PENDEKATAN SEMI-SUPERVISED LEARNING UNTUK PROSES PELABELAN PRODUK RETAIL UNTUK MENINGKATKAN PERFORMA DETEKSI OBJEK



Disusun Oleh:

N a m a : Pandu Avilba
NIM : 21523018

PROGRAM STUDI INFORMATIKA – PROGRAM SARJANA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
2025

HALAMAN PENGESAHAN DOSEN PEMBIMBING

PENDEKATAN SEMI-SUPERVISED LEARNING UNTUK PROSES PELABELAN PRODUK RETAIL UNTUK MENINGKATKAN PERFORMA DETEKSI OBJEK

TUGAS AKHIR

Disusun Oleh:

N a m a	: Pandu Avilba
NIM	: 21523018

Yogyakarta, 22 Oktober 2025

Pembimbing,

(Arrie Kurniawardhani, S.Si., M.Kom.)

HALAMAN PENGESAHAN DOSEN PENGUJI

PENDEKATAN SEMI-SUPERVISED LEARNING UNTUK PROSES PELABELAN PRODUK RETAIL UNTUK MENINGKATKAN PERFORMA DETEKSI OBJEK

TUGAS AKHIR

Telah dipertahankan di depan sidang penguji sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer dari Program Studi Informatika – Program Sarjana di Fakultas Teknologi Industri Universitas Islam Indonesia

Yogyakarta, 5 November 2025

Tim Penguji

Arrie Kurniawardhani, S.Si., M.Kom.

Anggota 1

Hari Setiaji, S.Kom., M.Eng.

Anggota 2

Mukhammad Andri Setiawan, S.T., M.Sc., Ph.D.

Mengetahui,

Ketua Program Studi Informatika – Program Sarjana

Fakultas Teknologi Industri

Universitas Islam Indonesia

(Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D.)

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini:

Nama : Pandu Avilba

NIM : 21523018

Tugas akhir dengan judul:

PENDEKATAN SEMI-SUPERVISED LEARNING UNTUK PROSES PELABELAN PRODUK RETAIL UNTUK MENINGKATKAN PERFORMA DETEKSI OBJEK

Menyatakan bahwa seluruh komponen dan isi dalam tugas akhir ini adalah hasil karya saya sendiri. Apabila di kemudian hari terbukti ada beberapa bagian dari karya ini adalah bukan hasil karya sendiri, tugas akhir yang diajukan sebagai hasil karya sendiri ini siap ditarik kembali dan siap menanggung risiko dan konsekuensi apapun.

Demikian surat pernyataan ini dibuat, semoga dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 22 Oktober 2025



(Pandu Avilba)

HALAMAN PERSEMBAHAN

Pada halaman persembahan ini, saya mengucapkan terima kasih sebesar-besarnya kepada Allah Subhanahu Wa Ta'ala yang telah memberikan saya kesempatan menempuh studi Informatika di Universitas Islam Indonesia. Saya juga mengucapkan terima kasih kepada kedua orang tua saya yang telah mengizinkan saya untuk belajar di Universitas Islam Indonesia. Terkahir, tak lupa saya ucapkan terima kasih kepada seluruh dosen dan teman-teman di Universitas Islam Indonesia yang telah memberikan saya inspirasi sehingga bisa menjadi pribadi yang lebih baik kedepannya.

HALAMAN MOTO

“Dan Dia mendapatimu sebagai seorang yang bingung, lalu Dia memberikan petunjuk.”
(Q.S Ad-Duha: 7)

KATA PENGANTAR

Puji Syukur kepada Allah SWT. karena atas rahmat-Nya penulis dapat menyelesaikan laporan tugas akhir yang berjudul "Pendekatan Semi-Supervised Learning untuk Proses Pelabelan Produk Retail guna Meningkatkan Performa Deteksi Objek" dengan baik.

Selama proses penyusunan laporan tugas akhir ini, penulis mendapatkan bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan rasa terima kasih sebesar-besarnya kepada:

1. Saparuddin dan Fitrialis selaku orang tua penulis yang selalu memberikan dukungan baik secara moral maupun finansial.
2. Bapak Prof. Fathul Wahid, S.T., M.Sc., Ph.D., selaku Rektor Universitas Islam Indonesia.
3. Bapak Dr. Raden Teduh Dirgahayu, S.T., M.Sc., selaku Ketua Jurusan Informatika.
4. Bapak Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D., selaku Ketua Program Studi Informatika Program Sarjana.
5. Ibu Arrie Kurniawardhani, S.Si. M.Kom., selaku Dosen Pembimbing Tugas Akhir, yang telah meluangkan waktu dan pikirannya untuk memberikan bimbingan dan arahan khususnya dalam penelitian dan penulisan Tugas Akhir ini.
6. Bapak Andhik Budi Cahyono, S.T., M.T. selaku dosen pembimbing akademik.
7. Seluruh Dosen dan Civitas Akademika di lingkungan Universitas Islam Indonesia.
8. Seluruh rekan saya Informatika 21.
9. Ucapan terima kasih kepada beberapa pihak yang telah membantu atau mendukung penyelesaian penulisan laporan atau pelaksanaan penelitian tugas akhir.

Semoga segala usaha yang dilakukan dalam penelitian ini bernilai di sisi Allah SWT. Penunis menyadari bahwa masih banyak kekurangan yang ada dalam laporan tugas akhir ini. Oleh sebab itu, segala saran dan kritik dibutuhkan untuk kedepannya. Akhir kata, semoga laporan tugas akhir ini memberikan manfaat yang besar-besarnya kepada semua pihak.

Yogyakarta, 21 Oktober 2025



(Pandu Avilba)

SARI

Penelitian ini mengimplementasikan pendekatan semi-supervised learning berbasis pseudo-labeling untuk meningkatkan performa deteksi objek produk retail dalam konteks On-Shelf Availability (OSA) dengan memanfaatkan kombinasi data berlabel dan tidak berlabel. Latar belakang penelitian adalah tingginya biaya dan waktu yang dibutuhkan untuk anotasi manual pada sistem deteksi objek yang melibatkan ratusan kategori produk dengan variasi visual tinggi. Penelitian menggunakan arsitektur YOLOv8n dan YOLOv11n yang dilatih pada dataset berisi 918 gambar dengan 174 kelas produk retail (total 11.161 objek), yang dibagi menjadi empat proporsi data berlabel: 20%, 40%, 60%, dan 80%.

Metodologi penelitian mencakup lima tahapan: pengumpulan dan persiapan data dengan anotasi manual format YOLO, supervised training untuk menghasilkan teacher model, pseudo-labeling dengan confidence threshold 0.7, semi-supervised training dengan tiga strategi inisialisasi (pretrained, teacher model, dan pseudo model), serta evaluasi menggunakan metrik precision, recall, mAP@0.5, mAP@0.5:0.95, F1-score, IoU, dan inference time.

Hasil penelitian menunjukkan bahwa proporsi 60% data berlabel menghasilkan performa optimal, dengan model Pretrained YOLOv11 mencapai mAP@0.5 sebesar 0.951 mendekati performa fully supervised dengan namun menghemat 40% effort anotasi.

Penelitian membuktikan bahwa semi-supervised learning efektif mengurangi ketergantungan terhadap anotasi manual dengan proporsi 60% sebagai titik optimal antara performa dan efisiensi, memberikan rekomendasi praktis untuk implementasi sistem deteksi objek retail yang cost-effective dengan performa mendekati fully supervised.

Kata kunci: semi-supervised learning, pseudo-labeling, deteksi objek, YOLO, produk retail, on-shelf availability, deep learning

GLOSARIUM

Anotasi	Proses pemberian label atau bounding box pada objek dalam gambar untuk keperluan pelatihan model machine learning.
Bounding Box	Kotak pembatas yang menandai lokasi objek dalam gambar, didefinisikan oleh koordinat pusat, lebar, dan tinggi.
Computer Vision	bidang ilmu komputer yang memungkinkan mesin untuk memahami dan menginterpretasi informasi visual dari gambar atau video.
Confidence Score	Nilai probabilitas yang menunjukkan tingkat keyakinan model terhadap prediksi keberadaan dan kelas suatu objek.
Epoch	Satu siklus lengkap pelatihan model di mana seluruh dataset training telah diproses satu kali.
Yolo	Arsitektur neural network untuk deteksi objek yang memproses gambar dalam satu tahap untuk kecepatan tinggi.

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PENGESAHAN DOSEN PEMBIMBING ii
HALAMAN PENGESAHAN DOSEN PENGUJI iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR iv
HALAMAN PERSEMBAHAN v
HALAMAN MOTO vi
KATA PENGANTAR vii
SARI viii
GLOSARIUM ix
DAFTAR ISI x
DAFTAR TABEL xii
DAFTAR GAMBAR xiii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 2
1.3 Tujuan Penelitian 2
1.4 Batasan Masalah 2
1.5 Manfaat Penelitian 2
1.6 Sistematika Penulisan 3
BAB II LANDASAN TEORI 4
2.1 On-Shelf Availability 4
2.2 Deteksi Objek 4
2.3 You Only Look Once (YOLO) 5
2.3.1 Yolov11 6
2.3.2 Yolov8 8
2.4 Semi-Supervised Learning 9
2.5 Pseudo-Labeling 9
2.6 Metrik Evaluasi 10
2.6.1 Precision 10
2.6.2 Recall 10
2.6.3 Mean Average Precision (mAP@0.5) 11
2.6.4 Mean Average Precision (mAP@0.5:0.95) 11
2.6.5 Inference time 11
2.6.6 Intersection over Union (IoU) 12
2.6.7 F1-Score 12
2.6.8 True Positive (TP) 13
2.6.9 False Positive (FP) 13
2.6.10 False Negative (FN) 13
2.6.11 Box Loss 13
2.6.12 Class Loss 14
2.6.13 Distribution Focal Loss (DFL Loss) 14
2.7 Studi Literatur 14
BAB III METODOLOGI 16
3.1 Pengumpulan dan Persiapan Data 16
3.2 Teacher Model Training 25
3.3 Pseudo-labeling 25
3.4 Student Model Training 26

3.5 Evaluation 27
BAB IV HASIL DAN PEMBAHASAN 28
4.1 Implementasi dan Hasil Pengumpulan dan Persiapan Data 28
4.2 Implementasi dan Hasil Teacher Model 31
4.3 Implementasi dan Hasil Pseudo-Labeling 41
4.3.1 Kualitas Pseudo-Label 42
4.3.2 Hasil Training Pseudo Model 49
4.4 Implementasi dan Hasil Student Model 58
4.5 Analisis dan Pembahasan Performa 102
4.5.1 Pengaruh Proporsi Data Berlabel terhadap Performa Model 102
4.5.2 Perbandingan Strategi Inisialisasi Model 102
4.5.3 Kualitas Pseudo-Label dan Dampaknya terhadap Performa 103
4.5.4 Perbandingan Visual Testing Semi Supervised Learning dengan Fully Supervised 103
4.5.5 Efektivitas Semi-Supervised Learning 104
BAB V KESIMPULAN DAN SARAN 105
5.1 Kesimpulan 105
5.2 Saran 105
DAFTAR PUSTAKA 106
LAMPIRAN 110

DAFTAR TABEL

Tabel 2. 1 Studi Literatur 14
Tabel 3. 1 Daftar Kelas 18
Tabel 4. 1 Proporsi Dataset 31
Tabel 4. 2 Hasil Train Fully Supervised 32
Tabel 4. 3 Hasil Training Teacher Model 40
Tabel 4. 4 Metrik Evaluasi Pseudo-label 47
Tabel 4. 5 Hasil Training Pseudo Model 57
Tabel 4. 6 Hasil Training Student model 20% 64
Tabel 4. 7 Hasil Visual Semi Supervised 20% 66
Tabel 4. 8 Hasil Training Student model 40% 75
Tabel 4. 9 Hasil Visual Semi Supervised 40% 77
Tabel 4. 10 Hasil Training Student model 60% 85
Tabel 4. 11 Hasil Visual Semi Supervised 60% 87
Tabel 4. 12 Hasil Training Student model 80% 95
Tabel 4. 13 Hasil Visual Semi Supervised 80% 97

