena itu, nilai operasional sebaiknya tidak dipandang sebagai angka tetap, melainkan parameter yang dapat dikalibrasi berdasarkan data lapangan, jenis rak, dan prioritas unit farmasi.

F1-Confidence Curve

This curve can be used to select the confidence threshold that balances precision and recall most effectively.



Gambar 5. Kurva F1-Confidence untuk pemilihan ambang operasional

Secara praktis, sistem berbasis CNN dapat ditempatkan sebagai alat bantu, bukan pengganti penuh validasi petugas pada tahap awal. Petugas tetap dapat memeriksa prediksi dengan *confidence* rendah atau objek yang tertutup sebagian sebelum data stok diperbarui. Pendekatan *human-in-the-loop* ini lebih realistis untuk lingkungan rumah sakit karena akurasi model harus berjalan bersama prosedur kerja, dokumentasi, dan tanggung jawab pengguna. Dengan integrasi basis data, setiap hasil deteksi dapat disimpan bersama *metadata* seperti waktu pengambilan citra, lokasi rak, pengguna, dan hasil validasi akhir [6].

Pada skenario penerimaan barang, model dapat membantu memeriksa apakah produk yang datang sesuai dengan daftar permintaan. Pada penyusunan rak, sistem dapat memberikan dokumentasi visual tentang jenis barang dan posisinya. Pada *stock* opname berkala, beberapa citra rak dapat diproses dan hasilnya direkap menjadi jumlah item per kelas. Pada audit, log citra dan hasil deteksi dapat menjadi bukti pendukung ketika terjadi perbedaan antara stok fisik dan catatan digital. Dengan demikian, manfaat model bukan hanya pada kecepatan membaca citra, tetapi juga pada peningkatan keterlacakan proses inventaris.

Dari perspektif manajerial, otomasi berbasis citra dapat membantu rumah sakit membangun tata kelola stok yang lebih objektif. Selama ini, pencatatan manual bergantung pada konsistensi pengamatan petugas. Dengan dokumentasi visual, setiap hasil hitung dapat diperiksa ulang. Namun, penerapan tetap perlu disertai prosedur operasional standar. SOP perlu mengatur cara pengambilan citra, jarak kamera, pencahayaan minimum, posisi rak, siapa yang melakukan validasi, serta kapan data hasil deteksi boleh memperbarui stok. Tanpa tata kelola tersebut, model yang akurat sekalipun dapat menghasilkan data yang kurang konsisten karena proses *input* berbeda-beda antarpetugas.

Jika sistem diterapkan bertahap, kelompok produk yang sudah masuk *dataset* dapat menjadi area percontohan. Pada tahap awal, model dapat digunakan untuk mendampingi proses *stock* opname, bukan langsung memperbarui jumlah stok secara otomatis. Setelah hasil deteksi dan validasi manual menunjukkan konsistensi, integrasi dapat diperluas ke modul transaksi, notifikasi stok minimum, atau riwayat audit. Pendekatan bertahap ini mengurangi risiko perubahan alur kerja yang terlalu mendadak dan memberi kesempatan kepada petugas untuk memahami batas kemampuan model.

Keluaran deteksi juga dapat membantu analisis historis. Apabila citra, hasil prediksi, dan hasil validasi disimpan secara berkala, pengelola farmasi dapat menelusuri kapan suatu kelas produk sering tidak terbaca, kelas mana yang sering membutuhkan koreksi manual, dan kondisi rak seperti apa yang memicu kesalahan. Informasi tersebut dapat dipakai untuk memperbaiki cara penataan barang, menambah data pelatihan pada kelas yang sulit, atau menyusun pedoman pengambilan gambar yang lebih konsisten. Dengan demikian, sistem deteksi tidak hanya menjadi alat hitung sesaat, tetapi juga sumber data untuk perbaikan proses inventaris secara berkelanjutan.

Aspek ini penting karena tujuan akhir otomasi bukan sekadar mengganti aktivitas manual, melainkan meningkatkan kualitas keputusan. Ketika hasil visual, rekap jumlah, dan koreksi pengguna terdokumentasi, rumah sakit memiliki dasar yang lebih kuat untuk mengevaluasi kebutuhan pembelian, mengidentifikasi produk yang sering bermasalah, dan merancang pengawasan stok yang lebih akuntabel. Data tersebut juga dapat memperkuat evaluasi berkala rutin.

