

ANALISIS KOMPARATIF KINERJA RF-DETR DAN YOLOV11 UNTUK DETEKSI DAN KLASIFIKASI KUALITAS CABAI

Karisma Wulan Sari^{*,1)}, Budi Sutomo²⁾

¹⁾ Program Studi Teknik Informatika, Fakultas Teknologi Bisnis dan Sains Dharma Wacana, Universitas Dharma Wacana Metro

Jl. Kenanga No.3, Mulyojati, Kec. Metro Barat, Kota Metro, Lampung 34121, Indonesia

Email: ¹Karismawulan20@gmail.com

Abstract

Increasing demands for the quality of horticultural commodities highlight the limitations of manual sorting methods, which tend to be subjective and inefficient. This study aims to develop and compare the YOLOv11 and RF-DETR models for the detection and classification of chili pepper quality using computer vision and deep learning. The research methods include dataset collection, instance segmentation annotation, preprocessing, data augmentation, hyperparameter tuning, model training, and evaluation using precision, recall, mean Average Precision (mAP), and inference latency metrics.

The results show that YOLOv11 achieved a precision of 0.87, a recall of 0.96, an mAP50 of 0.89, and a latency of ± 10 ms per image, making it suitable for real-time applications. Meanwhile, RF-DETR achieved a precision of 0.90, a recall of 0.94, and an mAP50 of 0.92 with a latency of ± 18 ms per image. RF-DETR demonstrated better detection accuracy and stability under complex visual conditions, while YOLOv11 excelled in computational efficiency. This study demonstrates that computer vision-based deep learning has the potential to improve the efficiency, consistency, and objectivity of the automated post-harvest chili grading process.

Keywords: Chili Quality Detection, Deep Learning, Object Detection, YOLOv11, RF-DETR

Abstrak

Peningkatan tuntutan terhadap kualitas komoditas hortikultura menyoroti keterbatasan metode sortasi manual yang cenderung subjektif dan tidak efisien. Penelitian ini bertujuan mengembangkan serta membandingkan model YOLOv11 dan RF-DETR untuk deteksi dan klasifikasi kualitas cabai berbasis computer vision dan deep learning. Metode penelitian meliputi pengumpulan dataset, anotasi instance segmentation, pra-pemrosesan, augmentasi data, hyperparameter tuning, pelatihan model, dan evaluasi menggunakan metrik precision, recall, mean Average Precision (mAP), serta latency inferensi.

Hasil penelitian menunjukkan bahwa YOLOv11 memperoleh precision sebesar 0.87, recall sebesar 0.96, mAP50 sebesar 0.89, dan latency ± 10 ms per citra sehingga cocok untuk aplikasi real-time. Sementara itu, RF-DETR memperoleh precision sebesar 0.90, recall sebesar 0.94, dan mAP50 sebesar 0.92 dengan latency ± 18 ms per citra. RF-DETR menunjukkan akurasi dan stabilitas deteksi yang lebih baik pada kondisi visual kompleks, sedangkan YOLOv11 unggul dalam efisiensi komputasi. Penelitian ini menunjukkan bahwa deep learning berbasis computer vision berpotensi meningkatkan efisiensi, konsistensi, dan objektivitas proses grading cabai pascapanen secara otomatis.

Kata Kunci: Deteksi Kualitas Cabai, Deep Learning, Deteksi Objek, YOLOv11, RF-DETR

1. Pendahuluan

Komoditas hortikultura memiliki peran strategis dalam sektor agrikultur Indonesia, terutama dalam menjaga ketahanan pangan, stabilitas ekonomi, dan distribusi bahan pangan nasional. Salah satu komoditas hortikultura yang memiliki tingkat konsumsi tinggi dan nilai ekonomi penting adalah cabai.[1] Cabai merupakan kebutuhan pokok masyarakat Indonesia yang permintaannya terus meningkat setiap tahun serta sering menjadi penyumbang inflasi nasional akibat fluktuasi harga yang tinggi [2]. Fluktuasi harga cabai dipengaruhi oleh berbagai faktor, seperti penurunan kualitas hasil panen, distribusi yang tidak stabil, serta tingginya tingkat kerusakan pascapanen. Pada kondisi tertentu, harga cabai dapat mengalami kenaikan drastis akibat keterbatasan pasokan di pasar. Kondisi tersebut menunjukkan bahwa kualitas dan kestabilan distribusi cabai menjadi aspek penting dalam menjaga kestabilan ekonomi pangan nasional. Oleh karena



itu, diperlukan sistem *grading* kualitas cabai yang mampu melakukan proses inspeksi secara cepat, objektif, dan konsisten untuk mendukung efisiensi distribusi hasil panen.

Di sisi lain, proses sortasi dan *grading* cabai di Indonesia masih banyak dilakukan secara manual dengan mengandalkan pengamatan visual manusia. Metode manual memiliki berbagai kelemahan, seperti subjektivitas penilaian, inkonsistensi hasil *grading*, serta membutuhkan waktu dan tenaga kerja yang besar [3]. Ketidakakuratan *grading* menyebabkan cabai berkualitas rendah tercampur dengan cabai berkualitas baik sehingga menurunkan nilai jual keseluruhan produk. Dampak lainnya adalah kerugian ekonomi bagi petani dan distributor karena hasil panen tidak dapat dihargai secara optimal sesuai kualitasnya. Selain itu, kesalahan klasifikasi cabai busuk atau belum matang dapat menyebabkan penurunan kualitas produk yang beredar di pasar dan meningkatkan risiko penolakan produk oleh konsumen maupun distributor. Permasalahan ini semakin kompleks karena cabai umumnya berada pada lingkungan visual yang tidak terstruktur, seperti pencahayaan yang tidak merata, latar belakang yang bervariasi, serta kondisi oklusi antarobjek [4]. Kondisi tersebut menjadi tantangan utama dalam penerapan sistem visi komputer untuk mendeteksi dan mengklasifikasikan kualitas cabai secara otomatis.