DAFTAR GAMBAR

Gambar 2. 1 Contoh Implementasi Deteksi Objek 5
Gambar 2. 2 Arsitektur Yolov11 6
Gambar 2. 3 Arsitektur Yolov8 8
Gambar 3. 1 Tahapan Penelitian 16
Gambar 3. 2 Contoh dataset 17
Gambar 3. 3 Label.txt 18
Gambar 3. 4. Struktur Folder Dataset 24
Gambar 4. 1 Kode Program Split Dataset 30
Gambar 4. 2 Kode Program Training Model 32
Gambar 4. 3 Teacher Model 20% Yolov11 33
Gambar 4. 4 Teacher model 20% Yolov8 33
Gambar 4. 5 Teacher Model 40% Yolov11 34
Gambar 4. 6 Teacher Model 40% Yolov8 35
Gambar 4. 7 Teacher Model 60% Yolov11 36
Gambar 4. 8 Teacher Model 60% Yolov8 37
Gambar 4. 9 Teacher Model 80% Yolov11 38
Gambar 4. 10 Teacher Model 80% Yolov8 39
Gambar 4. 11 Kode Program Generate Pseudo-label 41
Gambar 4. 12 Perbandingan Visual Pseudo-label 20% 43
Gambar 4. 13 Perbandingan Visual Pseudo-label 40% 44
Gambar 4. 14 Perbandingan Visual Pseudo-label 60% 45
Gambar 4. 15 Perbandingan Visual Pseudo-label 80% 46
Gambar 4. 16 Pseudo Model 20% Yolov11 50
Gambar 4. 17 Pseudo Model 20% Yolov8 51
Gambar 4. 18 Pseudo Model 40% Yolov11 52
Gambar 4. 19 Pseudo Model 40% Yolov8 53
Gambar 4. 20 Pseudo model 60% Yolov11 54
Gambar 4. 21 Pseudo Model 60% Yolov8 55
Gambar 4. 22 Pseudo model 80% Yolov11 56
Gambar 4. 23 Pseudo model 80% Yolov8 56
Gambar 4. 24 Student Model 20% dengan Pretrained Model Yolov11 59
Gambar 4. 25 Student Model 20% dengan Pretrained Model Yolov8 60

Gambar 4. 26 Student Model 20% dengan Teacher Model Yolov11 60
Gambar 4. 27 Student Model 20% dengan Teacher Model Yolov8 61
Gambar 4. 28 Student Model 20% dengan Pseudo Model Yolov11 62
Gambar 4. 29 Student Model 20% dengan Pseudo Model Yolov8 63
Gambar 4. 30 Visual Test Semi Supervised 20% 66
Gambar 4. 31 Student Model 40% dengan Pretrained Model Yolov11 71
Gambar 4. 32 Student Model 40% dengan Pretrained Model Yolov8 71
Gambar 4. 33 Student Model 40% dengan Teacher Model Yolov11 72
Gambar 4. 34 Student Model 40% dengan Teacher Model Yolov8 73
Gambar 4. 35 Student Model 40% dengan Pseudo Model Yolov11 74
Gambar 4. 36 Student Model 40% dengan Pseudo Model Yolov8 74
Gambar 4. 37 Visual Test Semi Supervised 40% 77
Gambar 4. 38 Student Model 60% dengan Pretrained Model Yolov11 81
Gambar 4. 39 Student Model 60% dengan Pretrained Model Yolov8 82
Gambar 4. 40 Student Model 60% dengan Teacher Model Yolov11 82
Gambar 4. 41 Student Model 60% dengan Teacher Model Yolov8 83
Gambar 4. 42 Student Model 60% dengan Pseudo Model Yolov11 84
Gambar 4. 43 Student Model 60% dengan Pseudo Model Yolov8 84
Gambar 4. 44 Visual Test Semi Supervised 60% 87
Gambar 4. 45 Student Model 80% dengan Pretrained Model Yolov11 91
Gambar 4. 46 Student Model 80% dengan Pretrained Model Yolov8 92
Gambar 4. 47 Student Model 80% dengan Teacher Model Yolov11 92
Gambar 4. 48 Student Model 80% dengan Teacher Model Yolov8 93
Gambar 4. 49 Student Model 80% dengan Pseudo Model Yolov11 94
Gambar 4. 50 Student Model 80% dengan Pseudo Model Yolov8 94
Gambar 4. 51 Visual Test Semi Supervised 80% 97

BAB I
PENDAHULUAN

1.1 Latar Belakang

Dalam industri retail, ketersediaan produk di rak (*On-Shelf Availability*/OSA) merupakan aspek penting dalam operasional toko karena berpengaruh langsung terhadap kepuasan pelanggan dan potensi penjualan. Dalam konteks ini, sistem berbasis *computer vision*, khususnya deteksi objek pada gambar rak toko, muncul sebagai solusi potensial untuk memantau ketersediaan produk, mengidentifikasi kekosongan rak, serta mengawasi tata letak barang (Cai et al. 2021)(Wei et al. 2020). Namun, penerapan sistem semacam ini menghadapi tantangan besar terkait kebutuhan data pelatihan beranotasi. Berbeda dengan klasifikasi gambar, deteksi objek memerlukan anotasi berupa *bounding box* untuk setiap objek, yang dilakukan secara manual dan memakan banyak waktu (G. Li et al. 2022). Tantangan ini semakin kompleks dalam lingkungan retail yang memiliki variasi kemasan, kemiripan visual antarproduk, serta perbedaan kondisi pencahayaan dan sudut pandang kamera. Kebutuhan untuk mencantumkan atribut produk secara detail seperti merek, rasa, ukuran, dan jenis juga menambah biaya dan kompleksitas proses pelabelan (Guimarães et al. 2023).

Untuk mengatasi tantangan keterbatasan data berlabel dan tingginya biaya anotasi pada tugas deteksi objek, pendekatan *semi-supervised learning* (SSL) telah menjadi perhatian dalam beberapa tahun terakhir (Ouali, Hudelot, and Tami 2020). SSL menawarkan solusi efisien dengan memanfaatkan kombinasi sebagian kecil data berlabel untuk pelatihan model (Sohn et al. 2022). Strategi ini memungkinkan peningkatan performa model secara signifikan, tanpa ketergantungan pada proses pelabelan manual yang mahal dan memakan waktu (Shehzadi et al. 2024). Salah satu teknik dalam SSL adalah *pseudo-labeling*, di mana model awal yang dilatih secara *supervised* digunakan untuk menghasilkan *pseudo-label* pada data tak berlabel (Qi, Nguyen, and Yan 2024)(Yang et al. 2023). Teknik ini relevan dalam konteks retail, yang umumnya melibatkan gambar dengan objek yang rapat dan kemiripan visual tinggi antar produk. Dalam skenario seperti ini, metode *semi-supervised object detection* (SSOD) menjadi pendekatan yang tepat, karena mampu memaksimalkan pemanfaatan data tak berlabel dalam meningkatkan kinerja deteksi objek (Qi et al. 2024).

Oleh karena itu, penelitian ini menggabungkan pendekatan SSL berbasis *pseudo-labeling* dengan menggunakan arsitektur *You Only Look Once* (YOLO) yaitu versi YOLOv8, yang

menurut (Santos Júnior, Paixão, and Alvarez 2025) menunjukkan performa lebih baik dibandingkan YOLOv9 dan YOLOv10. Selain itu, penelitian ini juga memanfaatkan YOLOv11 sebagai arsitektur terkini yang dirancang untuk memberikan keseimbangan optimal antara kecepatan inferensi, akurasi deteksi, dan efisiensi komputasi (Khanam and Hussain 2024). Meskipun metode *semi-supervised* dan *pseudo-labeling* telah terbukti memiliki potensi besar dalam meningkatkan efisiensi sistem deteksi objek di sektor retail, Sebagian besar studi SSOD sebelumnya masih menggunakan YOLOv5 - YOLOv7 atau *Faster R-CNN*, sehingga penelitian pada YOLOv8 dan YOLOv11 masih sangat terbatas.

Oleh karena itu, penelitian ini bertujuan untuk menerapkan pendekatan *semi-supervised* dengan *pseudo-labeling* pada YOLOv8 dan YOLOv11 dan mengevaluasi performanya dalam berbagai proporsi data berlabel (20%, 40%, 60%, dan 80%). Penelitian ini juga membandingkan hasil *pseudo-labeling* dengan *ground truth* untuk menilai kualitas *pseudo-label* yang dihasilkan.

1.2 Rumusan Masalah

a. Bagaimana perbandingan performa antara model yang dilatih secara *supervised* dengan model *semi-supervised* dalam konteks deteksi produk retail?
b. Berapa rasio optimal antara data berlabel dan tidak berlabel pada model semi-supervised learning yang diperlukan untuk mencapai performa deteksi yang optimal dalam skenario retail?

1.3 Tujuan Penelitian

a. Menganalisis dan membandingkan performa model *semi-supervised learning* dengan *supervised learning* dalam konteks deteksi produk retail.
b. Menentukan konfigurasi optimal rasio data berlabel dan tidak berlabel pada model semi-supervised learning untuk mencapai efisiensi maksimal dalam proses pelabelan produk retail.

1.4 Batasan Masalah

a. Penelitian dibatasi pada penggunaan arsitektur YOLOv8n dan YOLOv11n sebagai model deteksi objek.
b. *Dataset* yang digunakan terbatas pada produk retail dengan kategori susu bubuk.
c. Penelitian tidak mencakup implementasi *real-time deployment* pada sistem retail.

1.5 Manfaat Penelitian

a. Kontribusi ilmiah dalam pengembangan metode *semi-supervised learning* untuk deteksi objek retail.

b. Sistem deteksi produk yang lebih efisien dan efektif untuk industri retail.

1.6 Sistematika Penulisan

Sistematika Penulisan dari Penelitian ini terdiri dari lima bab, yakni:

a. BAB 1 PENDAHULUAN

Bab ini membahas latar belakang, rumusan masalah, tujuan masalah, batasan masalah, manfaat penelitan, dan sistematika penulisan.

b. BAB II LANDASAN TEORI

Bab ini menjelaskan teori dasar yang berkaitan dengan dengan penelitian ini, serta membahas penelitian sebelumnya yang relevan dengan penelitian ini.

c. BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan langkah-langkah penelitian, meliputi pengumpulan dan persiapan data, *teacher model training, pseudo-labeling, student model training*, dan evaluasi

d. BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil implementasi pendekatan *semi-supervised learning* pada *dataset* objek retail, serta menganalisis hasil pelatihan dan hasil pengujian model.

e. BAB V KESIMPULAN DAN SARAN

Bab ini menyajikan kesimpulan dari seluruh proses penelitian yang telah dilakukan dan memberikan saran untuk penelitian lebih lanjut di masa yang akan datang.

BAB II
LANDASAN TEORI

2.1 On-Shelf Availability

On-Shelf Availability (OSA) merujuk pada ketersediaan barang di rak yang siap dipilih dan dibeli oleh konsumen, sedangkan *Out of Stock* (OOS) didefinisikan sebagai persentase barang yang tidak tersedia di rak toko retail pada waktu tertentu. Dengan kata lain, OSA mengukur kemungkinan suatu produk tersedia dalam kondisi layak jual ketika dicari konsumen, sedangkan OOS menggambarkan situasi sebaliknya, yaitu ketika produk yang dicari tidak ditemukan, tidak layak jual, atau tidak berada di lokasi yang diharapkan (Rizzotti and Russo 2025).