3.4 Komparasi Model Pemanding

Untuk memperkuat posisi YOLO11s sebagai model utama, penelitian ini menambahkan eksperimen komparasi dengan beberapa arsitektur deteksi objek pada dataset dan pembagian data yang sama. Model pembanding yang digunakan meliputi YOLO11n sebagai model ringan, YOLO11m sebagai model berukuran lebih besar, YOLOv8s sebagai pembanding versi YOLO sebelumnya, RT-DETR-L sebagai detektor berbasis transformer, serta Faster R-CNN ResNet50-FPN sebagai pembanding *two-stage detector*. Komparasi dilakukan untuk melihat keseimbangan antara akurasi, stabilitas deteksi, ukuran model, waktu pelatihan, dan kecepatan inferensi [7], [8], [9], [10], [14], [15], [16]. Nilai pada Tabel 5 dan Tabel 6 berasal dari rangkaian eksperimen komparatif terpisah; karena itu, nilainya digunakan untuk menilai posisi relatif model dan tidak menggantikan hasil akhir model utama pada Tabel 4.

Tabel 5. Komparasi hasil validasi model pembanding

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Waktu latih	Ukuran bobot
YOLO11n	0,980	0,990	0,986	0,855	±1,10 jam	20,28 MB
YOLO11s	0,980	0,989	0,985	0,857	±1,57 jam	18,28 MB
YOLO11m	0,980	0,990	0,984	0,851	±2,34 jam	38,64 MB
YOLOv8s	0,980	0,991	0,986	0,858	±1,42 jam	21,47 MB
RT-DETR-L	0,974	0,988	0,983	0,850	±13,67 jam	63,19 MB
Faster R-CNN	-	-	0,990	0,835	±10 jam	315,14 MB

Berdasarkan Tabel 5, seluruh model pembanding selain YOLO11n pada pengujian awal validasi menunjukkan performa tinggi. YOLOv8s memperoleh nilai mAP@0.5:0.95 validasi tertinggi, yaitu 0,858, tetapi selisihnya terhadap YOLO11s hanya 0,001. YOLO11s memperoleh mAP@0.5:0.95 sebesar 0,857 dengan ukuran bobot sekitar 18,28 MB, sehingga lebih ringan dibanding YOLO11m, RT-DETR-L, dan Faster R-CNN. YOLO11m tidak memberikan peningkatan akurasi yang sebanding dengan ukuran modelnya, sedangkan RT-DETR-L membutuhkan waktu pelatihan paling lama. Faster R-CNN menghasilkan mAP@0.5 validasi yang tinggi, tetapi ukuran bobotnya jauh lebih besar dibanding model YOLO.

Tabel 6. Komparasi hasil pengujian pada data uji

Model	mAP@0.5 uji	F1 terbaik	Confidence F1	mAP@0.5:0.95 uji	Inferensi
YOLO11n	0,773	0,63	0,136	-	-
YOLO11s	0,991	0,99	0,715	0,872	±5,4 ms/citra
YOLO11m	0,991	0,99	0,714	-	-
YOLOv8s	0,991	0,99	0,661	-	-
RT-DETR-L	0,989	0,98	0,704	-	-
Faster R-CNN	0,993	-	-	0,849	55,59 ms/citra

Hasil uji pada Tabel 6 memperlihatkan bahwa YOLO11s, YOLO11m, dan YOLOv8s mencapai mAP@0.5 sebesar 0,991 dengan F1-score terbaik 0,99. RT-DETR-L tetap kompetitif dengan mAP@0.5 sebesar 0,989 dan F1-score 0,98. *Faster R-CNN* memperoleh mAP@0.5 tertinggi, yaitu 0,993, tetapi mAP@0.5:0.95 sebesar 0,849 masih lebih rendah dibanding YOLO11s yang mencapai 0,872. Selain itu, waktu inferensi *Faster R-CNN* sebesar 55,59 ms per citra jauh lebih lambat dibanding YOLO11s yang berada pada kisaran ±5,4 ms per citra. Hal ini menunjukkan bahwa *Faster R-CNN* mampu melakukan deteksi dengan baik, tetapi kurang efisien untuk kebutuhan sistem inventaris yang memerlukan pemrosesan cepat.

Pengujian ulang YOLO11n menunjukkan mAP@0.5 sebesar 0,773 dan F1-score terbaik 0,63 pada *confidence* 0,136. Hasil ini jauh lebih rendah dibanding YOLO11s, YOLO11m, YOLOv8s, RT-DETR-L, dan *Faster R-CNN*. Berdasarkan kurva dan *confusion matrix*, YOLO11n cenderung mengalami banyak *false negative*, terutama pada kelas yang memiliki kemasan kecil, tipis, atau visual mirip. Dengan demikian, meskipun YOLO11n memiliki arsitektur lebih ringan, kapasitas modelnya belum cukup stabil untuk kebutuhan deteksi produk farmasi pada *dataset* ini.

