

YOLOV8-BASED VISITOR DENSITY DETECTION AND CLASSIFICATION SYSTEM IN ACCESSORY SHOPPING CENTERS (CASE STUDY: SUPER DAZZLE GEJAYAN)

Loorius Forbin Dapa^{1*}, Supatman²

Program Studi Informatika, Universitas Mercu Buana Yogyakarta, Indonesia
dapaforbin@gmail.com^{1*}, supatman@mercubuana-yogya.ac.id²

**Corresponding author*

Received June 30, 2026; Revised July 11, 2026; Accepted July 13, 2026; Published July 31, 2026

ABSTRACT

Visitor density monitoring in commercial areas directly affects safety and facility comfort management. This study develops a visitor density detection and classification system based on a pretrained YOLOv8s model, with a case study at Super Dazzle Gejayan, Yogyakarta. The YOLOv8s model was used without retraining, leveraging weights pretrained on the COCO dataset to detect human objects in real time. Detection results were then classified into three density categories: Quiet (≤ 5 people), Moderate (6–10 people), and Crowded (> 10 people), using a threshold-based rule approach. Testing was conducted on 707 static images collected directly from the store environment under various visitor conditions. The system achieved an overall accuracy of 89.39%. Based on the classification report, the Crowded class achieved the best performance with an F1-score of 0.94, followed by the Quiet class (F1-score 0.91), while the Moderate class obtained the lowest F1-score of 0.82 due to its position between the two other classes. The main limitations of the system include occlusion conditions, lighting variation, and threshold sensitivity. The results indicate that the pretrained approach without retraining saves time and computational resources for the training process, while still producing competitive accuracy for visitor density classification applications in real-world environments.

Keywords: YOLOv8, object detection, density classification, people counting, computer vision.

ABSTRAK

Pemantauan kepadatan pengunjung di area komersial berpengaruh langsung dalam manajemen keamanan dan kenyamanan fasilitas. Penelitian ini mengembangkan sistem deteksi dan klasifikasi kepadatan pengunjung berbasis YOLOv8s pretrained dengan studi kasus di Super Dazzle Gejayan, Yogyakarta. Model YOLOv8s digunakan tanpa proses pelatihan ulang dengan memanfaatkan bobot hasil pelatihan pada dataset COCO untuk mendeteksi objek manusia secara real-time. Hasil deteksi kemudian diklasifikasikan ke dalam tiga kategori kepadatan Sepi (≤ 5 orang), Sedang (6–10 orang), dan Ramai (> 10 orang) menggunakan pendekatan berbasis ambang batas (threshold-based rule). Pengujian dilakukan terhadap 707 citra statis yang dikumpulkan langsung dari lingkungan toko pada berbagai kondisi kunjungan. Sistem berhasil mencapai akurasi keseluruhan sebesar 89,39%. Berdasarkan classification report, kelas Ramai memperoleh performa terbaik dengan F1-score 0,94, diikuti kelas Sepi (F1-score 0,91), sedangkan kelas Sedang memperoleh F1-score terendah sebesar 0,82 akibat posisinya yang berada di antara dua kelas lainnya. Keterbatasan utama sistem meliputi kondisi occlusion, variasi pencahayaan, dan sensitivitas nilai ambang batas. Hasil penelitian menunjukkan bahwa pendekatan pretrained tanpa pelatihan ulang merupakan alternatif yang menghemat waktu dan sumber daya komputasi untuk proses pelatihan secara komputasi namun tetap menghasilkan akurasi yang kompetitif untuk aplikasi klasifikasi kepadatan pengunjung di lingkungan nyata.

Kata kunci: YOLOv8, deteksi objek, klasifikasi kepadatan, people counting, computer vision.

PENDAHULUAN

Super Dazzle Gejayan merupakan salah satu pusat penjualan aksesoris telepon genggam, komputer, dan berbagai perangkat elektronik yang telah lama beroperasi di Kota Yogyakarta. Sebagai perusahaan ritel yang melayani pelanggan secara langsung, aktivitas operasionalnya cukup tinggi dan didukung oleh sumber daya manusia yang berperan dalam berbagai proses bisnis perusahaan (Samboga & Subkhan, 2025). Seiring meningkatnya kebutuhan masyarakat terhadap perangkat elektronik, aktivitas pelayanan di lingkungan toko semakin kompleks sehingga diperlukan pengelolaan operasional yang dapat merespons perubahan kepadatan pengunjung secara tepat waktu dan menghemat sumber daya, termasuk dalam memantau tingkat kepadatan pengunjung yang berubah-ubah pada waktu tertentu.

Pemantauan kepadatan pengunjung (*crowd density monitoring*) merupakan aspek penting dalam manajemen keamanan dan kenyamanan pada ruang publik maupun area komersial. Seiring meningkatnya kebutuhan akan sistem pemantauan secara real-time, metode penghitungan manual dinilai kurang memadai karena memerlukan sumber daya manusia yang besar serta rentan terhadap kesalahan. Penelitian terbaru oleh Wijayanti et al. (2025) mengembangkan sistem pemantauan kerumunan secara real-time berbasis YOLOv8 dan pelacakan objek yang mampu mencapai akurasi sebesar 94,2% dalam menghitung pengunjung pada lingkungan wisata. Penelitian tersebut juga menerapkan pendekatan yang menjaga privasi dengan mendeteksi bagian kepala sebagai pengganti wajah, sehingga tetap memungkinkan proses pemantauan tanpa mengorbankan identitas individu. Hasil tersebut menunjukkan bahwa teknologi berbasis visi komputer memiliki potensi yang tinggi untuk diterapkan pada sistem penghitung pengunjung otomatis yang tetap memperhatikan aspek privasi pengguna. Ketidakmampuan mendeteksi tingkat kepadatan secara dini dapat menyebabkan menurunnya kenyamanan pengunjung, terganggunya aktivitas operasional, hingga meningkatnya risiko keselamatan. Oleh karena itu, dibutuhkan sistem yang mampu memberikan informasi jumlah orang secara cepat dan akurat agar pengelola dapat mengambil keputusan yang tepat dalam pengelolaan kerumunan dan optimalisasi layanan (Prihandoko et al., 2025).

