

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Industri restoran dan *food service* merupakan salah satu sektor yang sangat bergantung pada kualitas pelayanan terhadap pelanggan. Salah satu aspek penting dalam pelayanan restoran adalah kecepatan dan ketepatan pelayan dalam menangani meja makan, termasuk mengangkat piring dan gelas yang telah selesai digunakan oleh tamu. Keterlambatan dalam proses pengangkatan peralatan makan dapat menurunkan kepuasan pelanggan dan mengurangi efisiensi operasional restoran secara keseluruhan. Penelitian di bidang sistem pemantauan restoran berbasis visi komputer menunjukkan bahwa pendekatan otomatis berbasis kamera mampu menggantikan sistem sensor fisik konvensional yang berbiaya tinggi dan tidak fleksibel, sekaligus meningkatkan efisiensi manajemen operasional secara signifikan [1].

Pada umumnya, pemantauan kondisi meja makan di restoran masih dilakukan secara manual oleh pelayan. Pendekatan ini memiliki keterbatasan yang inheren, antara lain bersifat subjektif, bergantung pada perhatian dan inisiatif masing-masing pelayan, serta tidak dapat dilakukan secara *real-time* dan konsisten terutama pada restoran dengan volume pelanggan yang tinggi. Kondisi ini menimbulkan kebutuhan nyata akan suatu sistem pemantauan otomatis yang mampu mendeteksi kondisi peralatan makan secara objektif, cepat, dan konsisten tanpa bergantung pada faktor manusia.

Perkembangan teknologi *Computer Vision* dan *Deep Learning* telah membuka peluang besar dalam otomatisasi pemantauan visual berbasis kamera. Melalui pemanfaatan *object detection*, sistem komputer dapat mengenali dan melokalisasi posisi objek dari dunia nyata secara otomatis [2]. Di dalam sektor industri, implementasi arsitektur berbasis jaringan saraf konvolusi (*Convolutional Neural Network/CNN*) juga telah terbukti andal dalam mengenali kondisi objek secara otomatis dengan tingkat akurasi yang tinggi [3]. Dalam skenario operasional restoran, teknologi deteksi objek ini dapat diintegrasikan untuk memantau keberadaan serta kondisi spesifik dari piring dan gelas di atas meja makan secara *real-time* guna mengotomatisasi pengiriman notifikasi penanganan kepada pelayan [1].

Penelitian sebelumnya menunjukkan bahwa *deep learning* mampu mengenali kondisi makanan secara visual dengan akurasi yang tinggi, termasuk membedakan kondisi piring berisi makanan dari piring yang sudah habis digunakan [4]. Temuan ini memperkuat kelayakan penerapan teknologi *computer vision* untuk otomatisasi pemantauan kondisi peralatan makan di lingkungan restoran.

Algoritma *You Only Look Once* (YOLO) merupakan keluarga algoritma *object detection* berbasis *deep learning* yang paling banyak diadopsi dalam penelitian terapan karena kemampuannya melakukan deteksi dalam satu tahap pemrosesan (*single-stage detector*) dengan keseimbangan yang baik antara kecepatan dan akurasi [2, 5]. Sejak pertama kali diperkenalkan, YOLO telah mengalami evolusi yang pesat melalui berbagai inovasi arsitektur [6]. YOLOv8 yang dirilis oleh Ultralytics pada Januari 2023 memperkenalkan arsitektur *anchor-free* dan modul C2f yang meningkatkan kemampuan ekstraksi fitur dibandingkan versi sebelumnya [7]. YOLOv9s yang dipresentasikan pada ECCV 2024 memperkenalkan *Programmable Gradient Information* (PGI) dan *Generalized Efficient Layer Aggregation Network* (GELAN) untuk mengatasi permasalahan *information loss* pada jaringan yang dalam [8]. YOLOv11 yang dirilis pada Oktober 2024 memperkenalkan blok C3k2 dan modul C2PSA yang meningkatkan efisiensi parameter sekaligus akurasi deteksi [9].

Pendekatan deteksi dalam penelitian ini difokuskan pada level objek (piring dan gelas) dibandingkan level kondisi meja secara keseluruhan, dengan beberapa pertimbangan. Pertama, deteksi berbasis objek memberikan informasi yang lebih spesifik dan *actionable* sistem dapat membedakan secara presisi objek mana yang perlu diangkat. Kedua, pendekatan berbasis kondisi meja secara keseluruhan rentan terhadap ambiguitas visual karena objek di atas meja dapat bermacam-macam bentuknya, sementara kondisi peralatan makan individual dapat didefinisikan dan dianotasi secara lebih konsisten [5]. Ketiga, sistem pemantauan restoran berbasis *computer vision* umumnya menggunakan deteksi objek sebagai komponen inti untuk menentukan status layanan meja secara otomatis [1].