Dari sisi penyebab, OSA dan OOS umumnya dipengaruhi oleh faktor operasional, seperti pemesanan, alokasi ruang rak, serta pencatatan persediaan. Eksekusi toko yang kurang optimal dapat mengurangi efektivitas rantai pasok dan strategi pemasaran, sehingga menurunkan ketersediaan produk di rak (Moussaoui et al. 2016). Kondisi ini sangat merugikan karena dapat menyebabkan hilangnya potensi penjualan dan berkurangnya kepuasan pelanggan. Salah satu bentuk OOS yang sering terjadi adalah *Shelf Out of Stock* (SOOS), yakni kondisi ketika produk tersedia di gudang, tetapi belum dipajang di rak penjualan akibat keterlambatan atau kesalahan pengisian ulang (Milella et al. 2020). Upaya seperti peningkatan frekuensi pemeriksaan rak dan pelatihan staf sering kali dilakukan untuk mengurangi risiko SOOS (Milella et al. 2020).

2.2 Deteksi Objek

Deteksi objek merupakan salah satu tugas utama dalam bidang *computer vision*, yang bertujuan menemukan objek dari kelas target tertentu dengan lokalisasi yang tepat pada sebuah gambar, serta memberikan label kelas yang sesuai untuk setiap objek yang terdeteksi. Berbeda dengan klasifikasi gambar yang hanya menentukan jenis objek dalam gambar secara keseluruhan, Deteksi objek tidak hanya berfokus pada pengenalan kategori objek, tetapi juga memprediksi lokasi setiap objek melalui penggunaan bounding box (Wu, Sahoo, and Hoi 2020). Contoh implementasi object detection dapat dilihat pada Gambar 2. 1 Contoh Implementasi .



Gambar 2. 1 Contoh Implementasi Deteksi Objek

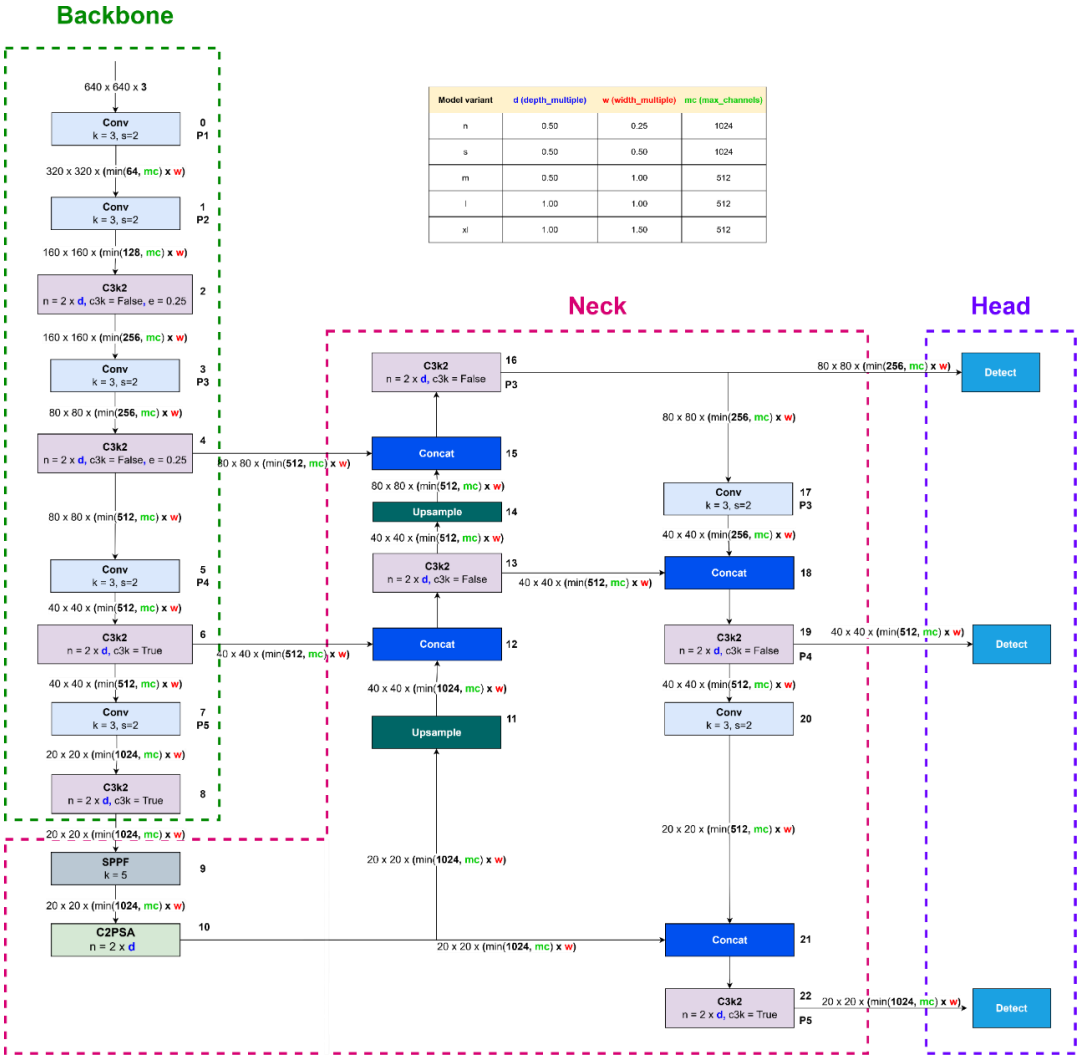
Secara umum, kerangka kerja deteksi objek berbasis deep learning secara umum dapat dibagi menjadi dua kelompok utama yaitu *two-stage detectors*, seperti *Region-based CNN (R-CNN)* beserta variannya (Girshick 2015), dan *one-stage detectors*, seperti YOLO (Redmon et al. 2016). *Two-stage detectors* pertama-tama menggunakan proposal generator untuk menghasilkan sejumlah proposal yang terbatas dan mengekstraksi fitur dari setiap proposal, kemudian dilanjutkan dengan *region classifier* yang memprediksi kategori dari *region* yang diusulkan. Sementara itu, *one-stage detectors* secara langsung melakukan prediksi kategori objek pada setiap lokasi di *feature maps* tanpa melalui tahap klasifikasi region secara bertahap (Wu et al. 2020) Pendekatan *one-stage*, khususnya keluarga YOLO, banyak digunakan dalam penelitian terkini karena menawarkan kecepatan inferensi yang tinggi dengan akurasi kompetitif, sehingga sesuai untuk aplikasi real-time seperti deteksi produk retail..

2.3 You Only Look Once (YOLO)

YOLO merupakan salah satu metode deteksi objek yang paling banyak digunakan dalam *computer vision*. Metode ini termasuk ke dalam kategori *one-stage detectors*, di mana proses deteksi dan klasifikasi objek dilakukan dalam satu tahap pemrosesan (Wu et al. 2020). Keunggulan utama YOLO terletak pada ukuran model yang relatif ringan dan kecepatan inferensinya yang tinggi. Struktur YOLO cukup sederhana karena dapat langsung menghasilkan posisi dan kategori *bounding box* melalui jaringan saraf. Kecepatan YOLO tergolong tinggi karena YOLO hanya memerlukan sekali proses *forward pass* untuk menghasilkan prediksi *bounding box* dan label kelas, sehingga dapat digunakan untuk deteksi

objek *real-time*, termasuk pada video (Jiang et al. 2021). Secara umum, arsitektur YOLO terdiri atas tiga komponen utama, yaitu *backbone* sebagai pengekstrak fitur, *neck* sebagai penggabung informasi multiskala, dan *head* sebagai komponen prediksi.

2.3.1 Yolov11



Gambar 2. 2 Arsitektur Yolov11 (Hidayatullah et al. 2025)

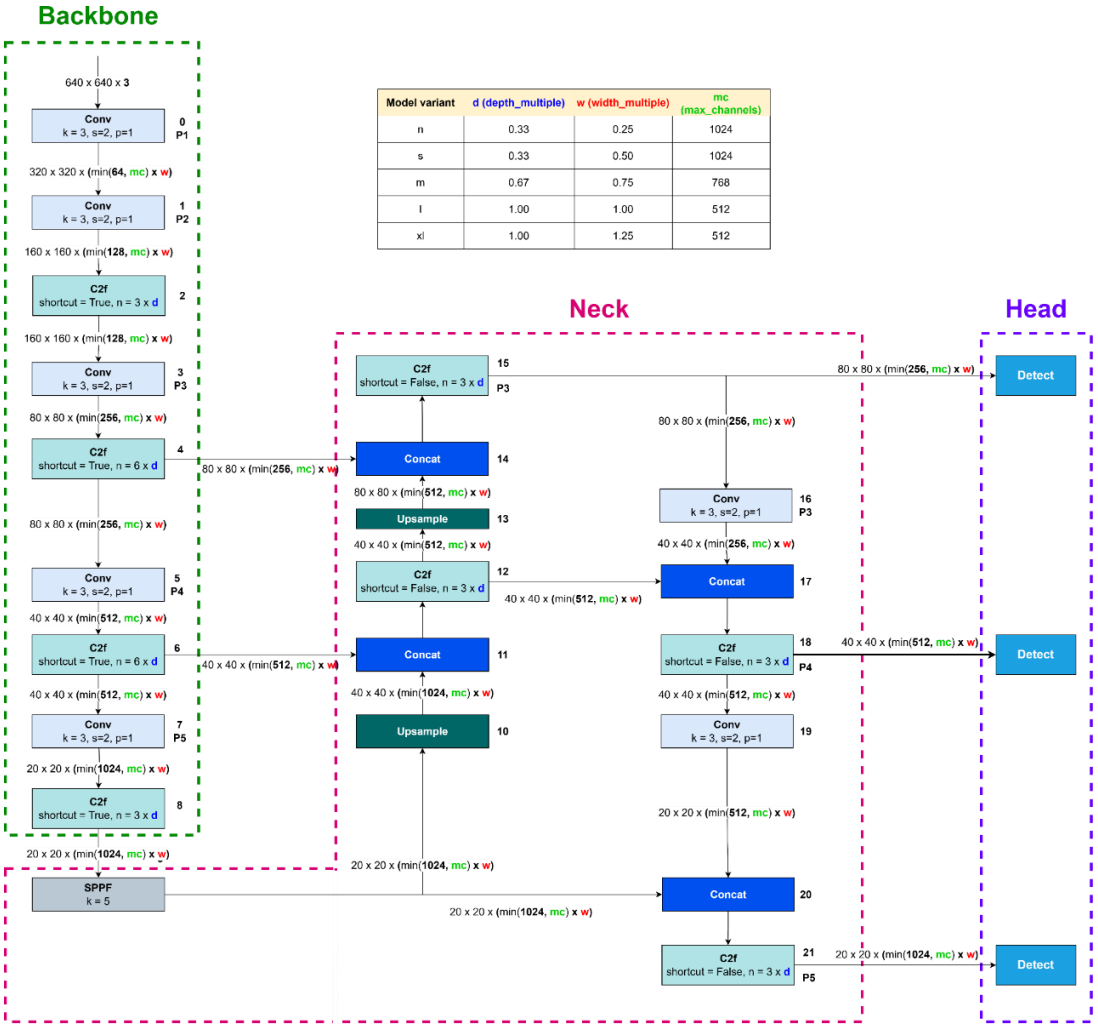
Pada YOLOv11, *backbone* tetap menggunakan lapisan konvolusi awal untuk melakukan *downsampling* gambar, namun diperkuat dengan blok C3k2 sebagai pengganti C2f yang digunakan pada YOLOv8. Blok ini merupakan implementasi yang lebih efisien dari *Cross Stage Partial (CSP) Bottleneck*, dengan dua konvolusi kecil menggantikan satu konvolusi besar sehingga proses ekstraksi fitur berjalan lebih cepat tanpa mengurangi performa. Selain itu, YOLOv11 tetap mempertahankan *Spatial Pyramid Pooling-Fast* (SPPF) untuk menghasilkan representasi multiskala, namun menambahkan blok baru yaitu *Cross Stage Partial with Spatial Attention* (C2PSA). Mekanisme ini memperkuat *spatial attention* sehingga model dapat lebih

fokus pada area penting dalam gambar dan meningkatkan akurasi deteksi untuk objek dengan ukuran bervariasi maupun yang sebagian tertutup (Khanam and Hussain 2024).