Perkembangan *computer vision* sebagai bidang kecerdasan buatan yang memungkinkan komputer memproses dan memahami informasi visual secara otomatis. Dio et al. (2022) telah menghadirkan solusi untuk permasalahan ini. Teknologi ini memanfaatkan Convolutional Neural Network (CNN) yang mampu mengekstrak fitur visual seperti pola, bentuk, dan tekstur objek secara otomatis (Mihuandayani et al., 2025). Salah satu implementasi terkemuka adalah algoritma You Only Look Once (YOLO), yang menggunakan pendekatan one-stage detector sehingga lokalisasi dan klasifikasi objek dilakukan dalam satu kali inferensi, menghasilkan kecepatan deteksi tinggi yang cocok untuk pemrosesan real-time (Sejak Terang Waruwu & Sipayung, 2026). Versi terbarunya, YOLOv8 yang dikembangkan oleh Ultralytics, menghadirkan peningkatan pada arsitektur backbone, neck, dan detection head sehingga lebih unggul dari versi sebelumnya dalam hal akurasi maupun kebutuhan sumber daya komputasi yang lebih rendah (Dwi Grevika Drantantiyas et al., 2023; Mahardhika & Pakereng, 2026;

Mihuandayani et al., 2025). Berbagai penelitian telah membuktikan efektivitasnya: implementasi YOLOv8 pada kampus STMIK Multicom mencapai akurasi 85% (Mihuandayani et al., 2025), sementara penerapannya di Stasiun Bekasi menghasilkan mAP@0.5 sebesar 91,7% dan F1-score 87% (Prihandoko et al., 2025). Dengan temuan serupa dikonfirmasi oleh Kusuma et al. (2021) dan Setiyadi et al. (2023).

Meskipun demikian, sebagian besar penelitian terdahulu masih berfokus pada estimasi jumlah orang atau visualisasi kepadatan secara kontinu, belum menghasilkan kategori kepadatan yang mudah dipahami oleh pengelola fasilitas (Aldo Dhavicky et al., 2025). Penelitian mengenai klasifikasi tingkat kepadatan berbasis *people counting* pada lingkungan pusat perbelanjaan aksesoris elektronik di Indonesia juga masih sangat terbatas, mengingat sebagian besar penelitian sebelumnya berfokus pada lokasi stasiun, terminal, maupun lingkungan kampus (Kusuma et al., 2021; Prihandoko et al., 2025).

Kebaruan penelitian ini terletak pada tiga hal. Pertama, penelitian ini menerapkan model YOLOv8s *pretrained* tanpa pelatihan ulang yang membutuhkan sumber daya komputasi besar seperti pada metode berbasis CNN yang memerlukan 2000 *epoch* selama tujuh hari (Aldo Dhavicky et al., 2025), sehingga pendekatan ini jauh lebih baik dan dapat dijalankan pada platform komputasi sederhana seperti Google Colaboratory. Kedua, hasil deteksi diklasifikasikan ke dalam kategori diskrit yaitu Sepi, Sedang, dan Ramai yang lebih mudah dipahami oleh pengelola fasilitas dibandingkan keluaran berupa *density map* kontinu yang memerlukan interpretasi teknis lebih lanjut (Aldo Dhavicky et al., 2025). Ketiga, pengujian dilakukan langsung pada lingkungan nyata Super Dazzle Gejayan, Yogyakarta, yang memiliki karakteristik tata ruang dan pola kunjungan berbeda dibandingkan stasiun maupun kampus yang menjadi fokus penelitian sebelumnya (Kusuma et al., 2021; Prihandoko et al., 2025; Samboga & Subkhan, 2025).

Berdasarkan kondisi tersebut, penelitian ini mengembangkan sistem klasifikasi kepadatan pengunjung menggunakan YOLOv8s *pretrained* yang mampu mengklasifikasikan hasil deteksi ke dalam tiga kategori yaitu Sepi (≤ 5 orang), Sedang (6–10 orang), dan Ramai (> 10 orang) tanpa memerlukan proses pelatihan ulang, dan diuji langsung pada lingkungan nyata Super Dazzle Gejayan, Yogyakarta. Sedangkan pada penelitian yang dilakukan oleh Aldo Dhavicky et al. (2025) dalam penelitiannya masih memerlukan pelatihan data yang terbilang lama, memakan waktu pelatihan hingga 7 hari dengan 2000 epochs. Sistem ditargetkan mampu mencapai akurasi minimal 80% dengan nilai F1-score yang seimbang pada ketiga kelas kepadatan, sehingga dapat menjadi dasar pengembangan sistem pemantauan kepadatan pengunjung yang praktis dan andal di lingkungan pusat perbelanjaan lokal Indonesia. Oleh karena itu, penelitian ini dibuat untuk mengembangkan kekurangan penelitian sebelumnya untuk mengklasifikasikan tingkat kepadatan penduduk berdasarkan objek yang terdeteksi.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *computer vision* berbasis deteksi objek untuk mengklasifikasikan tingkat kepadatan pengunjung berdasarkan jumlah orang yang terdeteksi dalam citra digital. Pendekatan kuantitatif

dipilih karena seluruh proses penilaian performa sistem didasarkan pada data numerik yang terukur dan dapat direplikasi, meliputi nilai akurasi, *precision*, *recall*, dan *F1-score*, sehingga hasil penelitian dapat dibandingkan secara objektif dengan penelitian lain (Mihuandayani et al., 2025; Setiyadi et al., 2023). Deteksi objek merupakan bagian dari *computer vision* yang tidak hanya mengenali jenis objek, tetapi juga menentukan lokasi objek dalam bentuk kotak pembatas (*bounding box*), sehingga cocok digunakan untuk memantau keberadaan manusia di suatu area sekaligus menghitung jumlahnya secara akurat sebagai dasar klasifikasi kepadatan (Dio et al., 2022; Indra et al., 2023; Prihandoko et al., 2025).