Perkembangan teknologi *Deep Learning* dan *Computer Vision* telah memberikan solusi yang signifikan dalam pengembangan sistem *grading* otomatis pada sektor pertanian [4]. Model berbasis *Convolutional Neural Network* (CNN), khususnya keluarga *You Only Look Once* (YOLO), telah banyak digunakan dalam penelitian deteksi objek *real-time* karena memiliki kombinasi akurasi dan kecepatan yang baik [5]. Penelitian sebelumnya berhasil mengembangkan sistem deteksi kualitas cabai menggunakan metode YOLO untuk mengklasifikasikan cabai ke dalam kategori baik, buruk, dan mentah sebagai solusi terhadap keterbatasan *grading* manual [6]. Selain itu, metode YOLO juga telah diterapkan pada klasifikasi penyakit tanaman cabai melalui pendekatan segmentasi citra daun dan buah dengan hasil deteksi yang baik [7]. Penelitian lain menggunakan YOLOv11 menunjukkan bahwa model tersebut mampu memberikan performa deteksi yang baik pada objek pertanian dengan efisiensi komputasi yang tinggi [8]. Selain itu, penerapan model YOLOv11 pada deteksi tingkat kematangan buah kopi, stroberi, dan pepaya menunjukkan bahwa teknologi *Deep Learning* mampu meningkatkan akurasi *grading* kualitas produk hortikultura secara otomatis [9], [10]. Pada penelitian lain, model YOLOv8 dan DETR juga menunjukkan kemampuan yang kompetitif dalam proses deteksi tingkat kematangan tomat berbasis visi komputer [11]. Hasil penelitian tersebut menunjukkan bahwa model *Deep Learning* memiliki potensi besar untuk diterapkan dalam proses *grading* kualitas cabai otomatis secara *real-time*.

Meskipun berbagai penelitian mengenai deteksi objek dan *grading* otomatis telah berkembang pesat, masih terdapat kesenjangan penelitian yang relevan untuk dikaji lebih lanjut. Sebagian besar penelitian terdahulu masih berfokus pada model CNN konvensional seperti YOLOv5, YOLOv7, dan YOLOv8, sedangkan penelitian yang membandingkan model transformer modern dengan model CNN terbaru pada domain kualitas cabai masih sangat terbatas [5], [11]. Selain itu, penelitian sebelumnya umumnya hanya menitikberatkan pada aspek akurasi tanpa mengevaluasi *trade-off* antara akurasi dan *latency* dalam implementasi *real-time*. Model berbasis CNN dikenal unggul dalam efisiensi komputasi dan kecepatan inferensi, tetapi memiliki keterbatasan dalam memahami hubungan spasial global pada objek yang ambigu atau teroklusi. Sebaliknya, arsitektur *Detection Transformer* (DETR) menawarkan kemampuan *global context modeling* melalui mekanisme *self-attention* yang memungkinkan model memahami dependensi jarak jauh antarfitur objek [11]. Perkembangan terbaru menghasilkan RF-DETR yang dioptimalkan untuk mencapai akurasi tinggi dengan performa *real-time* pada komputasi terbatas [12]. Namun, implementasi RF-DETR pada deteksi kualitas cabai pascapanen, khususnya untuk membedakan cabai segar, mentah, dan busuk pada lingkungan visual kompleks, masih belum banyak dikaji secara empiris. Oleh karena itu, diperlukan penelitian komparatif yang mampu mengevaluasi performa model transformer dan CNN modern secara langsung dalam konteks *grading* kualitas cabai otomatis.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan implementasi dua model *state-of-the-art*, yaitu RF-DETR dan YOLOv11 untuk mendeteksi dan mengklasifikasikan kualitas cabai pascapanen ke dalam tiga kategori, yaitu Segar, mentah, dan Busuk/Rusak. Pendekatan *instance segmentation* dipilih karena mampu memberikan informasi batas objek secara detail berbasis *mask* piksel, sehingga lebih efektif dalam membedakan area kerusakan, tekstur, dan tingkat kematangan cabai dibandingkan metode *bounding box* konvensional. YOLOv11 dipilih karena menawarkan efisiensi komputasi yang tinggi dengan jumlah parameter yang lebih sedikit dibanding generasi sebelumnya namun tetap mempertahankan performa deteksi yang optimal. Sementara itu, RF-DETR dipilih karena memiliki kemampuan pemodelan konteks global yang lebih baik melalui mekanisme transformer sehingga berpotensi meningkatkan akurasi pada kondisi visual kompleks dan ambigu. Proses pelatihan dan evaluasi model dilakukan menggunakan platform *Google Colaboratory* (*Google Colab*) yang mendukung komputasi berbasis GPU untuk mempercepat proses *training* model *Deep Learning*. Perbandingan kedua model ini diharapkan dapat memberikan pemahaman mengenai *trade-off* terbaik antara kecepatan inferensi dan akurasi deteksi untuk aplikasi *grading* cabai secara *real-time*.

Penelitian ini bertujuan untuk mengembangkan dataset *instance segmentation* cabai pascapanen berbasis anotasi *mask* piksel, mengimplementasikan model RF-DETR dan YOLOv11, serta melakukan analisis komparatif performa

kedua model berdasarkan metrik mAP50:95 dan *latency*. Kontribusi penelitian ini tidak hanya terletak pada evaluasi empiris dua arsitektur deteksi objek modern, tetapi juga pada pengembangan solusi *grading* otomatis yang dapat diterapkan pada sektor agrikultur modern dan industri pascapanen. Sistem *grading* berbasis *Deep Learning* diharapkan mampu meningkatkan konsistensi kualitas produk, mengurangi subjektivitas manusia dalam proses sortasi, mempercepat distribusi hasil panen, serta membantu petani dan distributor dalam meningkatkan nilai ekonomi komoditas cabai. Selain itu, penelitian ini juga memberikan kontribusi ilmiah berupa analisis penerapan model transformer dan CNN modern pada domain hortikultura yang masih relatif terbatas dalam literatur penelitian saat ini.

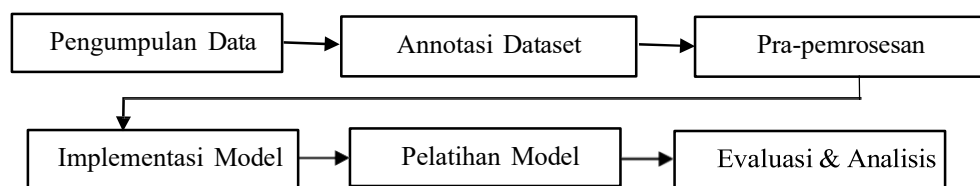
2. Metode

2.1 Desain Penelitian

Penelitian ini menggunakan pendekatan eksperimen kuantitatif dengan metode komparatif. Pendekatan ini dipilih untuk membandingkan kinerja dua model *object detection and segmentation* berbasis *deep learning*, yaitu RF-DETR Seg dan YOLOv11 Seg, dalam mendeteksi kualitas cabai pascapanen. Eksperimen dilakukan melalui tahapan sistematis mulai dari pengumpulan data, pelabelan (*annotation*), pra-pemrosesan, pelatihan model (*training*), evaluasi, hingga analisis hasil dengan metrik mAP50:95 dan *latency* (ms/inference).