Meskipun setiap versi YOLO mengklaim peningkatan performa, terdapat kesenjangan penelitian yang signifikan: belum ada studi yang secara khusus membandingkan YOLOv8, YOLOv9, dan YOLOv11 pada domain deteksi kondisi peralatan makan di lingkungan restoran. Penelitian komparatif yang ada umumnya berfokus pada dataset generik seperti MS COCO [8] atau domain yang tidak terkait seperti kendaraan [10] dan aplikasi industri [3], sehingga belum memberikan panduan yang valid untuk pemilihan model pada kasus *use case* spesifik restoran.

Selain itu, penelitian terdahulu pada domain makanan dan piring umumnya hanya mengevaluasi satu versi YOLO tanpa membandingkan *trade-off* antara aspek akurasi deteksi, kecepatan inferensi, dan efisiensi model secara sistematis [11]. Kesenjangan ini menjadi permasalahan praktis yang nyata: pengelola sistem restoran yang ingin mengadopsi teknologi *object detection* tidak memiliki referensi yang cukup untuk memilih versi YOLO yang paling sesuai dengan kebutuhan operasional mereka.

Berdasarkan kesenjangan tersebut, penelitian ini mengusulkan studi komparatif yang sistematis antara YOLOv8s, YOLOv9s, dan YOLOv11s untuk mendeteksi kondisi peralatan makan berupa piring dan gelas pada meja restoran. Perbandingan dilakukan secara menyeluruh berdasarkan dimensi utama metrik evaluasi yang meliputi akurasi deteksi (*precision*, *recall*, *mAP@50*, dan *mAP@50-95*), kecepatan inferensi (*inference time* dan FPS), serta efisiensi model komputasi (jumlah parameter dan GFLOPs). Penelitian ini diharapkan dapat mengisi kesenjangan literatur yang ada dan memberikan rekomendasi berbasis data yang dapat langsung diterapkan dalam pengembangan sistem pemantauan meja makan restoran secara *real-time*.

Urgensi dari penelitian ini didasarkan pada kebutuhan industri restoran untuk meningkatkan *table turnover rate* dan efisiensi operasional pramusaji. Keterlambatan dalam mengangkat piring kotor atau mendeteksi gelas kosong secara manual sering kali berdampak langsung pada penurunan tingkat kepuasan pelanggan. Melalui evaluasi komparatif multi-versi YOLO (YOLOv8, YOLOv9, dan YOLOv11), *trade-off* optimal antara akurasi deteksi (*precision*, *recall*, *mAP@50*, *mAP@50-95*), kecepatan inferensi (FPS), serta efisiensi komputasi (GFLOPs) dapat dipetakan secara terukur. Hal ini krusial untuk memberikan referensi berbasis data bagi pengembang sistem yang membutuhkan model dengan presisi tinggi guna meminimalkan kesalahan deteksi (*false alarm*) sebelum sistem diimplementasikan pada perangkat komputasi tepi (*edge devices*) di restoran.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana implementasi dan perbandingan kinerja YOLOv8s, YOLOv9s, dan YOLOv11s dalam mendeteksi kondisi peralatan makan (piring dan gelas) pada meja restoran secara *real-time*, ditinjau dari aspek akurasi

deteksi (*precision*, *recall*, *mAP@50*, dan *mAP@50-95*), kecepatan inferensi (*inference time/FPS*), serta efisiensi komputasi model (jumlah parameter dan GFLOPs), sehingga dapat ditentukan versi YOLO yang memberikan keseimbangan terbaik untuk diterapkan pada sistem deteksi restoran secara *real-time*?

1.3 Batasan Permasalahan

Agar penelitian lebih terfokus, ditetapkan batasan permasalahan sebagai berikut:

1. Kelas objek yang dideteksi terdiri dari lima kelas, yaitu *piring kotor* (piring dengan sisa makanan, siap diangkat), *piring aktif* (piring dengan makanan, tamu sedang makan), *piring kosong* (piring bersih belum digunakan), *gelas aktif* (gelas berisi minuman), dan *gelas kosong* (gelas yang telah habis isinya dan siap diangkat oleh pelayan). Penentuan kondisi kelas didasarkan pada status visual objek secara individual, bukan pada keberadaan atau ketidakhadiran tamu di sekitar meja.
2. Dataset yang digunakan merupakan kombinasi dataset publik dari Roboflow Universe dan data yang dikumpulkan secara mandiri, dengan total 1.000 gambar yang dibagi menjadi data latih (70%), validasi (20%), dan uji (10%).
3. Dataset yang dikumpulkan didominasi oleh peralatan makan berbentuk konvensional (piring lingkaran dan mangkok); variasi bentuk penyajian lain seperti kotak kertas, *hot plate*, atau wadah anyaman tidak diikutsertakan.
4. Sistem yang dikembangkan tidak terintegrasi dengan sistem manajemen restoran atau perangkat IoT secara langsung.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Membandingkan performa akurasi YOLOv8s, YOLOv9s, dan YOLOv11s dalam mendeteksi kondisi peralatan makan pada meja restoran secara konsisten menggunakan metrik *precision*, *recall*, *mAP@50*, dan *mAP@50-95*.

2. Membandingkan kecepatan inferensi (*inference time* dan FPS) serta efisiensi komputasi model (jumlah parameter dan GFLOPs) antara ketiga versi YOLO pada dataset dan perangkat keras yang sama.
3. Menentukan versi YOLO yang paling optimal untuk diterapkan pada sistem deteksi kondisi peralatan makan restoran berdasarkan *trade-off* yang seimbang antara akurasi deteksi (*precision*, *recall*, *mAP@50*, *mAP@50-95*), kecepatan inferensi (FPS), serta efisiensi parameter model (GFLOPs).

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan kontribusi ilmiah berupa studi komparatif yang sistematis antara YOLOv8s, YOLOv9s, dan YOLOv11s pada domain deteksi kondisi peralatan makan restoran yang belum pernah diteliti sebelumnya, sehingga dapat menjadi referensi bagi penelitian selanjutnya di bidang *computer vision* terapan.
2. Menyediakan rekomendasi berbasis data mengenai pemilihan versi YOLO yang tepat untuk aplikasi *real-time object detection* di lingkungan restoran, dengan mempertimbangkan *trade-off* komprehensif antara akurasi deteksi (*precision*, *recall*, *mAP@50*, *mAP@50-95*), kecepatan inferensi (FPS), dan efisiensi komputasi (GFLOPs).
3. Membantu pengelola restoran dalam memahami potensi penerapan teknologi *computer vision* untuk meningkatkan efisiensi dan kualitas layanan secara otomatis melalui sistem deteksi kondisi peralatan makan berbasis kamera.
4. Mendukung pengembangan sistem pemantauan meja makan otomatis yang dapat diintegrasikan ke dalam sistem operasional restoran di masa mendatang.

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini disusun untuk memberikan gambaran mengenai struktur dan isi laporan penelitian secara keseluruhan. Sistematika penulisan laporan adalah sebagai berikut:

- **Bab 1 PENDAHULUAN** membahas latar belakang permasalahan mengenai kebutuhan sistem deteksi kondisi peralatan makan di restoran secara otomatis, identifikasi kesenjangan penelitian, rumusan masalah, batasan permasalahan, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.
- **Bab 2 LANDASAN TEORI** membahas teori-teori yang mendasari penelitian, meliputi *computer vision* dalam konteks restoran, *deep learning* untuk deteksi objek, *object detection*, evolusi algoritma YOLO, serta arsitektur dan karakteristik YOLOv8, YOLOv9, dan YOLOv11.
- **Bab 3 METODOLOGI PENELITIAN** menjelaskan tahapan penelitian secara sistematis, meliputi pengumpulan dan anotasi dataset, *preprocessing* data, konfigurasi pelatihan model dengan *hyperparameter* identik, serta alur metode evaluasi kuantitatif berdasarkan metrik *precision*, *recall*, *mAP@50*, *mAP@50-95*, kecepatan inferensi, dan GFLOPs.
- **Bab 4 HASIL DAN DISKUSI** menyajikan hasil pelatihan dan pengujian komparatif ketiga model YOLOv8s, YOLOv9s, dan YOLOv11s pada data uji, analisis perbandingan performa terperinci berdasarkan metrik akurasi deteksi visual (*precision*, *recall*, *mAP@50*, *mAP@50-95*), kecepatan inferensi (FPS), dan beban efisiensi komputasi (GFLOPs), serta diskusi terhadap temuan penelitian.
- **Bab 5 SIMPULAN DAN SARAN** menyampaikan simpulan menyeluruh dari penelitian berdasarkan rangkuman metrik evaluasi final (*precision*, *recall*, *mAP@50*, *mAP@50-95*, FPS, GFLOPs), serta saran aplikatif untuk pengembangan penelitian selanjutnya.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A