Metode yang digunakan adalah YOLOv8s, varian *small* dari model *You Only Look Once* versi 8 yang dikembangkan oleh Ultralytics. YOLOv8s dipilih karena menggunakan pendekatan *one-stage detector* yang memproses deteksi dan klasifikasi objek dalam satu langkah sekaligus, menjadikannya lebih cepat dan memadai secara komputasi dibandingkan metode *two-stage detector* seperti Faster R-CNN (Dwi Grevika Drantantiyas et al., 2023). Selain itu, varian *small* dipilih karena menyeimbangkan antara kecepatan inferensi dan akurasi deteksi tanpa membutuhkan spesifikasi perangkat keras yang tinggi, sehingga sesuai dengan sumber daya komputasi yang tersedia pada platform Google Colaboratory (Setiyadi et al., 2023). Model ini telah dilatih menggunakan dataset COCO yang mencakup 80 kelas objek termasuk kelas *person*, sehingga mampu mendeteksi objek manusia secara langsung tanpa memerlukan proses pelatihan ulang yang membutuhkan sumber daya komputasi besar (Aldo Dhavicky et al., 2025).

Secara keseluruhan, penelitian ini dilakukan dalam empat tahap utama, yaitu mengumpulkan dataset uji, mendeteksi orang menggunakan YOLOv8s, mengklasifikasikan tingkat kepadatan berdasarkan jumlah orang yang terdeteksi, serta mengevaluasi kinerja sistem menggunakan metrik akurasi, *confusion matrix*, *precision*, *recall*, dan *F1-score*. Alur penelitian yang digunakan dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

Pengumpulan Dataset Test

Dataset yang digunakan dalam penelitian ini terdiri dari citra statis yang diperoleh dari kamera CCTV yang terpasang di lingkungan Super Dazzle Gejayan, Yogyakarta. Berdasarkan penelitian yang dilakukan pada toko ini, Super Dazzle Gejayan memiliki aktivitas operasional yang tinggi dengan melibatkan banyak sumber daya manusia dan pengunjung, sehingga menjadikannya lokasi yang representatif untuk pengujian sistem klasifikasi kepadatan pada berbagai kondisi kunjungan yang berbeda (Samboga & Subkhan, 2025).

Pengambilan data dilakukan selama dua minggu dengan memanfaatkan kamera CCTV yang sudah terpasang di area dalam toko. Citra diperoleh dengan cara melakukan tangkapan layar (*screenshot*) secara manual dari aplikasi pemantau CCTV setiap harinya mulai dari jam buka hingga jam tutup toko, sehingga data yang diperoleh mencakup berbagai kondisi kepadatan secara alami sepanjang jam operasional, mulai dari kondisi sepi pada jam-jam tertentu hingga kondisi ramai pada jam sibuk kunjungan. Pendekatan pengambilan data melalui CCTV dipilih karena lebih merepresentasikan kondisi nyata toko dibandingkan pengambilan foto secara langsung, serta mengurangi potensi bias akibat kehadiran pengambil gambar di area toko.

Jumlah total dataset yang digunakan adalah 707 citra yang terdistribusi ke dalam tiga kelas kepadatan. Penentuan jumlah citra didasarkan pada hasil seleksi dari keseluruhan tangkapan layar yang dikumpulkan selama periode pengambilan data dua minggu, dengan mempertimbangkan kualitas citra dan keterwakilan setiap kelas secara proporsional. Citra yang buram, terlalu gelap, atau tidak menampilkan area toko secara jelas tidak diikutsertakan dalam dataset. Distribusi dataset per kelas dapat dilihat pada Tabel 1 berikut.

Tabel 1. Distribusi Dataset per Kelas

Kelas	Jumlah Orang	Jumlah Citra
Sepi	≤ 5 orang	204 citra
Sedang	6 – 10 orang	205 citra
Ramai	> 10 orang	298 citra
Total		707 citra

Format citra yang digunakan adalah *.jpg*, *.jpeg*, dan *.png*. Pelabelan kelas dilakukan secara manual berdasarkan jumlah orang yang tampak pada setiap citra sebelum proses deteksi dilakukan, di mana nama folder secara langsung berfungsi sebagai *ground truth* atau label asli dalam proses evaluasi sistem. Setelah pelabelan, setiap citra kemudian dibaca menggunakan fungsi *cv2.imread()* dari pustaka OpenCV yang mengubah file gambar menjadi array numerik tiga dimensi sebelum diproses lebih lanjut oleh model.

Model Deteksi Objek

Algoritma YOLO menggunakan pendekatan *one-stage detector* yang memungkinkan deteksi dan klasifikasi objek dilakukan dalam satu kali inferensi,

sehingga jauh lebih cepat dibandingkan metode deteksi objek lainnya dan cocok untuk aplikasi yang membutuhkan efisiensi tinggi (Indra et al., 2023; Sejak Terang Waruwu & Sipayung, 2026). Dalam penelitian ini, model yang digunakan adalah YOLOv8s (small) dari Ultralytics, dipilih karena menyeimbangkan antara kecepatan inferensi dan akurasi deteksi tanpa membutuhkan spesifikasi perangkat keras yang tinggi (Mihuandayani et al., 2025). Model digunakan dalam mode pretrained dengan bobot hasil pelatihan pada dataset COCO yang mencakup 80 kelas objek termasuk kelas person pada indeks ke-0, sehingga tidak diperlukan proses pelatihan ulang yang membutuhkan sumber daya komputasi besar (Aldo Dhavicky et al., 2025). Nilai confidence threshold sebesar 0,3 dipilih untuk meningkatkan sensitivitas deteksi, khususnya pada kondisi pengunjung yang saling berdekatan atau sebagian tubuhnya tertutup oleh objek lain (Kusuma et al., 2021; Prihandoko et al., 2025).