2.2 Tahapan Penelitian

Secara umum, langkah-langkah penelitian ini ditunjukkan sebagai berikut:



Gambar 1. alur penelitian

2.3 Pengumpulan Data

Kualitas data merupakan faktor penentu dalam keberhasilan pelatihan model *deep learning*. Mengingat rencana untuk membuat *dataset* kustom melalui fotografi mandiri, protokol yang terstandarisasi harus diimplementasikan.

Sumber Data

Data utama yang digunakan dalam penelitian ini merupakan citra cabai pascapanen yang diperoleh dari dua sumber, yaitu:

- 1 Akuisisi langsung (data primer) melalui pengambilan gambar cabai menggunakan kamera digital, dan
- 2 *Dataset* sekunder dari sumber terbuka (jika diperlukan) untuk melengkapi variasi kondisi visual.

Data primer dikumpulkan di lingkungan pasar tradisional dan gudang penyortiran hasil pertanian, di mana cabai ditempatkan dalam berbagai kondisi alami, seperti:

- 1 Pencahayaan alami (*outdoor*)
- 2 Pencahayaan buatan (*indoor*)
- 3 Latar belakang beragam (meja, wadah, kain, atau permukaan kayu, kertas)

Jenis cabai yang difokuskan dalam penelitian ini adalah cabai merah keriting (*Capsicum annuum L.*) karena merupakan varietas yang paling umum dikonsumsi dan memiliki peran besar dalam fluktuasi harga pasar di Indonesia.

Ukuran *Dataset*

Untuk memastikan pelatihan *deep learning* yang stabil dan mengurangi risiko *overfitting*, meskipun *transfer learning* akan diterapkan, Untuk mendukung proses pelatihan model *deep learning* yang stabil dan mengurangi risiko *overfitting*, *dataset* penelitian dikembangkan melalui proses augmentasi sehingga variasi data menjadi lebih beragam.

Definisi Operasional Kelas Objek

Dataset kustom ini harus mendefinisikan secara operasional tiga kelas utama yang akan dideteksi oleh model: Segar (*Fresh/Good*), Mentah (*Unripe/Green*), dan Busuk/Rusak (*Spoiled/Defect*).

Table 1. Definisi Operasional Kelas Kualitas Cabai dalam *Dataset* Kustom

Kelas Objek	DESKRIPSI VISUAL DAN KRITERIA ANOTASI
Cabai Segar (<i>Fresh</i>)	Buah utuh, warna matang penuh, tidak ada cacat signifikan.
Cabai Mentah (<i>Unripe</i>)	Buah utuh, belum mencapai kematangan warna yang ditargetkan (hijau solid atau transisional).
Cabai busuk/rusak (<i>Defect</i>)	Memiliki kerusakan fisik atau biologis (pembusukan, bercak penyakit) melebihi ambang batas yang diizinkan untuk pasar.

2.4 Annotasi *Dataset*

Tahap berikutnya adalah pelabelan citra menggunakan perangkat lunak Roboflow. Proses anotasi dilakukan pada level *instance segmentation*, di mana setiap objek cabai diberi label dengan *mask* berbasis piksel sesuai kelasnya:

1. Label 0 : Cabai Segar
2. Label 1 : Cabai mentah
3. Label 2 : Cabai Busuk/Rusak

Setelah proses anotasi selesai, *dataset* terbagi menjadi tiga *subset*:

2.5 Pra-pemrosesan Data

Tahapan ini bertujuan menyiapkan *dataset* agar optimal untuk dilatih oleh model. Langkah- langkah meliputi:

Resize dan Normalisasi

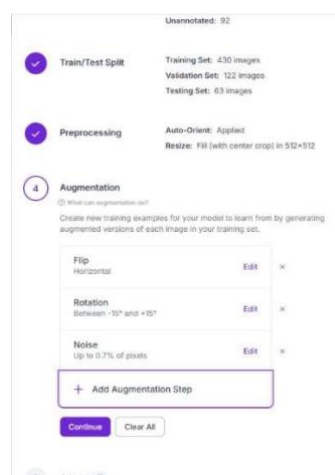
Setiap citra diubah ukurannya menjadi resolusi 640×640 piksel agar sesuai dengan input standar kedua model. Selain itu, dilakukan normalisasi nilai piksel (0–1) untuk mempercepat proses konvergensi model saat pelatihan.

Augmentasi Data

Untuk memperbanyak variasi citra dan meningkatkan kemampuan generalisasi model, dilakukan augmentasi data di Roboflow dengan beberapa teknik berikut:

1. *Flip horizontal* (membalik gambar ke arah kiri-kanan)
2. *Brightness & Contrast adjustment* (penyesuaian pencahayaan dan kontras)
3. *Rotation* ($\pm 15^\circ$) untuk variasi sudut pengambilan gambar
4. *Gaussian Blur* (0–1 px) untuk simulasi ketajaman gambar berbeda
5. *Zoom-in/out* (10–20%) agar model mengenali objek pada jarak berbeda

Dataset utama yang digunakan dalam penelitian ini terdiri dari 1.475 citra asli yang dikumpulkan secara mandiri. Untuk memperkaya variasi fitur, teknik augmentasi data seperti rotasi, *flipping*, dan penyesuaian kecerahan diterapkan secara manual melalui platform Roboflow sebelum proses pelatihan model dilakukan. Meskipun proses augmentasi manual ini menghasilkan variasi citra baru untuk memperkuat kemampuan generalisasi model, pencatatan dan pelaporan komposisi data pada draf ini tetap difokuskan pada 1.475 basis data asli (yang terbagi menjadi data pelatihan, validasi, dan pengujian seperti dirinci pada Tabel 3.1) untuk menjaga konsistensi dan objektivitas metrik evaluasi utama.

**Gambar 2.** Augmentasi *Dataset*

Pengaturan Format Anotasi

Roboflow menyediakan konversi otomatis ke berbagai format anotasi, seperti COCO JSON, YOLOv8/11 Segmentation, dan Pascal VOC XML. Dalam penelitian ini:

1. Untuk RF-DETR, digunakan format COCO JSON.
2. Untuk YOLOv11, digunakan format YOLOv8 Segmentation (.txt).

Langkah ini memastikan kompatibilitas antara *dataset* dan arsitektur model yang digunakan.