Pada bagian *neck*, YOLOv11 menggabungkan fitur dari berbagai skala melalui proses *upsampling* dan *concatenation*. Perubahan signifikan juga terlihat pada penggantian blok C2f dengan C3k2, yang terbukti lebih efisien dalam agregasi fitur. Selain itu, integrasi modul C2PSA semakin meningkatkan fokus model terhadap wilayah penting dalam gambar, menjadikannya lebih unggul dibandingkan YOLOv8 yang tidak memiliki mekanisme perhatian spasial tersebut (Khanam and Hussain 2024).

Sementara itu, *head* YOLOv11 berfungsi menghasilkan prediksi akhir berupa koordinat *bounding box*, skor *objectness*, serta klasifikasi objek. Pada bagian ini, digunakan beberapa blok C3k2 untuk memperhalus pemrosesan fitur multiskala. Blok ini bersifat fleksibel dengan kemampuan menyesuaikan struktur sesuai parameter yang digunakan, serta lebih efisien secara komputasi. Selain itu, ditambahkan pula *Convolution-BatchNorm-SiLU* (CBS) yang berperan menstabilkan aliran data dan meningkatkan performa model melalui fungsi aktivasi SiLU. Pada tahap akhir, lapisan konvolusi digunakan untuk mereduksi fitur sesuai kebutuhan output, kemudian lapisan *Detect* mengonsolidasikan prediksi berupa lokasi objek, probabilitas keberadaan objek, serta label kelas (Khanam and Hussain 2024). Sama seperti arsitektur YOLO lainnya, model ini juga dapat diskalakan menjadi berbagai varian (seperti n, s, m, l, x) dengan menyesuaikan parameter *depth_multiple* dan *width_multiple*. Hal ini memberikan fleksibilitas bagi pengguna untuk memilih model dengan keseimbangan yang tepat antara kecepatan inferensi dan akurasi deteksi sesuai dengan kebutuhan aplikasi mereka.

2.3.2 Yolov8



Gambar 2. 3 Arsitektur Yolov8 (Hidayatullah et al. 2025)

Arsitektur YOLOv8 secara struktural terbagi menjadi tiga komponen utama: *Backbone, Neck,* dan *Head*. Bagian pertama, *Backbone* pada YOLOv8 terdiri dari beberapa lapisan konvolusi yang tersusun secara berurutan untuk mengekstraksi fitur-fitur relevan dari input gambar. Modul baru C2f mengintegrasikan fitur tingkat tinggi dengan informasi kontekstual untuk meningkatkan akurasi deteksi. Blok SPPF (*Spatial Pyramid Pooling Fast*) yang berada pada bagian neck, tepat setelah blok terakhir pada backbone, berfungsi untuk memberikan representasi multi-skala dari feature map (Sohan, Sai Ram, and Rami Reddy 2024). Modul SPPF memungkinkan model untuk menangkap fitur pada berbagai tingkat abstraksi melalui proses pooling pada beberapa skala yang berbeda (Hidayatullah et al. 2025).

Setelah fitur diekstraksi, komponen *Neck* terdapat beberapa blok *concat* dan *upsample*. Proses *upsampling* berfungsi untuk meningkatkan resolusi *feature map*, dan YOLOv8 menggunakan teknik *nearest neighbor upsampling* (Hidayatullah et al. 2025).

Tahap terakhir adalah *Head*, bagian head mengambil *feature maps* yang dihasilkan oleh *backbone* dan memprosesnya lebih lanjut untuk menghasilkan keluaran akhir model dalam bentuk bounding box dan kelas objek. Head ini menghasilkan prediksi pada tiga skala berbeda yang diterima dari *Neck*, memungkinkannya untuk secara spesifik menargetkan objek berukuran kecil, sedang, dan besar (Hidayatullah et al. 2025). Sama seperti arsitektur YOLO lainnya, model ini juga dapat diskalakan menjadi berbagai varian (seperti n, s, m, l, x) dengan menyesuaikan parameter *depth_multiple* dan *width_multiple*. Hal ini memberikan fleksibilitas bagi pengguna untuk memilih model dengan keseimbangan yang tepat antara kecepatan inferensi dan akurasi deteksi sesuai dengan kebutuhan aplikasi mereka.

2.4 Semi-Supervised Learning

Semi-Supervised Learning (SSL) merupakan pendekatan dalam machine learning yang memanfaatkan kombinasi data berlabel dan data tak berlabel dalam proses pelatihan model. SSL menjadi solusi yang efektif ketika data berlabel sulit atau mahal untuk diperoleh, sementara data tak berlabel tersedia dalam jumlah besar (Smith 2025)(Yang et al. 2023). Dalam banyak kasus, model yang dilatih sepenuhnya menggunakan data berlabel memerlukan jumlah anotasi yang sangat besar dan memakan waktu. Oleh karena itu, SSL dirancang untuk memaksimalkan pemanfaatan informasi dari data tak berlabel guna meningkatkan generalisasi model dan efisiensi pelatihan (Mey and Loog 2023). Pendekatan ini sangat relevan untuk diterapkan dalam konteks dunia nyata seperti sistem pengenalan gambar, teks, dan suara, di mana kebutuhan terhadap akurasi tinggi bertemu dengan keterbatasan sumber daya anotasi. Penerapan SSL telah terbukti mampu meningkatkan performa model pada berbagai tugas klasifikasi dan deteksi, bahkan ketika hanya sebagian kecil dari keseluruhan data yang dilengkapi label. Dengan dukungan perkembangan arsitektur model dan strategi pelatihan modern, SSL telah berkembang menjadi pendekatan penting dalam pengembangan sistem kecerdasan buatan yang hemat data (Yang et al. 2023).

2.5 Pseudo-Labeling

Pseudo-labeling merupakan salah satu pendekatan paling umum dalam metode SSL, di mana data tak berlabel diberikan label prediksi oleh model awal *(teacher model)*, dan label tersebut digunakan seolah-olah label asli dalam proses pelatihan ulang *(student model)* (H. Li et al. 2022). Teknik ini memungkinkan pemanfaatan data tak berlabel secara langsung untuk

memperkaya data pelatihan, dengan asumsi bahwa model cukup andal menghasilkan prediksi awal yang akurat (Sohn, Zhang, et al. 2020).

Konsep dasar dari *pseudo-labeling* adalah menyaring prediksi berdasarkan tingkat kepercayaan tertentu, lalu hanya memasukkan data dengan probabilitas tertentu ke dalam pelatihan lanjutan. Pendekatan ini terbukti efektif meningkatkan performa model, khususnya ketika data berlabel sangat terbatas (H. Li et al. 2022). Proses ini juga dapat diulang secara iteratif dalam kerangka *teacher- student*, sehingga model mengalami penyempurnaan bertahap. Dalam konteks deteksi objek, *pseudo-labeling* lebih kompleks dibandingkan klasifikasi gambar karena memerlukan prediksi lokasi *bounding box* selain label kelas. Oleh karena itu, kualitas *pseudo-label* sangat mempengaruhi hasil akhir (Sohn, Zhang, et al. 2020) .

2.6 Metrik Evaluasi

2.6.1 Precision

Precision merupakan metrik yang digunakan untuk mengukur sejauh mana prediksi yang dihasilkan oleh model benar-benar sesuai dengan kondisi aktual. Metrik ini menjadi sangat penting terutama dalam konteks di mana kesalahan dalam prediksi *(False Positive)* harus ditekan seminimal mungkin. Nilai *precision* yang tinggi menunjukkan bahwa sebagian besar prediksi yang dibuat oleh model adalah akurat. Perhitungan precision dapat dilihat pada Persamaan (2.1).

$$Precision = \frac{TP}{TP + FP} \tag{2.1}$$

Keterangan:

- TP *(True Positive)*: Jumlah prediksi yang benar.
- FP *(False Positive)*: Jumlah prediksi yang salah.

2.6.2 Recall

Recall merupakan metrik evaluasi yang digunakan untuk mengukur sejauh mana model mampu mendeteksi semua objek yang sebenarnya ada dalam data. Metrik ini penting dalam situasi di mana kegagalan untuk mendeteksi suatu objek *(False Negative)* perlu diminimalkan, seperti dalam deteksi penyakit atau sistem pengawasan. Nilai *recall* yang tinggi menunjukkan bahwa model memiliki kemampuan deteksi yang baik terhadap objek yang relevan. Persamaan untuk recall ditunjukkan pada Persamaan (2.2).

$$Recall = \frac{TP}{TP + FN} \tag{2.2}$$

Keterangan:

- TP *(True Positive)*: Jumlah prediksi yang benar.
- FN *(False Negative)*: Jumlah prediksi negatif yang salah.

2.6.3 Mean Average Precision (mAP@0.5)

Mean Average Precision (mAP) pada *threshold Intersection over Union* (IoU) 0.5 merupakan salah satu metrik standar yang digunakan untuk mengevaluasi performa deteksi objek. mAP@0.5 menghitung nilai rata-rata dari presisi untuk setiap kelas objek, dengan suatu prediksi dianggap benar apabila nilai IoU antara prediksi dan *ground truth* melebihi 0.5. Metrik ini memperhitungkan baik kemampuan model dalam mendeteksi lokasi objek maupun klasifikasinya. Perhitungan mAP@0.5 dirumuskan dalam Persamaan (2.3).

$$\text{mAP@0.5} = \frac{1}{\text{N}}\sum_{\text{i=1}}^{\text{N}} \text{AP}_{\text{i}}^{(\text{IoU=0.5})} \tag{2.3}$$

Keterangan:

- AP: Nilai *Average Precision* per kelas.
- N: Jumlah kelas.

2.6.4 Mean Average Precision (mAP@0.5:0.95)

Mean Average Precision pada rentang IoU 0.5 hingga 0.95 merupakan metrik evaluasi yang lebih ketat dan menyeluruh dibandingkan mAP@0.5. mAP@0.5:0.95 menghitung rata-rata nilai *Average Precision* (AP) pada 10 threshold IoU yang berbeda, mulai dari 0.5 hingga 0.95 dengan interval 0.05. Metrik ini dirancang untuk mengevaluasi seberapa konsisten dan presisi model dalam melakukan deteksi objek pada berbagai tingkat tumpang tindih antara prediksi dan *ground truth.* Rumus perhitungannya ditunjukkan pada Persamaan (2.4).

$$\text{mAP@0.5:0.95} = \frac{1}{10\text{N}}\sum_{\text{t}\in\text{T}}\sum_{\text{i=1}}^{\text{N}} \text{AP}_{\text{i}}^{(\text{IoU=t})} \tag{2.4}$$

Keterangan:

- AP: Nilai *Average Precision* per kelas.
- N: Jumlah Kelas
- T: Himpunan *threshold* IoU

2.6.5 Inference time

Inference time adalah waktu yang dibutuhkan oleh model untuk menghasilkan prediksi dari satu gambar input, biasanya diukur dalam satuan milidetik (ms) atau detik (s). Metrik ini sangat penting terutama dalam aplikasi *real-time*, seperti sistem deteksi pada kamera pengawas,

kendaraan otonom, atau kasir otomatis. Model dengan *inference time* yang lebih rendah dianggap lebih efisien karena dapat memberikan hasil deteksi lebih cepat. Nilai ini biasanya dipengaruhi oleh kompleksitas arsitektur model serta spesifikasi perangkat keras yang digunakan saat inferensi.

$$Inference\ Time = \frac{Total\ Inference\ Time}{Number\ of\ Images} \tag{2.5}$$

Keterangan:

- *Total Inference Time*: Waktu total yang dibutuhkan model untuk memproses seluruh gambar test uji.
- *Number of Images*: Jumlah gambar uji.