Penghitungan Orang (*People Counting*)

Deteksi manusia merupakan penerapan deteksi objek yang bertujuan mengidentifikasi keberadaan orang dalam citra, di mana setiap orang yang terdeteksi ditandai dengan kotak pembatas dan informasi jumlahnya dapat dimanfaatkan untuk pemantauan pengunjung maupun pengendalian kerumunan (Dwi Grevika Drantantiyas et al., 2023; Mihuandayani et al., 2025). Dalam penelitian ini, *people counting* dilakukan dengan menghitung jumlah *bounding box* yang terklasifikasi sebagai kelas *person* (cls=0) dari hasil prediksi YOLOv8s pada setiap citra (Kusuma et al., 2021; Setiyadi et al., 2023). Nilai *person_count* yang dihasilkan kemudian menjadi masukan bagi fungsi klasifikasi kepadatan pada tahap berikutnya.

Klasifikasi Kepadatan

Klasifikasi tingkat kepadatan adalah proses mengelompokkan kondisi suatu area berdasarkan jumlah orang yang terdeteksi, dan informasi ini penting untuk mendukung pengambilan keputusan terkait pengelolaan pengunjung dan pengaturan kapasitas ruangan (Dwi Grevika Drantantiyas et al., 2023; Mahardhika & Pakereng, 2026; Mihuandayani et al., 2025). Hasil penghitungan orang diklasifikasikan ke dalam tiga kategori menggunakan fungsi berbasis aturan ambang batas (*threshold-based rule*). Pendekatan ini dipilih karena tidak memerlukan pelatihan model tambahan serta waktu banyak dan menghasilkan keluaran kategori yang mudah diinterpretasikan secara langsung oleh pengelola fasilitas (Aldo Dhavicky et al., 2025; Prihandoko et al., 2025). Penetapan ambang batas didasarkan pada pertimbangan kapasitas ruang dan karakteristik kunjungan di lingkungan Super Dazzle Gejayan , dengan aturan sebagai berikut:

$$\begin{cases} \text{sepi jika } n \leq 5 \\ \text{sedang jika } 6 \leq n \leq 10 \\ \text{ramai jika } n > 10 \end{cases}$$

Evaluasi Sistem

Evaluasi performa sistem dilakukan dengan membandingkan label prediksi terhadap label aktual dari seluruh citra uji pada ketiga kelas kepadatan. Metrik evaluasi yang digunakan mengacu pada pendekatan yang umum diterapkan dalam penelitian klasifikasi berbasis *computer vision* (Kusuma et al., 2021; Mihuandayani et al., 2025; Prihandoko et al., 2025; Setiyadi et al., 2023). Computer Vision adalah kemampuan sebuah mesin atau komputer untuk melihat atau mengenali gambar, dengan kemampuan yang sama atau bahkan lebih baik daripada penglihatan manusia (Mulya et al., 2023).

Akurasi

$$Akurasi = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Data Uji}} \times 100\%$$

Confusion Matrix

Confusion matrix digunakan untuk memvisualisasikan distribusi hasil prediksi benar dan salah pada masing-masing kelas (Aldo Dhavicky et al., 2025). Matriks berukuran 3×3 dihasilkan menggunakan fungsi `confusion_matrix` dari pustaka *Scikit-learn* dengan urutan label ["Sepi", "Sedang", "Ramai"].

Classification Report

Classification report digunakan untuk mengevaluasi kinerja model klasifikasi berdasarkan tiga metrik utama, yaitu precision, recall, dan F1-score. Precision menunjukkan tingkat ketepatan prediksi pada suatu kelas, *recall* menunjukkan kemampuan model dalam mengidentifikasi data yang benar pada kelas tersebut, sedangkan *F1-score* merupakan kombinasi antara precision dan recall yang digunakan untuk memberikan gambaran kinerja model secara keseluruhan (Kusuma et al., 2021; Mihuandayani et al., 2025).

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{precision} + \text{Recall}}$$

HASIL DAN PEMBAHASAN

Hasil

Hasil Deteksi YOLOv8s

Proses deteksi objek dilakukan menggunakan model YOLOv8s *pretrained* pada seluruh citra uji yang terdapat pada ketiga folder kelas kepadatan. Model berhasil mendeteksi objek manusia pada setiap citra dengan menandai setiap individu menggunakan *bounding box* beserta nilai *confidence* masing-masing. Gambar di bawah menunjukkan contoh hasil deteksi pada sampel citra dari kelas Sedang, sepi dan Ramai yang dilengkapi dengan informasi jumlah orang terdeteksi dan kategori kepadatan hasil klasifik



Gambar 2. Citra Ramai



Gambar 3. Citra Sedang



Gambar 4. Citra Sepi

Berdasarkan hasil deteksi yang diperoleh, model mampu mengidentifikasi objek manusia dengan cukup baik pada kondisi pencahayaan yang memadai. Namun pada beberapa citra dengan kondisi pengunjung yang saling berdekatan atau tertutup sebagian (*occlusion*), terdapat beberapa objek yang tidak terdeteksi sehingga memengaruhi jumlah orang yang terhitung.

Hasil Penghitungan dan Klasifikasi Per Citra

Seluruh hasil penghitungan orang dan klasifikasi kepadatan per citra disimpan dalam file `hasil_testing.csv`. Tabel 2 menampilkan sebagian sampel hasil pengujian yang mencakup gambar, label asli, label prediksi, dan jumlah orang yang terdeteksi oleh model.