2.6 Implementasi Model

Implementasi pelatihan model dalam penelitian ini dilakukan menggunakan Google Colab sebagai lingkungan komputasi berbasis *cloud* yang mendukung GPU. Platform ini dipilih karena menyediakan sumber daya komputasi yang cukup untuk melatih model berbasis *deep learning* tanpa memerlukan perangkat keras lokal berperforma tinggi. Model yang digunakan adalah RF-DETR dan YOLOv11, yang keduanya diimplementasikan melalui *notebook* pelatihan berbasis Python. Untuk RF-DETR, pelatihan dilakukan menggunakan referensi *notebook* dari repositori PyResearch dan Roboflow yang menyediakan panduan *fine-tuning* model pada *dataset* deteksi objek.

2.7 Pelatihan Model

Kedua Model RF-DETR dan YOLOv11 menggunakan *framework* PyTorch sebagai basis pelatihan, yang mendukung *training pipeline end-to-end* dengan metode *Distributed Data Parallel* (DDP) untuk efisiensi pada GPU. Proses pelatihan meliputi:

1. Inisialisasi model dengan bobot awal (*pre-trained weights*) dari *dataset* COCO.
2. Konfigurasi parameter pelatihan, seperti *batch size*, *learning rate*, *image size*, dan jumlah *epoch*.

Parameter pelatihan yang digunakan pada masing-masing model ditentukan berdasarkan kemampuan komputasi Google Colab serta hasil pengujian awal untuk menjaga stabilitas proses *training* dan mengurangi risiko *overfitting*. Konfigurasi hyperparameter yang digunakan dapat dilihat pada Tabel 2.2. Proses pelatihan dilakukan menggunakan *framework Ultralytics* untuk YOLOv11 dan repositori resmi untuk RF-DETR. Kedua model dilatih menggunakan Google Colab dengan spesifikasi GPU T4. Nilai parameter utama yang digunakan dalam proses pelatihan ini dirinci pada Tabel 2.2. Proses pelatihan kedua model dilakukan menggunakan *pendekatan Transfer Learning* dengan memanfaatkan *pre-trained weights* yang telah dilatih pada *dataset* COCO. Mengingat bobot fitur dasar arsitektur telah terbentuk secara optimal, proses *fine-tuning* pada penelitian ini dibatasi secara konstan pada 5 *epoch*. Pembatasan ini dilakukan untuk menghindari *overfitting* pada *dataset* spesifik cabai serta menguji efisiensi konvergensi cepat dari kedua model mutakhir tersebut dalam mengenali fitur geometris dan warna komoditas hortikultura.

Tabel 2. Hyperparameter Training Model

Parameter	YOLOv11	RF-DETR
<i>Epoch</i>	5	5
<i>Batch Size</i>	4	4
Resolusi Input	640 × 640	640 × 640
<i>Learning Rate</i>	Default Ultralytics	1×10 ⁻⁴
Gradient Accumulation	-	1
Device	GPU CUDA	GPU CUDA
<i>Pretrained Weight</i>	yolov11n.pt	RFDETRBase pretrained

Model YOLOv11 menggunakan pengaturan learning rate bawaan dari Ultralytics yang otomatis menyesuaikan selama proses *training*. Sementara itu, RF-DETR menggunakan *learning rate* sebesar 1×10⁻⁴ dengan grad accumulation sebesar 1 untuk menjaga kestabilan proses pelatihan model.

3. Pemanggilan *dataset* yang telah diproses dan diekspor melalui Roboflow dalam format YOLO/COCO.
4. Proses pelatihan (*training loop*) dijalankan untuk menyesuaikan bobot model terhadap *dataset* cabai.
5. Evaluasi performa model dilakukan secara otomatis menggunakan metrik *mean Average Precision* (mAP50:95).

Hasil pelatihan berupa model terlatih (.pt file) disimpan pada penyimpanan *cloud* Colab, dan dapat dievaluasi lebih lanjut untuk analisis performa serta perbandingan dengan model YOLOv11.

2.8 Evaluasi & Analisis

Tahap evaluasi dan analisis dilakukan untuk mengukur serta menafsirkan performa kedua model deteksi objek, yaitu RF-DETR dan YOLOv11, dalam mengidentifikasi serta mengklasifikasikan kualitas cabai pascapanen. Tujuan utamanya adalah menentukan model yang paling optimal berdasarkan keseimbangan antara akurasi, efisiensi komputasi, dan kemampuan generalisasi terhadap data uji.

Evaluasi Model

Evaluasi dilakukan menggunakan *subset* data validasi dan data uji yang telah dipisahkan dari *dataset* utama. Penilaian menggunakan metrik standar pada deteksi objek, meliputi:

1. *Mean Average Precision* (mAP50 dan mAP50:95)
mAP50 mengukur akurasi dengan ambang IoU ≥ 0.5 , sedangkan mAP50:95 merupakan rata-rata mAP pada sepuluh ambang (0.5–0.95). Metrik ini menggambarkan konsistensi performa model.
2. *Precision* (Presisi)
Mengukur proporsi prediksi benar terhadap seluruh prediksi positif. Nilai tinggi menunjukkan kemampuan menghindari *false positive*.
3. *Recall* (Sensitivitas)
Menunjukkan kemampuan model mendeteksi seluruh objek yang benar pada citra uji. Nilai tinggi berarti model efektif mengenali semua cabai.
4. *Latency* (*Inference Time*)
Mengukur kecepatan model mendeteksi per citra (ms/image). Parameter ini penting untuk penerapan waktu nyata (*real-time*).

Seluruh hasil evaluasi diperoleh otomatis melalui fungsi evaluasi pada *notebook* pelatihan di Google Colab. Setiap *epoch* menghasilkan nilai mAP, presisi, *recall*, serta grafik konvergensi model.

Analisis Model

Tahap analisis dilakukan setelah seluruh proses evaluasi selesai. Hasil Evaluasi Model yang diperoleh kemudian ditafsirkan untuk menentukan model yang paling sesuai dengan kebutuhan sistem *grading* cabai. Analisis difokuskan pada empat aspek utama, yaitu kinerja akurasi dan kecepatan, konsistensi deteksi, identifikasi kesalahan, serta rekomendasi model optimal.