Secara keseluruhan, komparasi ini memperkuat pemilihan YOLO11s sebagai model utama. YOLOv8s memang sedikit lebih tinggi pada mAP@0.5:0.95 validasi, tetapi selisihnya sangat kecil dan tidak menunjukkan keunggulan yang dominan pada data uji. YOLO11m memiliki performa setara pada mAP@0.5 *test*, tetapi ukuran modelnya lebih besar. RT-DETR-L memiliki performa baik, namun waktu pelatihannya jauh lebih panjang. *Faster R-CNN* unggul tipis pada mAP@0.5 *test*, tetapi lebih lambat dan kurang unggul pada mAP@0.5:0.95. Oleh karena itu, YOLO11s dipilih karena memberikan keseimbangan terbaik antara akurasi, ukuran model, dan kecepatan inferensi untuk prototipe otomasi inventaris farmasi RSGM.

3.5 Keterbatasan dan Arah Pengembangan

Meskipun hasil penelitian sangat baik, terdapat beberapa keterbatasan. Pertama, ruang lingkup data masih terbatas pada 11 kelas produk farmasi. Lingkungan rumah sakit yang sebenarnya dapat memiliki variasi item yang jauh lebih besar, termasuk perubahan desain kemasan dari produsen. Kedua, sebagian besar citra berasal dari kondisi yang relatif terkontrol. Pada implementasi nyata, rak yang sangat padat, occlusion berat, blur akibat gerak kamera, dan cahaya yang berubah-ubah dapat menurunkan performa. Ketiga, uji komparatif telah dilakukan pada beberapa arsitektur, tetapi tidak semua artefak pengujian menyediakan metrik yang sama secara lengkap, khususnya mAP@0.5:0.95 dan waktu inferensi pada seluruh model pembandingan. Oleh karena itu, komparasi lanjutan dengan format pelaporan metrik yang seragam tetap diperlukan [10], [11], [17], [18].

Pengembangan lanjutan perlu diarahkan pada perluasan *dataset*, penambahan kelas, pengambilan citra lintas perangkat, dan pengujian langsung pada ruang penyimpanan farmasi. Selain itu, integrasi dengan OCR, *barcode*, atau *QR code* dapat meningkatkan validasi identitas produk, terutama pada kemasan yang sangat mirip. Teknik peningkatan citra seperti CLAHE dan konsolidasi prediksi seperti *Weighted Boxes Fusion* juga dapat diuji apabila data lapangan menunjukkan masalah kontras, pantulan, atau kotak prediksi ganda [5], [10].

Keterbatasan lain adalah penelitian masih menggunakan citra statis. Pada penerapan yang lebih luas, rumah sakit mungkin membutuhkan pemindaian rak menggunakan kamera bergerak atau video agar area penyimpanan yang panjang dapat dicakup lebih cepat. Skenario video akan membawa tantangan baru seperti *motion blur*, pelacakan objek *antarframe*, duplikasi hitungan, dan kebutuhan inferensi *real-time*. Oleh karena itu, tahap berikutnya tidak cukup hanya menambah akurasi model, tetapi juga perlu menguji stabilitas sistem pada alur kerja yang lebih dinamis.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan prototipe otomasi inventaris farmasi RSGM berbasis CNN menggunakan YOLO11s. *Dataset* yang digunakan terdiri atas 2.562 citra dan 5.486 anotasi pada 11 kelas objek, dengan pembagian 70:15:15 menjadi sekitar 1.794 citra *training*, 384 citra validasi, dan 384 citra *test*. Pada data uji, model memperoleh *precision* 99,37%, *recall* 98,86%, mAP@0.5 99,41%, dan mAP@0.5:0.95 90,03%. Nilai tersebut

menunjukkan bahwa model mampu mengenali kelas produk dan menempatkan *bounding box* secara akurat pada data yang tidak digunakan selama pelatihan maupun pemilihan bobot.

Secara operasional, keluaran model dapat direkap menjadi jumlah item per kelas sehingga berpotensi mendukung *stock opname*, verifikasi barang, dan integrasi *dashboard* inventaris. Hasil komparasi menunjukkan bahwa YOLO11s memberikan keseimbangan yang baik antara akurasi dan efisiensi dibanding YOLO11n, YOLO11m, YOLOv8s, RT-DETR-L, dan Faster R-CNN. Meskipun Faster R-CNN memperoleh mAP@0.5 sedikit lebih tinggi, YOLO11s lebih unggul pada mAP@0.5:0.95 serta memiliki inferensi yang jauh lebih cepat. Keterbatasan utama penelitian adalah jumlah kelas yang masih terbatas, kondisi data yang belum sepenuhnya merepresentasikan lapangan, dan belum adanya uji *end-to-end* di rumah sakit. Penelitian berikutnya disarankan memperluas *dataset*, melakukan uji lapangan pada rak farmasi nyata, menambahkan validasi OCR atau *barcode*, serta mengevaluasi dampak sistem terhadap waktu kerja dan akurasi pencatatan inventaris.