Gambar	Label Asli	Prediksi	Jumlah Orang
--------	------------	----------	--------------

	Ramai	Ramai	23
	Ramai	Ramai	24

Dari tabel 2 tersebut dapat dilihat bahwa sebagian besar prediksi sistem sesuai dengan label asli yang telah ditetapkan secara manual. Kesalahan prediksi umumnya terjadi pada citra yang jumlah orangnya berada di sekitar nilai ambang batas antar kelas, misalnya citra dengan 5–6 orang yang berada di batas antara kelas Sepi dan Sedang, sehingga sedikit perbedaan jumlah deteksi dapat mengubah hasil klasifikasi.

Hasil Akurasi

Hasil evaluasi akurasi keseluruhan sistem ditampilkan pada Gambar 5. Sistem berhasil mencapai akurasi sebesar 89,39% dari total 707 citra uji yang terdiri dari tiga kelas kepadatan. Nilai akurasi ini menunjukkan bahwa model YOLOv8s *pretrained* tanpa pelatihan ulang mampu mengklasifikasikan kepadatan pengunjung dengan baik di lingkungan nyata Super Dazzle Gejayan. Hasil ini lebih tinggi dibandingkan penelitian penghitungan pengunjung berbasis YOLOv8 di lingkungan kampus yang mencapai akurasi 85% (Mihuandayani et al., 2025), bahkan melampaui sistem integrasi YOLOv8 dan *Streamlit* di Stasiun Bekasi yang memperoleh akurasi rata-rata 80,37% (Prihandoko et al., 2025).

Nilai akurasi yang dihasilkan dari penelitian ini dipengaruhi oleh beberapa faktor. Dari sisi positif, model YOLOv8s yang telah dilatih pada dataset COCO memiliki kemampuan generalisasi yang baik karena mencakup ribuan gambar manusia dalam berbagai kondisi lingkungan, sehingga mampu beradaptasi pada kondisi toko yang belum pernah dilihat sebelumnya. Selain itu, penetapan nilai *confidence threshold* sebesar 0,3 turut meningkatkan sensitivitas deteksi sehingga objek manusia yang terdeteksi dengan keyakinan rendah tetap dihitung, mengurangi risiko *undercounting* terutama pada kondisi pengunjung yang saling berdekatan. Pendekatan klasifikasi berbasis ambang batas yang digunakan juga tidak bergantung pada model tambahan, sehingga tidak ada potensi kesalahan dari proses klasifikasi sekunder dan keputusan kategori didasarkan langsung pada nilai numerik yang objektif.

Di sisi lain, terdapat beberapa faktor yang membatasi capaian akurasi sistem. Kondisi *occlusion* pada kepadatan tinggi menyebabkan sebagian tubuh manusia tidak terdeteksi model sehingga jumlah orang yang terhitung lebih rendah dari kondisi sebenarnya. Variasi pencahayaan di dalam ruang toko yang tidak merata juga memengaruhi konsistensi deteksi, terutama pada area yang kurang mendapat cahaya. Selain itu, sensitivitas nilai ambang batas pada citra yang jumlah orangnya berada tepat di batas antar kelas menjadi penyebab utama kesalahan klasifikasi, khususnya pada kelas Sedang yang memperoleh F1-score terendah sebesar 0,82 dari ketiga kelas yang diuji.

```

...
=====
AKURASI
=====
89.39%

```

Gambar 5. Hasil Akurasi

Dengan perhitungan menggunakan rumus akurasi sebagai berikut;

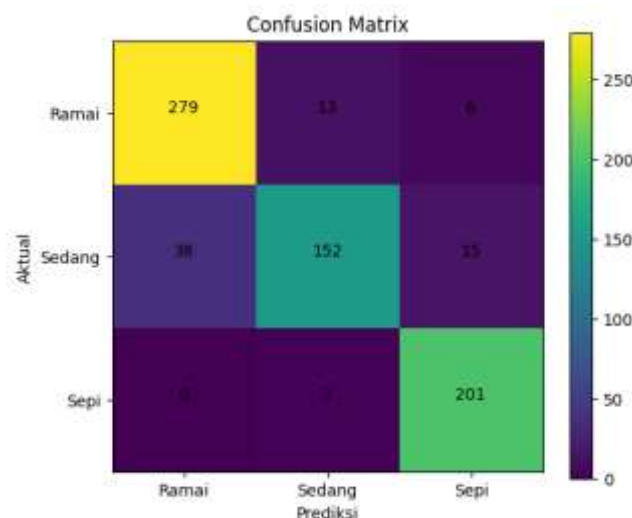
$$Akurasi = \frac{201 + 152 + 279}{707} \times 100\%$$

$$\frac{632}{707} \times 100\% = 89,39\%$$

Hasil perhitungan manual tersebut sesuai dengan nilai akurasi yang dihasilkan oleh fungsi *accuracy_score* dari pustaka *Scikit-learn*, sehingga membuktikan bahwa proses evaluasi sistem berjalan dengan benar dan dapat dipertanggungjawabkan

Confusion Matrix

Gambar 6 menampilkan *confusion matrix* hasil pengujian sistem pada ketiga kelas kepadatan terhadap 707 citra uji.