Perbandingan kinerja antara RF-DETR dan YOLOv11 dilakukan berdasarkan nilai mAP50:95, presisi, *recall*, dan waktu inferensi. Analisis ini digunakan untuk melihat *trade-off* antara tingkat akurasi yang tinggi (seperti pada RF-DETR) dengan efisiensi komputasi yang lebih cepat (seperti pada YOLOv11). Selain itu, model juga diuji terhadap berbagai kondisi pencahayaan dan orientasi objek guna menilai stabilitas prediksi dalam situasi yang tidak ideal.

Selanjutnya dilakukan analisis kesalahan (*error analysis*) untuk mengidentifikasi potensi kelemahan dari masing-masing model, seperti munculnya *false positive*, *false negative*, atau kesalahan klasifikasi. RF-DETR diharapkan memiliki keunggulan dalam mendeteksi objek yang tumpang tindih dan detail kecil, sedangkan YOLOv11 diprediksi lebih efisien dalam kecepatan namun cenderung memiliki akurasi yang sedikit lebih rendah.

Berdasarkan hasil analisis akhir, model dengan keseimbangan terbaik antara akurasi (mAP50:95) dan efisiensi waktu inferensi akan direkomendasikan sebagai model unggulan untuk sistem otomatisasi *grading* cabai pascapanen. Model terpilih diharapkan mampu memberikan hasil deteksi yang konsisten, cepat, dan akurat untuk mendukung implementasi pada sistem berbasis visi komputer di bidang pertanian digital.

3 Hasil dan Pembahasan

3.1 Implementasi Sistem

Implementasi sistem deteksi kualitas cabai dilakukan menggunakan Google Colab sebagai lingkungan komputasi berbasis *cloud*. Platform ini dipilih karena menyediakan GPU Tesla T4 yang mampu mempercepat proses pelatihan model *deep learning*. Penelitian ini menggunakan dua model, yaitu YOLOv11 berbasis CNN dan RF-DETR berbasis Transformer.

Tahap persiapan sistem meliputi beberapa langkah utama:

1. Mount Google Drive untuk akses *dataset* dan penyimpanan model
2. Instalasi library Ultralytics dan PyTorch
3. Pengecekan GPU dan CUDA

Dataset cabai hasil anotasi Roboflow dibagi menjadi:

1. Data *training*: 87%
2. Data *validation*: 8%
3. Data *testing*: 4%

Table 3. Pembagian *Dataset*

Jenis Data	PERSENTASE	Jumlah Perkiraan	Tujuan
Train Set	87%	1290 gambar	Digunakan untuk melatih model
Validation Set	8%	122 gambar	Untuk memantau kinerja saat <i>training</i>
Test Set	4%	63 gambar	Untuk evaluasi akhir model

Dataset terdiri dari tiga kelas, yaitu cabai segar, cabai mentah, dan cabai busuk/rusak. Format *dataset* disesuaikan dengan model yang digunakan, yaitu format YOLO untuk YOLOv11 dan format COCO JSON untuk RF-DETR. Model YOLOv11 menggunakan bobot awal yolov11n.pt yang telah dilatih pada *dataset* COCO. Proses *training* dilakukan dengan parameter utama sebagai berikut:

1. *Epoch*: 5
2. *Batch size*: 4
3. Ukuran citra: 640×640

Sementara itu, model RF-DETR dilatih menggunakan metode *fine-tuning*, di mana bobot pretrained tetap digunakan dan layer akhir disesuaikan dengan jumlah kelas.

Distribusi *Dataset*

Distribusi data dilakukan untuk mengetahui keseimbangan jumlah data pada masing-masing kelas sehingga dapat dianalisis apakah *dataset* mengalami imbalance atau tidak.

Tabel 4. Distribusi Data Pelatihan (Train Set) Berdasarkan Kelas

Kelas	JUMLAH GAMBAR	Persentase
Cabai Segar (<i>Fresh</i>)	430	33%
Cabai Mentah (<i>Unripe</i>)	420	32%
Cabai busuk/rusak (<i>Defect</i>)	440	35%

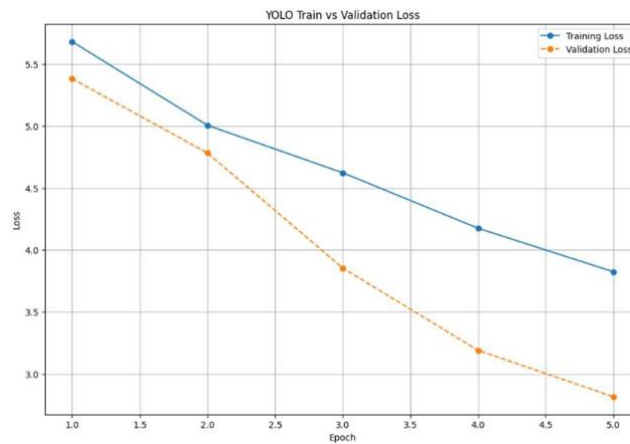
3.2 Hasil *Training* Model

Hasil YOLOv11

Berdasarkan hasil evaluasi, model YOLOv11 memperoleh:

Tabel 5. Hasil Evaluasi YOLOv11

Metrik	NILAI
<i>Precision</i>	0.87
<i>Recall</i>	0.96
mAP50	0.89
mAP50-95	0.82
<i>Latency</i>	±10 ms



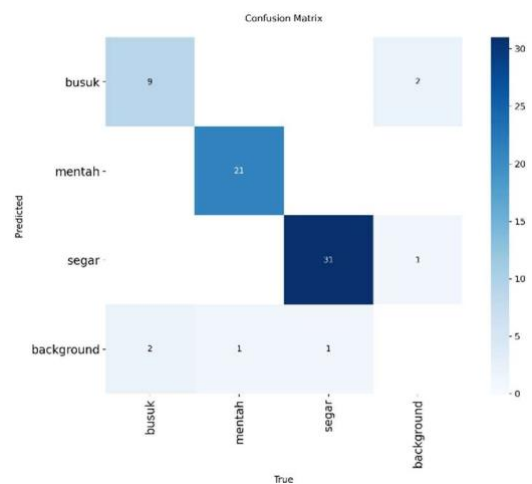
Gambar 3. *Validation Loss YOLOv11*

Nilai *recall* yang tinggi menunjukkan bahwa model mampu mendeteksi hampir seluruh objek cabai pada data uji. *Precision* yang cukup tinggi juga menunjukkan bahwa sebagian besar prediksi yang dihasilkan model merupakan prediksi yang benar.

Grafik *training loss* dan *validation loss* menunjukkan tren penurunan yang stabil, sehingga dapat disimpulkan bahwa proses pembelajaran berjalan dengan baik dan tidak terjadi *overfitting*.