2.6.6 Intersection over Union (IoU)

Intersection over Union (IoU) adalah metrik yang digunakan untuk mengukur tingkat kesesuaian antara *bounding box* hasil prediksi model dengan *bounding box* sebenarnya *(ground truth)*. IoU dihitung sebagai rasio antara luas area irisan (*intersection*) dan luas area gabungan (*union*) dari dua *bounding box*. Nilai IoU berkisar dari 0 hingga 1, di mana nilai yang lebih tinggi menunjukkan tingkat tumpang tindih yang lebih baik. Metrik ini sangat penting untuk menentukan apakah suatu prediksi deteksi objek dianggap benar atau tidak. Perhitungan IoU ditunjukkan dalam Persamaan (2.5).

$$IoU = \frac{Area_{overlap}}{Area_{union}} \tag{2.6}$$

Keterangan:

- IoU: Nilai rasio antara area tumpang tindih dan area gabungan dari dua *bounding box.*
- *Area of Overlap*: Luas area yang merupakan irisan antara prediksi dan *ground truth*.
- *Area of Union*: Total luas gabungan dari dua *bounding box*.

2.6.7 F1-Score

F1-Score adalah metrik evaluasi yang menggabungkan *precision* dan *recall* ke dalam satu nilai tunggal untuk memberikan gambaran menyeluruh mengenai keseimbangan antara keduanya. *F1-Score* dihitung sebagai rata-rata harmonik dari *precision* dan *recall*, sehingga metrik ini sangat berguna dalam situasi di mana diperlukan *trade-off* antara kemampuan model mendeteksi objek dengan benar dan menghindari kesalahan prediksi. Nilai *F1-Score* yang

tinggi menunjukkan bahwa model memiliki kinerja yang seimbang dalam mendeteksi objek secara akurat dan lengkap. Perhitungannya ditunjukkan pada Persamaan (2.7).

$$F1\text{-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{2.7}$$

Keterangan:

- *F1-Score:* Nilai gabungan dari *precision* dan *recall.*
- *Precision:* Proporsi prediksi yang benar.
- *Recall:* Proporsi objek sebenarnya yang berhasil terdeteksi.

2.6.8 True Positive (TP)

True Positive (TP) adalah jumlah prediksi di mana model berhasil mendeteksi objek dan hasil deteksinya sesuai dengan *ground truth*, baik dari segi kelas maupun lokasi. TP menunjukkan bahwa model melakukan deteksi yang benar terhadap objek yang memang ada dalam gambar. Nilai TP yang tinggi menunjukkan bahwa model mampu mengenali objek-objek target secara akurat. TP merupakan komponen dasar dalam perhitungan metrik lain seperti *precision, recall,* dan *F1-score.*

2.6.9 False Positive (FP)

False Positive (FP) terjadi ketika model mendeteksi adanya objek padahal objek tersebut tidak benar-benar ada dalam *ground truth*, atau model mengklasifikasikan objek dengan kelas yang salah. Kesalahan ini menyebabkan penurunan nilai *precision* karena model membuat prediksi yang tidak sesuai dengan kenyataan. Nilai FP yang tinggi menunjukkan bahwa model cenderung menghasilkan deteksi yang keliru atau berlebihan.

2.6.10 False Negative (FN)

False Negative (FN) terjadi ketika model gagal mendeteksi objek yang sebenarnya ada dalam gambar. Ini berarti objek tidak terdeteksi meskipun terdapat dalam *ground truth*, sehingga dapat menurunkan nilai *recall*. Nilai FN yang tinggi menunjukkan bahwa model kurang sensitif dalam mengenali keberadaan objek dan sering melewatkan deteksi yang seharusnya dilakukan.

2.6.11 Box Loss

Box Loss adalah metrik yang digunakan untuk mengukur seberapa besar kesalahan prediksi model dalam menentukan posisi *bounding box* dibandingkan dengan posisi sebenarnya (*ground truth*). Metrik ini biasanya dihitung berdasarkan selisih koordinat pusat, lebar, dan

tinggi antara *bounding box* prediksi dan yang sebenarnya. Semakin kecil nilai *Box Loss*, semakin akurat model dalam memperkirakan lokasi objek.

2.6.12 Class Loss

Class Loss adalah metrik yang digunakan untuk menghitung kesalahan model dalam mengklasifikasikan kelas objek yang terdeteksi. Metrik ini menunjukkan seberapa besar perbedaan antara prediksi kelas oleh model dan label kelas yang sebenarnya. *Class Loss* biasanya dihitung menggunakan fungsi *cross-entropy loss*, yang banyak digunakan pada tugas klasifikasi.

2.6.13 Distribution Focal Loss (DFL Loss)

Distribution Focal Loss (DFL) adalah fungsi kerugian yang dirancang untuk meningkatkan kualitas prediksi regresi *bounding box* dalam model deteksi objek. Berbeda dengan metode regresi tradisional yang memprediksi satu nilai, DFL mengubah tugas regresi menjadi prediksi distribusi terdiskretisasi, memungkinkan model untuk belajar representasi spasial yang lebih kaya. DFL bekerja dengan cara memperlakukan koordinat box sebagai distribusi probabilitas, lalu menerapkan *loss* yang berfokus pada distribusi nilai target yang lebih relevan, sehingga membantu meningkatkan ketelitian posisi prediksi.

2.7 Studi Literatur

Tabel 2. 1 Studi Literatur

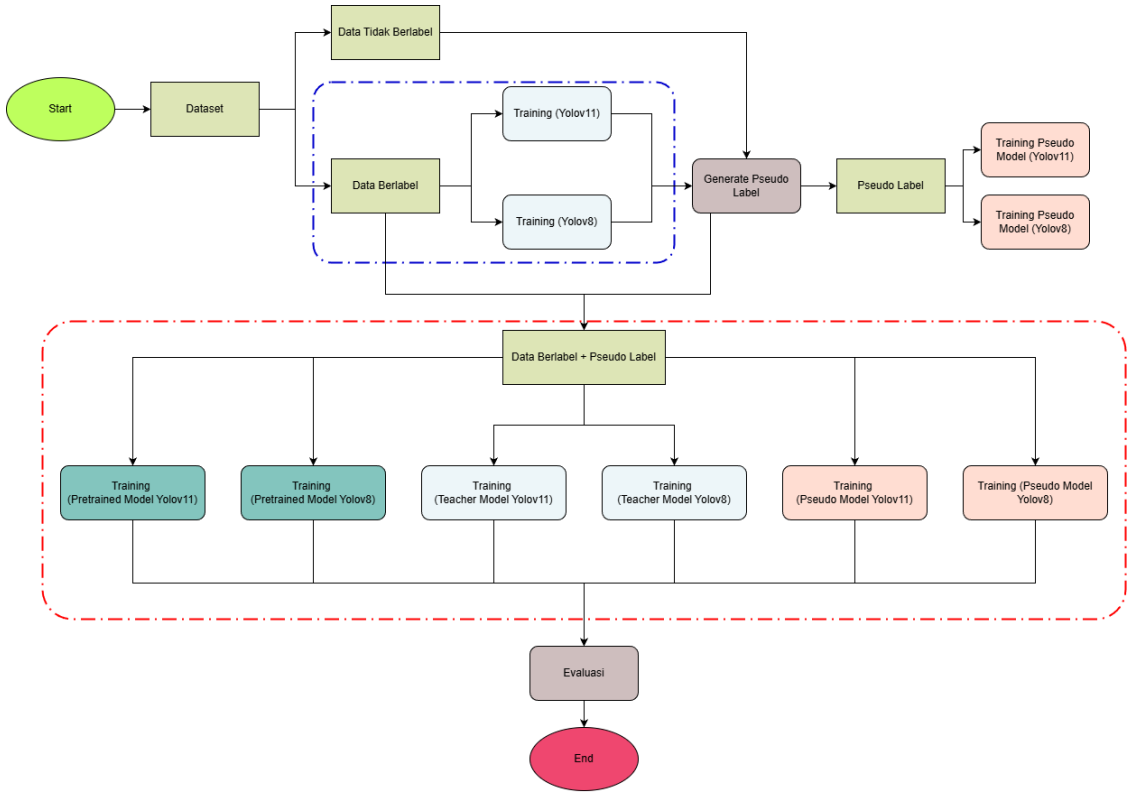
Peneliti	Model	Dataset	Metode	Hasil
R. Yilmazer	RetinaNet, Yolov3, Yolov4	3.153 gambar, 5 kelas: Beverage, Breakfast, Food, Empty Shelf, Almost Empty Shelf	Semi-Supervised Lerning	Yolov4 mendapatkan mAP = 0.9187
D. H. Fudholi, A. Kurniawardhani, G. I. Andaru, A. A. Alhanafi, and N. Najmudin	Yolov5n, Yolov6n, Yolov7-tiny, Yolov8n	7.697 gambar, 125 kelas produk	Supervised Learning	Yolov8 mAP50 = 0.933.

J. Chauhan, S. Varadarajan, and M. M. Srivastava	GFL + Resnet-50	9000 berlabel + 10K - 200K tak berlabel	Semi-Supervised	mAP = 0.403 dengan 200k pseudo-label
D. Jha, A. Mahjoubfar, and A. Joshi,	EfficientDet (D0 - D1), YOLOv5 (n6)	1.000 gambar, 1 kelas (Empty Shelf)	Supervised Learning	Hasil Yolov5n6 mAP = 0.669
R. Digo Saputra and D. Hatta Fudholi	YOLOv4-tiny	4.637 gambar, 106 kelas (produk susu bubuk)	Supervised Learning	mAP = 92.14

Berdasarkan Tabel 2. 1, terlihat bahwa mayoritas penelitian menggunakan keluarga model YOLO karena keunggulannya dalam kecepatan inferensi dan efisiensi komputasi, sehingga cocok untuk skenario *real-time*. Beberapa studi, seperti yang dilakukan oleh (Yilmazer and Birant 2021) serta (Digo Saputra and Hatta Fudholi 2023), menggunakan arsitektur YOLOv4 (termasuk varian tiny) dan berhasil memperoleh performa yang tinggi (mAP > 0.9). Sementara itu, penelitian oleh (Fudholi et al. 2024) menunjukkan bahwa YOLOv8n mampu memberikan hasil terbaik di antara varian YOLO generasi sebelumnya. Di sisi lain, studi (Chauhan, Varadarajan, and Srivastava 2021). Dan (Yilmazer and Birant 2021) mengeksplorasi pendekatan *semi-supervised learning*, yang memungkinkan pemanfaatan data tak berlabel dalam jumlah besar melalui *pseudo-labeling*. Hasilnya, meskipun peningkatannya tidak signifikan dibandingkan metode *fully-supervised*, pendekatan ini menawarkan solusi praktis untuk mengatasi keterbatasan data berlabel. Penelitian (Jha, Mahjoubfar, and Joshi 2022) juga menegaskan keunggulan YOLO dibandingkan *EfficientDet* dalam konteks dataset kecil, dengan YOLOv5n6 menghasilkan metrik yang lebih baik. Sejauh pengetahuan penulis, belum terdapat studi yang secara khusus mengimplementasikan *semi-supervised learning* berbasis YOLO terbaru (misalnya YOLOv11) pada domain *on-shelf availability* (OSA).

BAB III
METODOLOGI

Penelitian ini dilaksanakan melalui lima tahapan utama untuk menerapkan pendekatan *semi-supervised learning* berbasis *pseudo-labeling* pada arsitektur YOLOv11 dan YOLOv8 dalam tugas deteksi objek produk retail. Tahapan tersebut mencakup proses pengumpulan data, pelatihan *supervised*, *pseudo-labeling,* pelatihan *semi-supervised*, serta evaluasi performa model dan kualitas *pseudo-label*. Diagram alur tahapan penelitian disajikan pada Gambar 3. 1 Tahapan Penelitian.



Gambar 3. 1 Tahapan Penelitian

Kotak Biru (*Teacher Model Training*), Kotak Merah (*Student Model Training*)

3.1 Pengumpulan dan Persiapan Data

Tahap awal dari penelitian ini adalah pengumpulan dan persiapan *dataset* yang akan digunakan dalam pelatihan dan pengujian model deteksi objek. Proses pengumpulan data dalam penelitian ini dilakukan dengan mengambil gambar produk secara langsung dari toko. Gambar yang diambil berfokus pada yang tersusun di rak, seperti yang ada pada Gambar 3. 2.