Gambar 6. Confusion Matrix

Berdasarkan *confusion matrix*, kelas Sepi memiliki tingkat kesalahan yang paling rendah dengan 201 citra berhasil diklasifikasikan dengan benar dan hanya 3 citra yang salah diprediksi sebagai kelas Sedang. Kelas Sedang menunjukkan tingkat kesalahan tertinggi, di mana dari 205 citra hanya 152 citra yang terklasifikasi dengan benar, sedangkan sisanya salah diprediksi sebagai Sepi maupun Ramai. Hal ini menunjukkan bahwa kelas Sedang lebih sulit dibedakan karena berada di antara dua tingkat kepadatan lainnya. Sementara itu, kelas Ramai memiliki performa yang sangat baik dengan 279 citra berhasil diprediksi dengan benar dari total 298 citra, sedangkan kesalahan yang terjadi umumnya disebabkan oleh kondisi occlusion atau objek yang saling menutupi sehingga jumlah orang yang terdeteksi lebih sedikit dari kondisi sebenarnya. Secara keseluruhan, hasil *confusion matrix* menunjukkan bahwa model mampu mengklasifikasikan kelas Sepi dan Ramai dengan baik, namun masih mengalami kesulitan dalam membedakan kelas Sedang.

Classification Report

Gambar 7 menampilkan *classification report* yang memuat nilai *precision*, *recall*, dan *F1-score* untuk masing-masing kelas kepadatan.

	precision	recall	f1-score	support
Ramai	0.88	0.94	0.91	298
Sedang	0.90	0.74	0.82	205
Sepi	0.91	0.99	0.94	204
accuracy			0.89	707
macro avg	0.90	0.89	0.89	707
weighted avg	0.89	0.89	0.89	707

Gambar 7. Hasil *Classification Report*

Berdasarkan hasil *classification report*, sistem menunjukkan performa yang baik dalam mengklasifikasikan tingkat kepadatan pengunjung pada ketiga kelas yang diuji. Interpretasi nilai *precision*, *recall*, dan *F1-score* pada masing-masing kelas diuraikan sebagai berikut.

Kelas Ramai memperoleh nilai *precision* 0,88, *recall* 0,94, dan *F1-score* 0,91. Nilai *recall* sebesar 0,94 menunjukkan bahwa sistem mampu mengidentifikasi 94% dari seluruh citra yang sebenarnya berkelas Ramai dengan benar, artinya sistem jarang melewatkan kondisi ramai yang sesungguhnya hal ini penting secara praktis karena kondisi Ramai yang tidak terdeteksi dapat berdampak langsung pada keselamatan dan kenyamanan pengunjung di dalam toko. Nilai *precision* sebesar 0,88 menunjukkan bahwa dari seluruh citra yang diprediksi sebagai Ramai, 88% di antaranya memang benar berkelas Ramai, sedangkan 12% sisanya merupakan citra dari kelas lain (kemungkinan

Sedang) yang ikut terprediksi sebagai Ramai akibat jumlah orang yang berada tepat di sekitar batas ambang.

Kelas Sedang memperoleh nilai precision 0,90, recall 0,74, dan F1-score 0,82, menjadikannya kelas dengan performa terendah di antara ketiga kelas. Nilai recall sebesar 0,74 menunjukkan bahwa sistem hanya berhasil mengidentifikasi 74% dari seluruh citra berkelas Sedang dengan benar, artinya 26% citra Sedang masih terklasifikasi ke kelas lain. Kondisi ini terjadi karena kelas Sedang berada di antara dua kelas lainnya, sehingga kesalahan deteksi satu hingga dua orang akibat *occlusion* sudah cukup mengubah hasil klasifikasi ke kelas Sepi maupun Ramai. Meskipun demikian, nilai precision yang cukup tinggi sebesar 0,90 menunjukkan bahwa ketika sistem memprediksi suatu citra sebagai Sedang, prediksi tersebut sebagian besar benar.

Kelas Sepi memperoleh nilai precision 0,91, recall 0,99, dan F1-score 0,94, menjadikannya kelas dengan performa terbaik. Nilai recall sebesar 0,99 menunjukkan bahwa sistem hampir selalu berhasil mengidentifikasi kondisi sepi dengan benar, artinya sistem sangat jarang salah mengklasifikasikan citra Sepi ke kelas lain. Nilai precision sebesar 0,91 menunjukkan bahwa dari seluruh citra yang diprediksi sebagai Sepi, 91% di antaranya memang benar berkelas Sepi, sehingga kesalahan klasifikasi pada kelas ini relatif kecil dan tidak signifikan memengaruhi keandalan sistem secara keseluruhan. Secara keseluruhan sistem memperoleh nilai *macro average* (precision 0,90; recall 0,89; F1-score 0,89) dan *weighted average* (0,89 pada ketiga metrik) yang konsisten pada kisaran 0,89–0,90. Konsistensi antara kedua nilai rata-rata tersebut menunjukkan bahwa performa sistem tidak hanya unggul pada satu kelas tertentu, melainkan merata di seluruh kelas kepadatan meskipun terdapat ketidakseimbangan jumlah data antar kelas. Hal ini membuktikan bahwa sistem memiliki kemampuan klasifikasi yang seimbang dan dapat diandalkan untuk diterapkan pada kondisi nyata di lingkungan Super Dazzle Gejayan (Dwi Grevika Drantantiyas et al., 2023; Mihuandayani et al., 2025).

Pembahasan

Secara keseluruhan sistem yang dikembangkan menunjukkan performa yang baik dengan akurasi 89,39% pada 707 citra uji. Kelas Ramai memiliki performa terbaik dengan F1-score 0,94, diikuti kelas Sepi dengan F1-score 0,91, sedangkan kelas Sedang menjadi kelas dengan performa terendah dengan F1-score 0,82. Rendahnya performa pada kelas Sedang dapat dijelaskan oleh posisinya yang berada di antara dua kelas lainnya, sehingga kesalahan deteksi satu hingga dua orang saja sudah cukup untuk mengubah hasil klasifikasi ke kelas yang berdekatan. Hal serupa juga ditemukan pada penelitian klasifikasi tingkat kepadatan kerumunan berbasis XGBoost pada konteks ibadah massal (Illaihiyah et al., 2026), yang menunjukkan bahwa kelas dengan posisi menengah pada skala rendah–sedang–tinggi secara konsisten lebih sulit dipisahkan dibandingkan kelas pada dua ujung skala meskipun akurasi keseluruhan modelnya tinggi.