Tabel 6. *Training Loss dan Validation Loss YOLOv11*

<i>Epoch</i>	<i>TRAINING LOSS</i>	<i>VALIDATION LOSS</i>
1	1.245	1.382
2	0.932	1.041
3	0.721	0.854
4	0.603	0.742
5	0.511	0.668



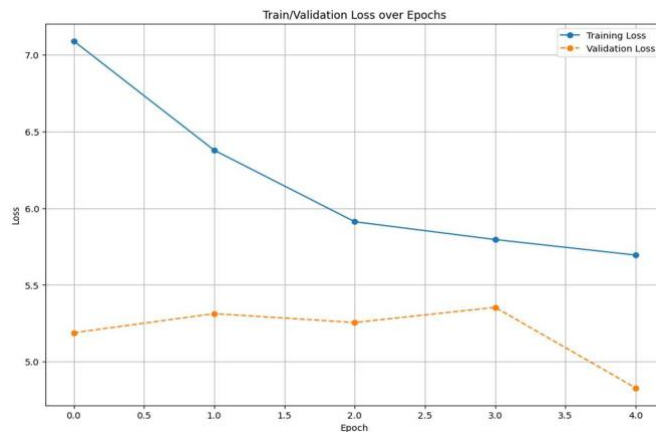
Gambar 4. *Confusion Matrix YOLOv11*

Hasil RF-DETR

Model RF-DETR menghasilkan performa sebagai berikut:

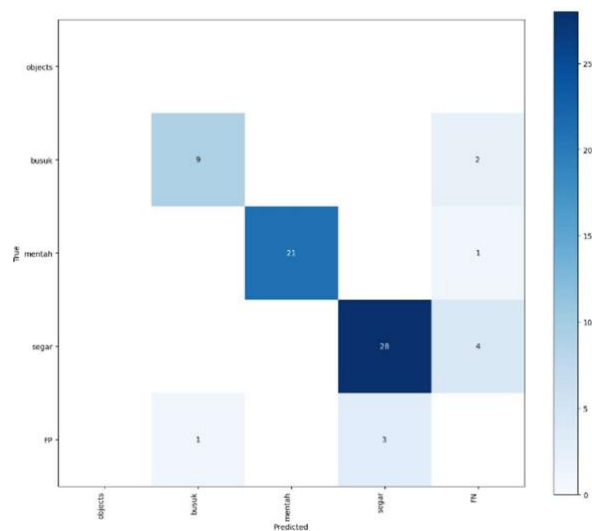
Tabel 7. Hasil Evaluasi RF-DETR

Metrik	NILAI
<i>Precision</i>	0.90
<i>Recall</i>	0.94
mAP50	0.92
mAP50-95	0.86
<i>Latency</i>	18 ms per citra



Gambar 5. Validation Loss RF-DETR

Nilai mAP yang lebih tinggi menunjukkan bahwa RF-DETR memiliki akurasi lebih baik dibandingkan YOLOv11. Model ini juga menunjukkan stabilitas yang lebih baik pada kondisi citra kompleks, seperti objek saling bertumpuk dan pencahayaan kurang optimal.



Gambar 6. Confusion Matrix RF-DETR

Tabel 8. *Training Loss* dan *Validation Loss* RF-DETR

<i>Epoch</i>	<i>TRAINING LOSS</i>	<i>VALIDATION LOSS</i>
1	1.102	1.241
2	0.851	0.973
3	0.663	0.781
4	0.521	0.648
5	0.438	0.574

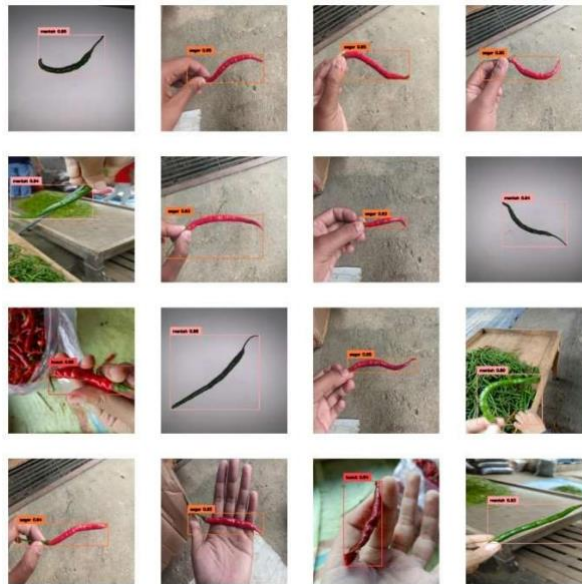
Grafik *training loss* dan *validation loss* pada RF-DETR menunjukkan penurunan yang stabil pada setiap *epoch*. Hal ini menandakan bahwa model mampu mempelajari pola *dataset* dengan baik dan tidak mengalami *overfitting* yang signifikan. Nilai *validation loss* yang terus menurun juga menunjukkan kemampuan generalisasi model yang cukup baik terhadap data validasi.

3.3 Visualisasi Hasil Deteksi

**Gambar 7.** Pengujian Visual Cabai Segar dan Mentah YOLOv11

Pengujian visual dilakukan menggunakan data uji secara acak. Pada model YOLOv11, hasil visualisasi menunjukkan:

1. Cabai segar terdeteksi dengan *confidence* hingga 0.94
2. Cabai mentah terdeteksi dengan *confidence* 0.87
3. Topeng segmentasi (*mask*) terbentuk dengan cukup presisi mengikuti kontur fisik objek.



Gambar 8. Pengujian Visualiasi RF-DETR

Gambar 3.8 Pengujian Visualiasi RF-DETR Sementara pada RF-DETR, hasil visualisasi menunjukkan:

1. Deteksi lebih stabil pada objek bertumpuk
2. Kesalahan deteksi (*false detection*) lebih sedikit
3. Masker segmentasi piksel terpetakan secara lebih presisi pada tepi objek.

3.4 Perbandingan Kinerja Model

Berdasarkan hasil evaluasi, perbandingan kinerja kedua model dapat dilihat sebagai berikut:

Tabel 9. Perbandingan Kinerja YOLOv11 dan RF-DETR

Aspek	YOLOv11	RF-DETR
mAP50	0.89	0.92
mAP50-95	0.82	0.86
Latency	±10 ms	±18 ms
Stabilitas <i>overlap</i>	Cukup	Sangat baik

YOLOv11 unggul dari sisi kecepatan inferensi sehingga cocok untuk aplikasi *real-time*. Sebaliknya, RF-DETR unggul dari sisi akurasi dan stabilitas deteksi.