Gambar 3. 2 Contoh dataset

Tabel 3. 1 Karakteristik Dataset

Aspek	Spesifikasi
Format File Gambar	JPEG, JPG, PNG
Resolusi	1024x768 (*landscape*), 768x1024 (*potrait*)
Ukuran File	100 KB - 300 KB
Jumlah Objek per Gambar	Bervariasi
Format File Anotasi	YOLO (.txt)
Kategori Produk	Susu Bubuk

Karakteristik teknis dataset yang digunakan dirangkum pada Tabel 3. 1, yang mencakup format file, resolusi gambar, ukuran file, jumlah objek per gambar, format anotasi, serta kategori produk. Dalam prosesnya, data gambar yang dikumpulkan bervariasi dalam beberapa aspek untuk meningkatkan ketahanan model. Varisasi tersebut mencakup tingkat kemiringan, jarak kamera dari objek, dan tingkat pencahayaan sesuai dengan kondisi ruangan. Meskipun berbagai variasi diterapkan, posisi tegak lurus dan jarak yang tidak terlalu jauh tetap diprioritaskan untuk memastikan detail produk dapat terlihat agar dapat dikenali secara optimal. Seluruh gambar yang telah dikumpulkan kemudian diberi anotasi atau label untuk membedakan setiap objek yang terdapat dalam gambar. Daftar kelas yang digunakan dalam penelitian ini disajikan pada Tabel 3. 2. Proses pelabelan dilakukan menggunakan perangkat bantu *LabelImg* (Tzutalin 2015), dan hasil anotasi disimpan dalam format YOLO (.txt) dengan

struktur seperti yang ditunjukkan pada Gambar 3. 3, setiap baris merepresentasikan satu objek pada gambar dengan format kelas seperti berikut: *class, x_center, y_center, width, height*.

```
92 0.130859 0.189941 0.217448 0.235352
103 0.176432 0.441406 0.194010 0.181641
102 0.509766 0.456055 0.149740 0.144531
93 0.326172 0.194824 0.196615 0.239258
94 0.672526 0.443848 0.173177 0.176758
94 0.845052 0.444824 0.187500 0.180664
94 0.777995 0.638184 0.160156 0.147461
105 0.481120 0.640137 0.149740 0.143555
105 0.628255 0.639648 0.141927 0.140625
106 0.183594 0.828125 0.151042 0.128906
99 0.466146 0.845703 0.138021 0.109375
99 0.602865 0.835449 0.130208 0.116211
97 0.729818 0.832031 0.134115 0.109375
10 0.729167 0.194336 0.208333 0.244141
```

Gambar 3. 3 Label.txt

Tabel 3. 2 Daftar Kelas

ID	Class Name
0	LACTOGROW 3 Happynutri Van 12x750g A
1	LACTOGROW 3 Happynutri 12x750g A
2	LACTOGROW 4 Happynutri Honey12x750g A
3	LACTOGROW 4 Happynutri Van 12x750g A
4	LACTOGROW 3 Happynutri Honey 24x350g A
5	LACTOGEN 1 Happynutri 40x180g A
6	LACTOGEN 2 Happynutri 40x180g A
7	DANCOW 5+ Van Advn ExcNutr 24x400g B
8	NESTLE BATITA 1+ Madu+Iron 12x850g
9	LACTOGEN 1 Happynutri 12x750g A
10	DANCOW 1+ Madu Advn ExcNutr 12x800g B
11	LACTOGEN 2 Happynutri 12x750g A
12	LACTOGROW 3 Happynutri Honey 12x750g A
13	DANCOW 3+ Madu Advn ExcNutr 12x800g B
14	DANCOW 5+ Madu Advn ExcNutr 12x800g B

15	LACTOGROW 3 Happynutri 24x350g A
16	LACTOGEN 2 Happynutri 24x350g A
17	LACTOGEN 1 Happynutri 24x350g A
18	LACTOGROW 3 Happynutri Van 24x350g A
19	LACTOGROW 3 Happynutri Honey 40x180g A
20	SGM
21	MORINAGA
22	VIDORAN
23	FRISIAN FLAG
24	ENFAGROW
25	DANCOW 5+ Van Advn ExcNutr 12x800g B
26	DANCOW 3+ Cok Advn ExcNutr 12x800g B
27	DANCOW 3+ Van Advn ExcNutr 12x800g B
28	DANCOW 3+ Van Advn ExcNutr 24x400g B
29	DANCOW 5+ Madu Advn ExcNutr 24x400g B
30	DANCOW 5+ Cok Advn ExcNutr 12x800g B
31	DANCOW 1+ Madu Advn ExcNutr 12x1000g B
32	DANCOW 1+ Van Advn ExcNutr 12x1000g B
33	DANCOW 3+ Madu Advn ExcNutr 12x1000g B
34	DANCOW 3+ Van Advn ExcNutr 12x1000g B
35	DANCOW 5+ Madu Advn ExcNutr 12x1000g B
36	DANCOW 1+ Van Advn ExcNutr 12x800g N1 B
37	DANCOW 1+ Van Advn ExcNutr 24x400g B
38	DANCOW 3+ Madu Advn ExcNutr 24x400g B
39	DANCOW 1+ Madu Advn ExcNutr 24x400g B
40	DANCOW 1+ Cok Advn ExcNutr 12x800g B
41	DANCOW 3+ Madu Advn ExcNutr 40x200g B
42	S-26 PROCAL Pouch 12x700g A
43	LACTOGEN 1 Happynutri 12x1kg A
44	LACTOGEN 2 Happynutri 12x1kg A
45	LACTOGROW 3 Happynutri Honey 12x1kg A

46	LACTOGROW 3 Happynutri Van 12x1kg A
47	LACTOGROW 4 Happynutri Honey12x1kg A
48	LACTOGROW 4 Happynutri Van 12x1kg A
49	LACTOGROW 4 Happynutri Van 12x1kg B
50	LACTOGROW 4 Happynutri Honey12x1kg A
51	LACTOGROW 3 Happynutri Van 12x1kg A
52	DANCOW 1+ Madu Advn ExcNutr 40x200g A
53	DANCOW 1+ Van Advn ExcNutr 40x200g B
54	LACTOGEN 1 Happynutri 12x750g B
55	LACTOGEN 2 Happynutri 12x750g B
56	LACTOGROW 3 Happynutri Honey 12x750g B
57	LACTOGROW 4 Happynutri Honey12x750g B
58	S-26 ULTIMA PROCAL Can Top 6x850g
59	S-26 ULTIMA PROMIL 2 Can Top 6x850g
60	S-26 ULTIMA PROMIL 1 Can Top 6x850g
61	S-26 GOLD PROMIL 1 Can Top 6x900g
62	S-26 PROCAL GOLD Can Top 6x900g
63	S-26 PROMISE GOLD Can Top 6x900g
64	S-26 PROCAL GOLD Can 6x1.6kg
65	S-26 PROMISE GOLD Can 6x1.6kg
66	S-26 PROCAL GOLD Can 24x400g
67	S-26 PROMIL GOLD 2 Can 24x400g
68	S-26 PROMIL GOLD 1 Can 24x400g
69	S-26 PROMISE Pouch 12x400g A
70	S-26 PROCAL Honey Pouch 12x400g
71	S-26 PROMIL 1 Pouch 12x700g
72	S-26 PROMIL 2 Pouch 12x700g
73	S-26 PROMISE Pouch 12x700g
74	S-26 PROCAL Pouch 12x400g
75	S-26 PROCAL Pouch 6(2x700g)
76	S-26 GOLD PROMIL 2 Can Top 6x900g

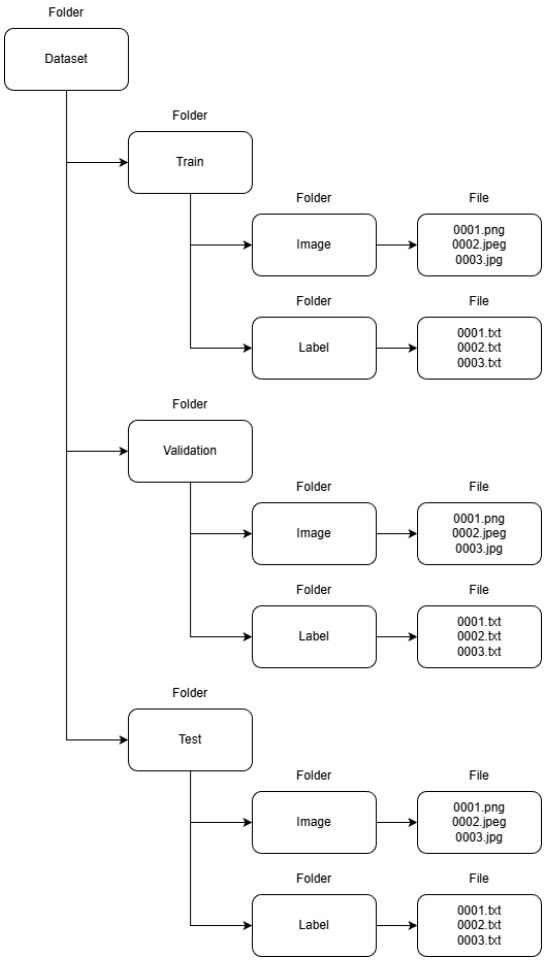
77	S-26 PROMISE Pouch 6(2x700g)
78	S-26 PROMIL 2 Pouch 12x400g
79	S-26 PROMIL 1 Pouch 12x400g
80	LACTOGROW 3 Happynutri Van 12x750g B
81	LACTOGROW 3 Happynutri 12x750g B
82	DANCOW BABY 1 Box 40x120g
83	DANCOW BABY 2 Box 40x120g
84	LACTOGEN 1 Happynutri 24x350g B
85	LACTOGEN 2 Happynutri 24x350g B
86	LACTOGROW 3 Happynutri Honey 24x350g B
87	LACTOGEN 1 Happynutri 40x180g B
88	DANCOW 1+ Van Advn ExcNutr 24x400g C
89	DANCOW 3+ Van Advn ExcNutr 24x400g C
90	DANCOW 5+ Madu Advn ExcNutr 24x400g C
91	DANCOW BABY 2 Box 12x550g
92	DANCOW 5+ Madu Advn ExcNutr 12x800g C
93	DANCOW 3+ Madu Advn ExcNutr 12x800g C
94	DANCOW 1+ Van Advn ExcNutr 12x800g N1 C
95	DANCOW BABY 1 Box 12x550g
96	DANCOW 1+ Van Advn ExcNutr 12x1000g C
97	DANCOW 1+ Madu Advn ExcNutr 12x1000g C
98	DANCOW 3+ Van Advn ExcNutr 12x1000g C
99	DANCOW 3+ Madu Advn ExcNutr 12x1000g C
100	DANCOW 1+ Van Advn ExcNutr 40x200g C
101	DANCOW 1+ Madu Advn ExcNutr 40x200g C
102	DANCOW 3+ Van Advn ExcNutr 12x800g C
103	DANCOW 5+ Van Advn ExcNutr 12x800g C
104	DANCOW 5+ Madu Advn ExcNutr 12x1000g C
105	DANCOW 3+ Cok Advn ExcNutr 12x800g C
106	DANCOW 5+ Cok Advn ExcNutr 12x800g C
107	LACTOGEN 2 Happynutri 40x180g B