REFERENCES

- [1] A. Almeman, "The digital transformation in pharmacy: embracing online platforms and the cosmeceutical paradigm shift," *J. Heal. Popul. Nutr.*, vol. 43, no. 1, pp. 1–22, 2024, doi: 10.1186/s41043-024-00550-2.
- [2] E. P. B. Abimanyu, S. Satibi, and D. Endarti, "Digital Technologies in Hospital Pharmacy: A Systematic Review of Their Impact on Efficiency, Safety, and Inventory Management," *J. Prev. Med. Public Health*, pp. 1–11, 2025, doi: 10.3961/jpmph.25.495.
- [3] M. D. Simpson and H. S. Qasim, "Clinical and Operational Applications of Artificial Intelligence and Machine Learning in Pharmacy: A Narrative Review of Real-World Applications," *Pharmacy*, vol. 13, no. 2, p. 41, 2025, doi: 10.3390/pharmacy13020041.
- [4] S. Al-Hourani and D. Weraikat, "A Systematic Review of Artificial Intelligence (AI) and Machine Learning (ML) in Pharmaceutical Supply Chain (PSC) Resilience: Current Trends and Future Directions," *Sustain.*, vol. 17, no. 14, pp. 1–27, 2025, doi: 10.3390/su17146591.
- [5] S. S. Dhurwey, P. Kumar, U. Sendur, and A. Verma, *AI-Driven Inventory Management System for Pharmacy Stores*, no. Icaaa. Atlantis Press International BV, 2025. doi: 10.2991/978-94-6463-738-0_83.
- [6] R. K. Sridharan, A. Xhafa, S. Chaothari, and S. Putsala, "Automated Vial and Pre-Filled Syringe Counting in the Pharmaceutical Industry Using YOLO and SAHI Techniques," *Int. J. Innov. Sci. Res. Technol.*, vol. 9, no. 9, pp. 3127–3134, 2024, doi: 10.38124/ijisrt/IJISRT24SEP831.
- [7] S. Kang, Z. Hu, L. Liu, K. Zhang, and Z. Cao, "Object Detection YOLO Algorithms and Their Industrial Applications: Overview and Comparative Analysis," *Electron.*, vol. 14, no. 6, pp. 1–36, 2025, doi: 10.3390/electronics14061104.
- [8] N. A. Megantara and E. Utami, "Object Detection using YOLOv8 : A Systematic Review," *Sistemasi*, vol. 14, no. 3, p. 1186, 2025, doi: 10.32520/stmsi.v14i3.5081.
- [9] J. Lungido and E. H. Rachmawanto, "Real-Time Drug Classification Using YOLOv11 for Reducing Medication Errors," *J. Appl. Informatics Comput.*, vol. 9, no. 4, pp. 1191–1200, 2025, doi: 10.30871/jaic.v9i4.10117.
- [10] M. G. Ragab *et al.*, "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," *IEEE Access*, vol. 12, no. April, pp. 57815–57836, 2024, doi: 10.1109/ACCESS.2024.3386826.
- [11] K. Alomar, H. I. Aysel, and X. Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J. Imaging*, vol. 9, no. 2, 2023, doi: 10.3390/jimaging9020046.
- [12] J.-H. Oh, H.-G. Kim, and K. M. Lee, "Developing and evaluating deep learning algorithms for object detection: Key points for achieving superior model performance," *Korean Journal of Radiology*, vol. 24, no. 7, pp. 698–714, 2023, doi: 10.3348/kjr.2022.0765.
- [13] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A comparative analysis of object detection metrics with a companion open-source toolkit," *Electronics*, vol. 10, no. 3, Art. no. 279, 2021, doi: 10.3390/electronics10030279.
- [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, Art. no. 336, 2024, doi: 10.3390/computers13120336.
- [15] J. Sobek *et al.*, "MedYOLO: A medical image object detection framework," *Journal of Imaging Informatics in Medicine*, vol. 37, no. 6, pp. 3208–3216, 2024, doi: 10.1007/s10278-024-01138-2.
- [16] C. Albuquerque, R. Henriques, and M. Castelli, "Deep learning-based object detection algorithms in medical imaging: Systematic review," *Heliyon*, vol. 11, no. 1, Art. no. e41137, 2025, doi: 10.1016/j.heliyon.2024.e41137.
- [17] C. Chen, N. A. Mat Isa, and X. Liu, "A review of convolutional neural network based methods for medical image classification," *Computers in Biology and Medicine*, vol. 185, Art. no. 109507, 2025, doi: 10.1016/j.compbiomed.2024.109507.
- [18] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.