3.5 Pembahasan

Hasil penelitian menunjukkan bahwa kedua model memiliki keunggulan masing-masing. YOLOv11 lebih unggul dari sisi efisiensi dan kecepatan komputasi, sehingga lebih sesuai untuk sistem *grading* cabai berbasis *real-time*. RF-DETR, di sisi lain, lebih unggul dari sisi akurasi dan ketahanan terhadap kondisi visual yang kompleks.

Dengan demikian, pemilihan model terbaik sangat bergantung pada kebutuhan sistem. Jika sistem mengutamakan kecepatan, maka YOLOv11 menjadi pilihan yang tepat. Namun, jika sistem lebih mengutamakan akurasi, maka RF-DETR lebih direkomendasikan.

3.6 Pengaplikasian System pada Model dengan Streamlit

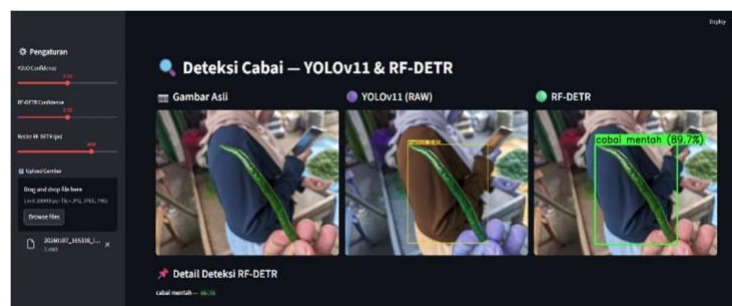
Pengaplikasian sistem dilakukan dengan membangun antarmuka berbasis web menggunakan Streamlit. Streamlit dipilih karena merupakan *framework* Python yang ringan, mudah digunakan, dan sangat cocok untuk menampilkan hasil inferensi model *deep learning* secara interaktif tanpa memerlukan konfigurasi frontend yang kompleks.

Aplikasi ini dirancang untuk memungkinkan pengguna mengunggah citra cabai, kemudian sistem akan secara otomatis melakukan proses deteksi dan klasifikasi kualitas cabai menggunakan model yang telah dilatih, yaitu YOLOv11 dan RF-DETR.



Gambar 9. Pengaplikasian system model dengan streamlit (cabai segar)

Setelah sistem berhasil diimplementasikan menggunakan Streamlit, dilakukan perbandingan kinerja antara model YOLOv11 dan RF-DETR berdasarkan hasil deteksi yang dihasilkan pada citra cabai. Perbandingan ini difokuskan pada aspek ketepatan *mask* segmentasi, konsistensi klasifikasi kelas, serta nilai *confidence* yang dihasilkan oleh masing-masing model.



Gambar 10. Pengaplikasian system model dengan stramlit (cabai mentah)

Berdasarkan hasil pengujian visual, model YOLOv11 mampu mendeteksi objek cabai dengan waktu inferensi yang lebih cepat. Namun, pada beberapa citra terlihat bahwa poligon *mask* yang dihasilkan YOLOv11 masih sedikit kurang menutup seluruh area tepi cabai. Selain itu, nilai *confidence* yang dihasilkan oleh YOLOv11 relatif lebih fluktuatif, khususnya pada kondisi pencahayaan yang kurang ideal.



Gambar 11. Pengaplikasian sistem model dengan streamlit (cabai busuk)

Berbeda dengan YOLOv11, model RF-DETR menunjukkan hasil deteksi yang lebih stabil dan konsisten. *Mask* segmentasi yang dihasilkan oleh RF-DETR terlihat lebih rapat, solid, dan sesuai dengan lekukan bentuk fisik objek cabai, baik pada kondisi cabai segar, mentah, maupun busuk. Selain itu, RF-DETR menghasilkan nilai *confidence* yang relatif tinggi dan konsisten pada berbagai skenario pengujian, yang menunjukkan kemampuan generalisasi model yang lebih baik terhadap variasi data.

Pada pengujian klasifikasi kualitas cabai, RF-DETR juga menunjukkan keunggulan dibandingkan YOLOv11. Model ini mampu membedakan kategori kualitas cabai dengan lebih jelas, ditunjukkan oleh label kelas yang sesuai dengan kondisi visual objek serta nilai *confidence* yang tinggi, seperti pada deteksi cabai segar, cabai mentah, dan cabai busuk. Hal ini menunjukkan bahwa RF-DETR lebih andal dalam menangkap fitur visual yang kompleks dibandingkan YOLOv11.

Berdasarkan hasil perbandingan tersebut, dapat disimpulkan bahwa meskipun YOLOv11 unggul dari segi kecepatan inferensi, RF-DETR memiliki performa yang lebih baik dalam hal akurasi deteksi dan kestabilan klasifikasi. Oleh karena

itu, RF-DETR lebih sesuai digunakan sebagai model utama dalam sistem deteksi dan klasifikasi kualitas cabai, sedangkan YOLOv11 dapat dimanfaatkan sebagai model pembandingan atau *baseline*.

4 Kesimpulan dan saran

4.1 Kesimpulan

Penelitian ini bertujuan untuk mengembangkan serta mengevaluasi sistem deteksi dan klasifikasi kualitas cabai pascapanen menggunakan pendekatan *computer vision* berbasis *deep learning*, serta melakukan analisis komparatif terhadap dua arsitektur *instance segmentation*, yaitu YOLOv11 dan RF-DETR. Berdasarkan hasil eksperimen kedua model menunjukkan performa yang baik dalam mensegmentasi dan mengklasifikasikan cabai ke dalam beberapa kategori kualitas.

Model YOLOv11 memiliki keunggulan dalam kecepatan inferensi dengan waktu sekitar ± 10 ms per citra, sehingga lebih sesuai untuk implementasi sistem *real-time*. Selain itu, YOLOv11 juga menunjukkan proses pelatihan yang stabil dengan penurunan *training loss* dan *validation loss* yang konsisten pada setiap *epoch*. Sementara itu, model RF-DETR menunjukkan performa yang lebih tinggi dalam hal akurasi, ditunjukkan oleh nilai *precision*, *recall*, dan *mean Average Precision* (mAP) yang lebih baik dibandingkan YOLOv11. RF-DETR juga memiliki kemampuan lokalisasi *mask* yang lebih stabil pada kondisi citra kompleks, seperti objek yang saling bertumpuk, pencahayaan tidak merata, serta variasi bentuk dan tekstur cabai.