108	DANCOW 3+ Madu Advn ExcNutr 24x400g C
109	LACTOGEN 1 Happynutri 12x1kg B
110	DANCOW 1+ Madu Advn ExcNutr 24x400g C
111	DANCOW 1+ Madu Advn ExcNutr 12x800g C
112	LACTOGROW 3 Happynutri Van 24x350g B
113	LACTOGROW 4 Happynutri Van 12x750g B
114	NESTLE DATITA 3+Vanilla+Iron 12x900g A
115	NESTLE DATITA 3+ Madu+Iron 12x900g A
116	NESTLE BATITA 1+ Madu+Iron 12x900g A
117	NESTLE BATITA 1+ Vanilla+Iron 12x900g A
118	NESTLE DATITA 5+ Madu+Iron 12x900g A
119	LACTOGEN 2 Happynutri 12x1kg B
120	LACTOGROW 3 Happynutri Van 12x1kg B
121	LACTOGROW 3 Happynutri Honey 12x1kg B
122	LACTOGROW 4 Happynutri Van 12x1kg C
123	LACTOGROW 3 Happynutri 24x350g B
124	LACTOGROW 4 Happynutri Honey12x1kg B
125	S-26 PROCAL 3 Pouch 15x200g
126	S-26 PROCAL GOLD Can 24x400g A
127	S-26 PROMIL GOLD 2 Can 24x400g A
128	S-26 PROMIL GOLD 1 Can 24x400g A
129	S-26 GOLD PROMIL 1 Can Top 6x900g B
130	S-26 GOLD PROMIL 2 Can Top 6x900g B
131	S-26 PROCAL GOLD Can 6x1.6kg B
132	BATITA 3+ Madu 12x850g
133	BATITA 3+ Madu 24x350g
134	BATITA 1+ Madu 24x350g
135	BATITA 1+ Madu 12x850g
136	BATITA 1+ Vanilla 24x350g
137	BATITA 1+ Vanilla 12x850g
138	BATITA 3+ Vanilla 12x850g

139	BATITA 5+ Madu 12x850g
140	S-26 PROMISE GOLD Can 6x1.6kg B
141	DANCOW 5+ Van Advn ExcNutr 24x400g C
142	LACTOGROW PRO 1+ Van 12x1kg
143	LACTOGEN PRO 0-6 mo 40x180g
144	LACTOGEN PRO 0-6 mo 24x350g
145	LACTOGEN PRO 0-6 mo 12x735g
146	LACTOGEN PRO 0-6 mo 12x1kg
147	LACTOGROW PRO 3+ Van 12x735g
148	LACTOGROW PRO 3+ Honey12x735g
149	LACTOGROW PRO 1+ Honey 12x1kg
150	LACTOGROW PRO 3+ Van 12x1kg
151	LACTOGROW PRO 3+ Honey 12x1kg
152	LACTOGROW PRO 1+ Honey 24x350g
153	LACTOGROW PRO 1+ Van 24x350g
154	LACTOGROW PRO 1+ 24x350g
155	LACTOGROW PRO 1+ Honey 40x145g
156	LACTOGROW PRO 1+ Honey 12x735g
157	LACTOGROW PRO 1+ 12x735g
158	LACTOGEN PRO 6-12 mo 40x180g
159	LACTOGEN PRO 6-12 mo 24x350g
160	LACTOGEN PRO 6-12 mo 12x735g
161	LACTOGEN PRO 6-12 mo 12x1kg
162	LACTOGROW PRO 1+ Van 12x735g
163	LACTOGROW 3 Happynutri 12x1kg
164	LACTOGROW 4 Happynutri Van 24x350g
165	S-26 PROMISE Pouch 12x400g N4
166	S-26 PROMISE Pouch 12x700g N4
167	NUTRICIA
168	DANCOW 1+ Madu Advn ExcNutr 24x400g A
169	LACTOGROW 3 Happynutri Honey 40x180g B

170	LACTOGROW 4 Happynutri Honey12x735g
171	S-26 PROCAL GOLD Can Top 6x900g B
172	DANCOW 3+ Van Advn ExcNutr 12x800g
173	DANCOW 5+ Cok Advn ExcNutr 12x1000g B

Setelah itu, *dataset* akan dibagi menjadi *training, test, validation* yang nantinya folder dataset akan terlihat seperti pada Gambar 3. 4. Proses ini biasanya disebut juga *data splitting, data train* digunakan untuk melatih model agar dapat mengenali pola dari data, sementara *data test* dan *validation* digunakan untuk mengevaluasai performa model pada data yang belum pernah dilihat sebelumnya. Untuk pendekatan *semi-supervised learning* nantinya akan ada 4 eksperimen dengan proporsi 20%, 40%, 60%, dan 80% data berlabel.



Gambar 3. 4. Struktur Folder Dataset

3.2 Teacher Model Training

Setelah tahap pengumpulan dan persiapan data, langkah berikutnya adalah pelatihan awal model deteksi objek menggunakan data berlabel manual. Model yang digunakan dalam penelitian ini adalah YOLOv11n dan YOLOv8n, dengan pertimbangan bahwa kedua arsitektur tersebut memiliki reputasi kuat dalam menyeimbangkan kecepatan inferensi dan akurasi deteksi objek, sedangkan pemilihan varian nano dilakukan untuk mengevaluasi performa dari versi model paling dasar. Pada tahap ini, pelatihan dilakukan secara terpisah untuk setiap skenario proporsi data berlabel (20%, 40%, 60%, dan 80%) untuk membandingkan performa antar skenario, dan juga dilakukan pelatihan *supervised* sebagai pembanding terhadap pendekatan *semi-supervised*. Proses pelatihan pada dasarnya mengikuti parameter *default* YOLO, termasuk *learning rate schedule, batch size,* dan penggunaan *optimizer AdamW*. Namun terdapat beberapa penyesuaian, yaitu jumlah *epochs* ditingkatkan menjadi 500 agar model memiliki cukup waktu untuk mencapai kondisi stabil, kemudian ditambahkan *patience* = 50 sehingga proses pelatihan akan berhenti otomatis apabila tidak terjadi peningkatan performa selama 50 *epochs* berturut-turut, indikasi bahwa model telah mencapai titik stabil tanpa harus memaksakan sisa *epochs*. Selain itu, *warmup epochs* = 10 diterapkan untuk memastikan peningkatan bobot berjalan stabil pada fase awal pelatihan. Model yang dihasilkan pada tahap ini didefinisikan sebagai *teacher model*, yang selanjutnya digunakan untuk menghasilkan *pseudo-label* pada data tidak berlabel. Sebagai tambahan skenario pengujian, dilakukan pula pelatihan *Pseudo Model*, yaitu proses supervised training menggunakan hanya data *pseudo-label* dengan tujuan mengamati apakah model ini mampu mendekati atau bahkan menggantikan *teacher model* hasil pelatihan dengan data berlabel manual, terutama pada proporsi data berlabel kecil seperti 20%.

3.3 Pseudo-labeling

Proses *pseudo-labeling* dilakukan setelah *supervised training* selesai dengan memanfaatkan model hasil pelatihan sebagai *teacher model* pada skema *semi-supervised learning*. Model *teacher* ini digunakan untuk mendeteksi objek pada data tidak berlabel, menghasilkan file berformat YOLO (.txt) yang memuat koordinat *bounding box* dan kelas objek terdeteksi. Untuk menjaga kualitas *pseudo-label*, hanya prediksi dengan tingkat kepercayaan di atas 0.7 (*confidence threshold*) yang digunakan, prediksi dengan nilai kepercayaan lebih rendah diabaikan. Pemilihan *confidence threshold* 0.7 mengacu pada

penelitian *FixMatch* (Sohn, Berthelot, et al. 2020) yang menjelaskan bahwa semakin rendah nilai *threshold* akan semakin banyak kesalahan *pseudo-label* akibat tingginya *noise* pada prediksi, sementara dua studi pendukung lainnya, yaitu *Unbiased Teacher* (Liu et al. 2021) dan *In Defense of Pseudo-Labeling* (Rizve et al. 2021), juga menggunakan nilai 0.7 sebagai *confidence threshold*. Selain itu, uji coba empiris yang dilakukan dalam penelitian ini menggunakan *threshold* 0.3, 0.5, 0.7, dan 0.9 menunjukkan bahwa nilai 0.7 memberikan hasil terbaik dibanding nilai lainnya. Format keluaran dari *pseudo-labeling* sama persis dengan data berlabel manual sehingga dapat langsung digabungkan dalam proses pelatihan ulang. Tahapan ini menjadi elemen penting dalam kerangka *semi-supervised learning* karena model awal berperan sebagai *teacher* untuk memperkuat kemampuan model akhir (*student*) melalui pemanfaatan data tidak berlabel sebagai sumber informasi tambahan.

3.4 Student Model Training

Setelah *pseudo-label* berhasil dihasilkan pada data tak berlabel, tahap berikutnya adalah pelatihan ulang model dengan menggabungkan data berlabel manual dan *pseudo-label*. Proses ini disebut *semi-supervised training*, yang bertujuan memanfaatkan keseluruhan data latih secara optimal, baik yang dianotasi oleh manusia maupun yang dihasilkan oleh *teacher model*.

Pada penelitian ini, dilakukan 6 eksperimen *student training* dengan variasi sebagai berikut:

a. *Pretrained Model* YOLOv11n : Model dasar YOLOv11n
b. *Pretrained Model* YOLOv8n : Model dasar YOLOv8n
c. *Teacher Model* YOLOv11n : model hasil *supervised training* pada data berlabel digunakan sebagai titik awal pelatihan *student*.
d. *Teacher Model* YOLOv8n : serupa dengan skenario ketiga, tetapi menggunakan arsitektur YOLOv8n.
e. *Pseudo Model* YOLOv11n : model hasil *supervised training* hanya pada data *pseudo-label*.
f. *Pseudo Model* YOLOv8n : serupa dengan skenario kelima, tetapi menggunakan arsitektur YOLOv8n.

Seluruh 6 eksperimen ini dijalankan pada setiap proporsi data berlabel (20%, 40%, 60%, dan 80%), sehingga memungkinkan analisis performa model terhadap variasi jumlah data berlabel dalam kerangka *semi-supervised learning*.

3.5 Evaluation

Tahap terakhir dalam penelitian ini adalah melakukan evaluasi terhadap performa *student model* hasil pelatihan *semi-supervised*. Evaluasi bertujuan untuk mengukur kemampuan model dalam mendeteksi objek secara akurat pada data uji, serta menilai sejauh mana kontribusi *pseudo-label* dapat meningkatkan performa deteksi. Proses evaluasi dilakukan menggunakan 92 gambar uji yang telah dianotasi secara manual, sehingga hasil pengujian dapat dibandingkan secara objektif dengan *ground truth*. Evaluasi yang dilakukan meliputi:

a. Model hasil *supervised training*
b. Model hasil *semi-supervised training*

Selain itu, dilakukan juga evaluasi terhadap kualitas pseudo-label yang dihasillan oleh *teacher model*. Evaluasi mencakup:

a. Evaluasi visual, yaitu membandingkan hasil *pseudo-label* dengan *ground truth* secara langsung pada gambar, untuk melihat kesesuaian *bounding box* serta kelas.
b. Evaluasi metrik, yaitu menghitung skor metrik dengan menjadikan *ground-truth* sebagai acuan, sehingga dapat diketahui sejauh mana *pseudo-label* mendekati kualitas label manual.

Evaluasi juga dilakukan terhadap grafik training untuk menganalisis pola pembelajaran model selama proses pelatihan. Grafik training yang dievaluasi meliputi kurva loss (training dan validation loss) serta metrik performa seperti mAP selama epoch berlangsung. Analisis grafik ini bertujuan untuk mendeteksi adanya overfitting, underfitting, atau konvergensi model, serta membandingkan stabilitas dan efisiensi pembelajaran antara supervised dan semi-supervised training.

Metrik evaluasi yang digunakan mencakup *Precision, Recall, mAP@0.5, mAP@0.5:0.95, F1-Score, IoU,* dan *Inference time.* Dengan menggunakan kombinasi evaluasi performa model dan kualitas *pseudo-label*, penelitian ini diharapkan dapat memberikan gambaran menyeluruh mengenai efektivitas pendekatan *semi-supervised* dibandingkan *supervised training*, sekaligus menilai kontribusi nyata *pseudo-label* terhadap peningkatan performa deteksi objek.