Dari sisi efisiensi komputasi, pendekatan pretrained tanpa pelatihan ulang yang diusulkan dalam penelitian ini terbukti lebih baik dibandingkan metode berbasis CNN yang membutuhkan proses pelatihan hingga 2000 epoch selama tujuh hari (Aldo

Dhavicky et al., 2025) untuk mengestimasi jumlah dan kepadatan orang pada citra kerumunan tunggal. Perbandingan ini juga sejalan dengan penelitian penghitungan pengunjung objek wisata menggunakan MobileNet-SSD (Saputra et al., 2022), yang memanfaatkan model deep learning yang telah dilatih sebelumnya (pretrained) tanpa membangun ulang arsitektur jaringan dari nol, sehingga sistem dapat langsung diterapkan pada kasus penghitungan orang secara real-time dengan beban komputasi yang jauh lebih ringan dibandingkan pelatihan CNN penuh. Dengan demikian, meskipun pendekatan berbasis pelatihan ulang seperti pada Aldo Dhavicky et al. (2025) berpotensi menghasilkan model yang lebih spesifik terhadap karakteristik dataset, pendekatan pretrained pada penelitian ini menawarkan keseimbangan yang lebih baik antara efisiensi waktu pengembangan dan akurasi yang tetap kompetitif, terutama untuk kebutuhan implementasi pada skala kecil hingga menengah seperti pemantauan kepadatan di dalam ruang toko.

Dari sisi manfaat praktis, pendekatan pretrained tanpa pelatihan ulang ini membuka peluang implementasi yang lebih luas pada berbagai fasilitas publik selain ruang toko, seperti terminal, stasiun, area ibadah, taman kota, maupun objek wisata, karena sistem tidak memerlukan proses pelatihan model dari nol setiap kali diterapkan pada lokasi baru dengan karakteristik lingkungan yang berbeda. Peluang ini didukung oleh penelitian terdahulu yang menunjukkan bahwa model pretrained mampu menghitung jumlah pengunjung yang keluar-masuk secara otomatis di objek wisata guna mendukung pembatasan kapasitas pengunjung (Saputra et al., 2022), serta terbukti efektif dalam meningkatkan efisiensi dan akurasi pemantauan jumlah pengunjung di ruang publik (Mihundayani et al., 2025). Hal ini menjadikan sistem lebih hemat biaya dan waktu pengembangan dibandingkan pendekatan berbasis pelatihan penuh, sehingga lebih realistis diadopsi oleh pengelola fasilitas dengan sumber daya teknis dan infrastruktur komputasi yang terbatas. Selain itu, kemampuan sistem dalam mengklasifikasikan kepadatan ke dalam kategori yang mudah dipahami (sepi, sedang, ramai) berpotensi mendukung pengambilan keputusan operasional secara cepat, misalnya untuk pengaturan jumlah pengunjung, penjadwalan petugas keamanan, atau pemberian notifikasi dini kepada pengelola ketika kepadatan mendekati kondisi ramai, sebagai langkah mitigasi terhadap risiko penumpukan pengunjung (overcrowding) yang dapat mengganggu kenyamanan maupun keselamatan pada fasilitas publik secara lebih luas.

Keterbatasan utama sistem ini terletak pada kondisi occlusion pada kepadatan tinggi, variasi pencahayaan di dalam ruang toko, serta sensitivitas nilai ambang batas klasifikasi terhadap citra yang jumlah orangnya berada tepat di batas antar kelas. Tantangan occlusion dan variasi pencahayaan ini juga dilaporkan pada penelitian pemantauan kerumunan wisatawan berbasis YOLOv8 dengan deteksi kepala (Wijayanti et al., 2025), yang secara khusus memilih pendekatan deteksi kepala alih-alih deteksi seluruh tubuh karena kepala relatif lebih tahan terhadap oklusi antarindividu pada kerumunan padat, meskipun sistem tersebut tetap dilatih pada lebih dari 3.000 frame beranotasi untuk mengakomodasi variasi pencahayaan dan sudut pandang yang beragam. Perbandingan ini menunjukkan bahwa strategi pemilihan objek deteksi (mis. kepala vs. seluruh tubuh) dapat menjadi salah satu arah pengembangan lanjutan untuk mengurangi dampak occlusion pada sistem yang dikembangkan dalam penelitian ini.

KESIMPULAN

Penelitian ini berhasil membuktikan bahwa model YOLOv8s *pretrained* tanpa pelatihan ulang mampu diterapkan untuk mengklasifikasikan kepadatan pengunjung di lingkungan pusat perbelanjaan nyata Super Dazzle Gejayan, Yogyakarta, dengan akurasi sebesar 89,39% dari 707 citra uji. Pendekatan klasifikasi berbasis ambang batas jumlah orang yang terdeteksi terbukti menghasilkan kategori kepadatan yang praktis dan mudah diinterpretasikan, dengan kelas Ramai mencapai performa terbaik (F1-score 0,94), diikuti kelas Sepi (F1-score 0,91), sementara kelas Sedang memperoleh F1-score terendah sebesar 0,82 yang disebabkan oleh sensitivitas nilai ambang batas dan kondisi *occlusion* pada citra dengan jumlah pengunjung di sekitar batas antar kelas. Untuk penelitian selanjutnya, disarankan agar dilakukan *fine-tuning* model pada dataset lokal yang lebih beragam guna meningkatkan akurasi deteksi pada kondisi *occlusion* dan pencahayaan rendah, serta mengkaji penetapan nilai ambang batas menggunakan pendekatan berbasis data seperti *clustering* agar lebih adaptif terhadap karakteristik lokasi yang berbeda, dan mengembangkan sistem secara *real-time* menggunakan kamera CCTV untuk penerapan langsung di lingkungan komersial lokal Indonesia.