Hasil penelitian juga menunjukkan bahwa penerapan *hyperparameter tuning* dan augmentasi data mampu membantu meningkatkan performa model serta mengurangi risiko *overfitting*. Berdasarkan analisis *training loss* dan *validation loss*, kedua model menunjukkan kemampuan generalisasi yang cukup baik terhadap data validasi karena tidak terdapat gap *loss* yang terlalu besar selama proses pelatihan.

Implikasi dari penelitian ini menunjukkan bahwa penerapan teknologi *deep learning* dalam bidang pertanian, khususnya pada proses *grading* komoditas hortikultura, dapat meningkatkan efisiensi, konsistensi, dan objektivitas dibandingkan metode manual. Sistem segmentasi berbasis visi komputer memungkinkan proses identifikasi kualitas cabai dilakukan secara lebih cepat dan akurat hingga ke batas fisik objek. Selain itu, perbandingan antara model YOLOv11 dan RF-DETR memberikan gambaran bahwa pemilihan model sangat bergantung pada kebutuhan sistem, baik dari segi kecepatan maupun akurasi.

Dengan demikian, hasil penelitian ini dapat menjadi referensi dalam pengembangan sistem otomatis berbasis visi komputer untuk mendukung proses produksi, sortasi, dan distribusi hasil panen pada sektor agrikultur modern.

4.2 Saran

Penelitian ini masih memiliki beberapa keterbatasan yang dapat dikembangkan lebih lanjut pada penelitian berikutnya. Pertama, jumlah *dataset* yang digunakan masih terbatas dan kondisi pengambilan citra belum sepenuhnya merepresentasikan variasi lingkungan yang lebih luas, sehingga kemampuan generalisasi model masih perlu ditingkatkan. Oleh karena itu, penelitian selanjutnya disarankan menggunakan *dataset* yang lebih besar dan lebih beragam, baik dari sisi pencahayaan, latar belakang, maupun kondisi objek cabai.

Kedua, penelitian ini hanya membandingkan dua arsitektur *instance segmentation*, yaitu YOLOv11 dan RF-DETR. Penelitian selanjutnya dapat mengeksplorasi penggunaan model *deep learning* lain yang lebih mutakhir, seperti YOLOv12, RT-DETR, EfficientDet, maupun kombinasi beberapa metode (*ensemble learning*) untuk memperoleh performa deteksi yang lebih optimal.

Selain itu, proses *hyperparameter tuning* pada penelitian ini masih dilakukan secara manual. Penelitian berikutnya disarankan menerapkan metode optimasi hyperparameter otomatis, seperti *Grid Search*, *Random Search*, atau *Bayesian Optimization* agar diperoleh konfigurasi parameter yang lebih optimal.

Pengembangan selanjutnya juga dapat diarahkan pada integrasi sistem deteksi dengan perangkat keras berbasis *Internet of Things* (IoT), kamera *real-time*, maupun conveyor otomatis sehingga sistem *grading* cabai dapat diterapkan secara langsung pada lingkungan industri pertanian dan distribusi hasil panen.

Daftar Pustaka

- [1] A. Lukas *et al.*, “Fresh Chili Agribusiness: Opportunities and Problems in Indonesia,” in *Sustainable Development*, vol. 16, O. Özçatalbaş, Ed., IntechOpen, 2024. doi: 10.5772/intechopen.112786.
- [2] M. Bhattarai and J. Mariyono, “The economic aspects of chilli production in Central Java,” *Econ. J. Emerg. Mark.*, vol. 8, no. 2, pp. 85–97, Oct. 2016, doi: 10.20885/ejem.vol8.iss2.art1.
- [3] N. Khuriyati, A. P. Pamungkas, and A. A. P., “The Sorting and Grading of Red Chilli Peppers (*Capsicum annuum* L.) Using Digital Image Processing,” *Int. J. Agric. Environ. Sci.*, vol. 6, no. 4, pp. 17–23, Aug. 2019, doi: 10.14445/23942568/IJAES-V6I4P104.
- [4] Sudianto, Y. Herdiyeni, A. Haristu, and M. Hardhienata, “Chilli Quality Classification using Deep Learning,” in *2020 International Conference on Computer Science and Its Application in Agriculture (ICOSICA)*, Bogor, Indonesia: IEEE, Sep. 2020, pp. 1–5. doi: 10.1109/ICOSICA49951.2020.9243176.
- [5] R. Sapkota *et al.*, “YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series,” *Artif. Intell. Rev.*, vol. 58, no. 9, p. 274, Jun. 2025, doi: 10.1007/s10462-025-11253-3.
- [6] A. A. Abdul Manan, M. A. Mohd Razman, I. Mohd Khairuddin, and M. N. A. Shapiee, “Chili Plant Classification using Transfer Learning models through Object Detection,” *MEKATRONIKA*, vol. 2, no. 2, pp. 23–27, Dec. 2020, doi: 10.15282/mekatronika.v2i2.6743.
- [7] M. Anwar, Y. Kristian, and E. Setyati, “Klasifikasi Penyakit Tanaman Cabai Rawit Dilengkapi Dengan Segmentasi Citra Daun dan Buah Menggunakan Yolo v7,” *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 6, no. 1, pp. 540–548, Jun. 2023, doi: 10.31539/intecom.v6i1.6071.
- [8] H. Zhu and X. Xie, “Enhanced YOLOv11n for small object detection in UAV imagery: higher accuracy with fewer parameters,” *Sci. Rep.*, vol. 16, no. 1, p. 5536, Jan. 2026, doi: 10.1038/s41598-026-35301-2.
- [9] F. Maulana, “Detection of Papaya Fruit Ripeness Using the YOLOv11s Algorithm Based on Digital Images,” vol. 6, 2026.
- [10] N. M. Handayani and Y. Azhar, “Implementation of YOLOv11 for Detection and Identification of Strawberry Ripeness Stages,” vol. 10, no. 2.
- [11] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, “End-to-End Object Detection with Transformers,” May 28, 2020, *arXiv*: arXiv:2005.12872. doi: 10.48550/arXiv.2005.12872.
- [12] I. Robinson, P. Robicheaux, M. Popov, D. Ramanan, and N. Peri, “RF-DETR: Neural Architecture Search for Real-Time Detection Transformers,” 2025, *arXiv*. doi: 10.48550/ARXIV.2511.09554.