REFERENSI

- Aldo Dhavicky, D., Fuad Rachmadi, R., & Mardi Susiki Nugroho, S. (2025). Estimasi Jumlah dan Kepadatan Orang Pada Citra Kerumunan Tunggal Menggunakan Metode Convolutional Neural Network (CNN). *Journal of Computer Engineering, Network, and Intelligent Multimedia*, 3(2), 93–103. <https://doi.org/10.59378/jcenim.v3i2.74>
- Dio, M. D. R. P., Bayu Priyatna, B. P., April Lia Hananto, A. L. H., & Shofa Shofiah Hilabi, S. S. H. (2022). Deteksi Objek Kecelakaan Pada Kendaraan Roda Empat Menggunakan Algoritma YOLOv5. *Teknologi: Jurnal Ilmiah Sistem Informasi*, 12(2), 15–26. <https://doi.org/10.26594/teknologi.v12i2.3260>
- Dwi Grevika Drantantiyas, N., Yulita, W., Taufiq Ridwan, N., Arta Ramadhani, U., Indra Kesuma, R., Zahri Rakhman, A., Bagaskara, R., Miranto, A., & Mufidah, Z. (2023). Performasi Deteksi Jumlah Manusia Menggunakan YOLOv8. *Jurnal Aplikasi Sains, Informasi, Elektronika dan Komputer*, 5(2), 63–68. <https://doi.org/10.26905/jasiek.v5i2.11605>
- Illaihiyah, R., Abdurrahman, G., & Daryanto. (2026). Implementasi Algoritma XGBoost Classifier untuk Klasifikasi Kerumunan Jemaah Haji dan Umrah. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 6(1), 106–115. <https://doi.org/10.57152/malcom.v6i1.2245>
- Indra, D., Herman, H., & Budi, F. S. (2023). Implementasi Sistem Penghitung Kendaraan Otomatis Berbasis Computer Vision. *Komputika: Jurnal Sistem Komputer*, 12(1), 53–62. <https://doi.org/10.34010/komputika.v12i1.9082>
- Kusuma, T. A. A. H., Usman, K., & Saidah, S. (2021). People Counting for Public Transportations Using You Only Look Once Method. *Jurnal Teknik Informatika (Jutif)*, 2(1), 57–66. <https://doi.org/10.20884/1.jutif.2021.2.2.77>

- Mahardhika, D. S., & Pakereng, M. A. I. (2026). Deteksi Objek Secara Real-Time Berbasis YOLOv8 dan Algoritma DeepSORT. *Jurnal Media Computer Science*, 5(1), 671–686. <https://doi.org/10.37676/jmcs.v5i1.10038>
- Mihuandayani, N., Helmi, W., Pesik, S., Mamuaya, S., Putra, A. N., & Mulyaman, E. (2025). Implementasi YOLOv8 untuk Deteksi dan Hitung Objek Manusia dengan Convolutional Neural. *JURNAL BUFFER INFORMATIKA*, 11(1), 27–35. <https://doi.org/10.25134/buffer.v11i1.356>
- Mulya, M. A., Zaenul Arif, & Syefudin. (2023). Tinjauan Pustaka Sistematis: Penerapan Metode Gabor Wavelet Pada Computer Vision. *Journal Of Computer Science and Technology (JOCSTEC)*, 1(2), 83–88. <https://doi.org/10.59435/jocstec.v1i2.78>
- Prihandoko, P., Rumapea, S. A., & Pratama, A. H. (2025). Integrasi Algoritma YOLOv8 dan Streamlit untuk Visualisasi Real-Time dan Akurat dalam Penghitungan Kerumunan di Kawasan Stasiun Bekasi. *METHOMIKA Jurnal Manajemen Informatika Dan Komputerisasi Akuntansi*, 9(1), 179–187. <https://doi.org/10.46880/jmika>.
- Samboga, R., & Subkhan, M. (2025). Pengaruh Kompensasi dan Lingkungan Kerja terhadap Kinerja Karyawan dengan Kepuasan Kerja Sebagai Variabel Mediasi pada Super Dazzle Gejayan. *EBISNIS (Jurnal Ilmiah Ekonomi Dan Bisnis)*, 18(1), 49–62. <https://doi.org/10.51903/e-bisnis.v18i1.2253>
- Saputra, M. H., Erwanto, D., & Rizal, R. F. (2022). Penghitung Jumlah Pengunjung Objek Wisata dengan Metode Deep Learning MobileNet-SSD. *Techné Jurnal Ilmiah Elektroteknika*, 21(2), 145–154. <https://doi.org/10.31358/techne.v21i2.313>
- Sejak Terang Waruwu, N., & Sipayung, Y. R. (2026). Prediksi Tingkat Kepadatan Kendaraan Menggunakan Metode YOLO Berbasis Image Processing pada Video Jalan Raya. *Jurnal Algoritma*, 23(1), 390–399. <https://doi.org/10.33364/algoritma/v.23-1.3152>
- Setiyadi, A., Utami, E., & Ariatmanto, D. (2023). Analisa Kemampuan Algoritma YOLOv8 Dalam Deteksi Objek Manusia Dengan Metode Modifikasi Arsitektur. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 7(2), 891–901. <https://doi.org/10.30645/j-sakti.v7i2.694>
- Wijayanti, K., Mutiara, G. A., Suryawardani, B., Erwina, E., & Kusuma, G. P. (2025). Non-Intrusive Real-Time Tourist Crowd Monitoring for Overtourism Mitigation using YOLOv8-Based Head Detection and Tracking. *Journal of Robotics and Control*, 6(4), 1985–2004. <https://doi.org/10.18196/jrc.v6i4.26396>