BAB IV
HASIL DAN PEMBAHASAN

4.1 Implementasi dan Hasil Pengumpulan dan Persiapan Data

Tahap awal penelitian adalah pengumpulan dan persiapan *dataset*. Dari proses ini, berhasil dikumpulkan total 918 gambar. Semua gambar tersebut kemudian dianotasi dan dikelompokkan ke dalam 174 kelas produk yang berbeda dengan total jumlah objek 11.161. Lalu, datasetnya dibagi menjadi 734 gambar untuk *train set*, 92 gambar untuk *test*, dan 92 gambar *validation.* Untuk mendukukung pendekatan *semi-supervised learing*, data *training set* dibagi lagi ke dalam data berlabel dan tidak berlabel, dengan empat skenario. Skrip pada Gambar 4. 1 digunakan untuk membagi dataset secara merata dan memastikan setiap kelas memiliki minimal jumlah objek di dalam set data berlabel.

```
# ==== PARAMETER ====
label_dir = "/kaggle/input/dataset1/dataset/test/labels"
image_dir = "/kaggle/input/dataset1/dataset/test/images"
min_obj_per_class = 1     # minimal objek per kelas
labeled_ratio = 0.2      # proporsi dataset berlabel (misal 20%)

output_paths = {
    "labeled_img": "/kaggle/working/dataset/labeled/images",
    "labeled_lbl": "/kaggle/working/dataset/labeled/labels",
    "unlabeled_gt_img": "/kaggle/working/dataset/unlabeled_gt/images",
    "unlabeled_gt_lbl": "/kaggle/working/dataset/unlabeled_gt/labels",
    "unlabeled_img": "/kaggle/working/dataset/unlabeled/images"
}
for path in output_paths.values():
    os.makedirs(path, exist_ok=True)

def find_image_file(base_name):
    for ext in ['.jpg', '.jpeg', '.png', '.JPG', '.JPEG', '.PNG']:
        img_path = os.path.join(image_dir, base_name + ext)
        if os.path.isfile(img_path):
            return img_path
    return None

all_files_by_class = defaultdict(list)
all_base_names = []

for lbl_file in os.listdir(label_dir):
    if not lbl_file.endswith('.txt'):
        continue

    base_name = os.path.splitext(lbl_file)[0]
    all_base_names.append(base_name)
    lbl_path = os.path.join(label_dir, lbl_file)

    with open(lbl_path, 'r') as f:
        lines = f.readlines()
```

```
    for line in lines:
        if line.strip() == "":
            continue
        class_id = line.strip().split()[0]
        all_files_by_class[class_id].append(base_name)

for class_id in all_files_by_class:
    all_files_by_class[class_id] = list(set(all_files_by_class[class_id]))

labeled_files_initial = set()
for class_id, files_list in all_files_by_class.items():
    random.shuffle(files_list)

    count_per_class = 0
    for file_base_name in files_list:
        if count_per_class >= min_obj_per_class:
            break

        lbl_path = os.path.join(label_dir, file_base_name + '.txt')
        with open(lbl_path, 'r') as f:
            lines = f.readlines()

        temp_class_counts = defaultdict(int)
        for line in lines:
            if line.strip() == "":
                continue
            cid = line.strip().split()[0]
            temp_class_counts[cid] += 1

        if class_id in temp_class_counts:
            labeled_files_initial.add(file_base_name)
            count_per_class += temp_class_counts[class_id]

total_images = len(all_base_names)
target_labeled_count = int(total_images * labeled_ratio)
unlabeled_files = list(set(all_base_names) - labeled_files_initial)

random.shuffle(unlabeled_files)
needed_from_unlabeled = target_labeled_count - len(labeled_files_initial)

if needed_from_unlabeled > 0:
    labeled_files_final       =       list(labeled_files_initial)       +
unlabeled_files[:needed_from_unlabeled]
else:
    labeled_files_final = list(labeled_files_initial)
    random.shuffle(labeled_files_final)
    labeled_files_final = labeled_files_final[:target_labeled_count]

unlabeled_files_final = list(set(all_base_names) - set(labeled_files_final))
labeled_files_final = set(labeled_files_final)
unlabeled_files_final = set(unlabeled_files_final)

labeled_class_counts = defaultdict(int)

# Salin labeled files
for file_base_name in labeled_files_final:
    src_lbl = os.path.join(label_dir, file_base_name + '.txt')
    src_img = find_image_file(file_base_name)
```

```
    if src_img:
        shutil.copy2(src_lbl,     os.path.join(output_paths["labeled_lbl"],
os.path.basename(src_lbl)))
        shutil.copy2(src_img,     os.path.join(output_paths["labeled_img"],
os.path.basename(src_img)))

        with open(src_lbl, 'r') as f:
            lines = f.readlines()
        for line in lines:
            if line.strip() != "":
                class_id = line.strip().split()[0]
                labeled_class_counts[class_id] += 1
    else:
        print(f"[WARNING] Gambar tidak ditemukan: {file_base_name}")

# Salin unlabeled_gt files
for file_base_name in unlabeled_files_final:
    src_lbl = os.path.join(label_dir, file_base_name + '.txt')
    src_img = find_image_file(file_base_name)

    if src_img:
        shutil.copy2(src_lbl,
os.path.join(output_paths["unlabeled_gt_lbl"], os.path.basename(src_lbl)))
        shutil.copy2(src_img,
os.path.join(output_paths["unlabeled_gt_img"], os.path.basename(src_img)))
    else:
        print(f"[WARNING] Gambar tidak ditemukan: {file_base_name}")

# ==== STEP 5: Salin gambar unlabeled_gt ke unlabeled (tanpa label) ====
for img_file in os.listdir(output_paths["unlabeled_gt_img"]):
    src_img = os.path.join(output_paths["unlabeled_gt_img"], img_file)
    dst_img = os.path.join(output_paths["unlabeled_img"], img_file)
    shutil.copy2(src_img, dst_img)
```

Gambar 4. 1 Kode Program Split Dataset

Kode di atas berfungsi untuk melakukan pembagian dataset menjadi subset berlabel dan tidak berlabel sesuai kebutuhan eksperimen semi-supervised learning. Prosesnya terdiri dari beberapa tahapan: pertama, program membaca semua file label (.txt) dan mengelompokkan gambar berdasarkan kelas objeknya, sehingga diketahui distribusi data tiap kelas. Selanjutnya, pada Step 2, program memastikan bahwa setiap kelas memiliki minimal jumlah objek tertentu (dalam kode ini min_obj_per_class=7) dengan cara memilih file yang mengandung objek dari kelas tersebut. Setelah kebutuhan minimum tercapai, pada Step 3 dilakukan penyesuaian agar proporsi dataset berlabel sesuai dengan target (misalnya 20% dari total dataset melalui parameter labeled_ratio=0.2). Jika jumlah gambar berlabel masih kurang, sistem akan menambahkan file acak dari sisanya; jika kelebihan, maka file akan diacak ulang lalu dipangkas.

Hasil akhirnya berupa dua kelompok utama: labeled dataset yang digunakan untuk melatih teacher model, serta unlabeled dataset yang akan diberi pseudo-label. Sebagai tambahan,

salinan dataset unlabeled_gt (gambar dengan label aslinya) juga dipertahankan untuk keperluan evaluasi. Pada Step 4, file-file yang masuk subset berlabel disalin ke direktori khusus sekaligus dihitung distribusi objek per kelas agar dapat dipastikan seimbang. Sedangkan file yang termasuk subset tidak berlabel disalin ke folder unlabeled_gt (lengkap dengan label) dan unlabeled (gambar saja tanpa label).

Dengan mekanisme ini, peneliti dapat mengontrol komposisi dataset berlabel dan tidak berlabel, memastikan adanya representasi minimum tiap kelas, sekaligus mempersiapkan data untuk pseudo-labeling dan semi-supervised training. Setelah pembagian data selesai, tiap eksperimen akan memiliki komposisi data berlabel dan tidak berlabel seperti pada Tabel 4. 1.

Tabel 4. 1 Proporsi Dataset

Proposorsi Data Berlabel	Jumlah Gambar Berlabel	Jumlah Gambar Tidak Berlabel
20%	147	587
40%	294	440
60%	440	294
80%	587	147

4.2 Implementasi dan Hasil Teacher Model

Tahap kedua dalam penelitian ini adalah implementasi pelatihan model deteksi objek menggunakan data berlabel dengan pendekatan *supervised learning.* Model yang dihasilkan pada tahap ini berperan sebagai *teacher model*, yang nantinya digunakan untuk menghasilkan *pseudo-label* pada data tidak berlabel dalam kerangka *semi-supervised learning*. Selain itu, pada skenario ini juga dilakukan pelatihan model dengan 100% data berlabel (full supervised) sebagai *baseline* pembanding terhadap model *semi-supervised*. Dengan demikian, performa dari pendekatan semi-supervised dapat dievaluasi secara lebih obyektif terhadap hasil yang mungkin dicapai apabila seluruh data berlabel tersedia.

```
def train_model():
    model = YOLO("/kaggle/input/modelyolo/other/default/1/yolo11n.pt")
    model.train(
        data='/kaggle/input/config_yaml/other/default/1/dataconfig1.yaml',
# Dataset configuration
        epochs=500,  # Total training epochs
        imgsz=640,  # Image size
        batch=16,  # Batch size
        device='0,1',  # Use dual GPU
        optimizer='AdamW',  # Optimizer
        val=True,  # Enable validation
        patience=50, #Early Stopping
        warmup_epochs=10,
    )
```

```
if __name__ == "__main__":
    train_model()
```

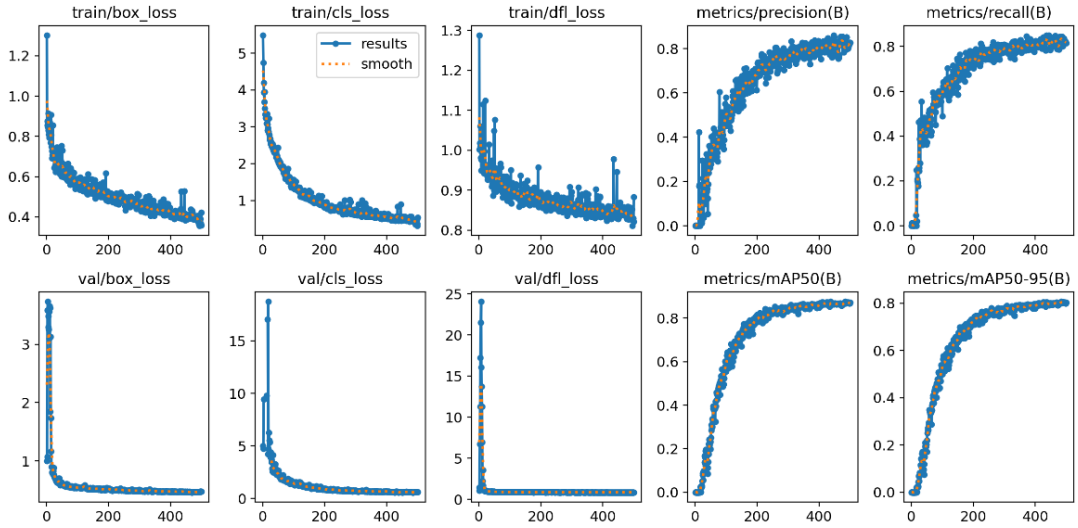
Gambar 4. 2 Kode Program Training Model

Berdasarkan potongan kode pada Gambar 4. 2 pelatihan model menggunakan arsitektur YOLOv11 dan YOLOv8 dengan versi nano karena memiliki ukuran yang ringan, sehingga efisien dalam hal waktu pelatihan dan penggunaan memori, tetapi tetap cukup akurat untuk mendeteksi objek. Lalu model dilatih menggunakan konfigurasi dataset yang telah ditentukan pada file dataconfig1.yaml, dengan parameter pelatihan epochs 500, ukuran gambar 640 piksel, batch size 16, patience= 50, dan warmup_epochs=10. Perangkat komputasi yang digunakan device='0.1' karena menggunakan dual GPU untuk mempercepat proses pelatihan. Opimizer yang dipilih adalah AdamW.

Tabel 4. 2 Hasil Train Fully Supervised

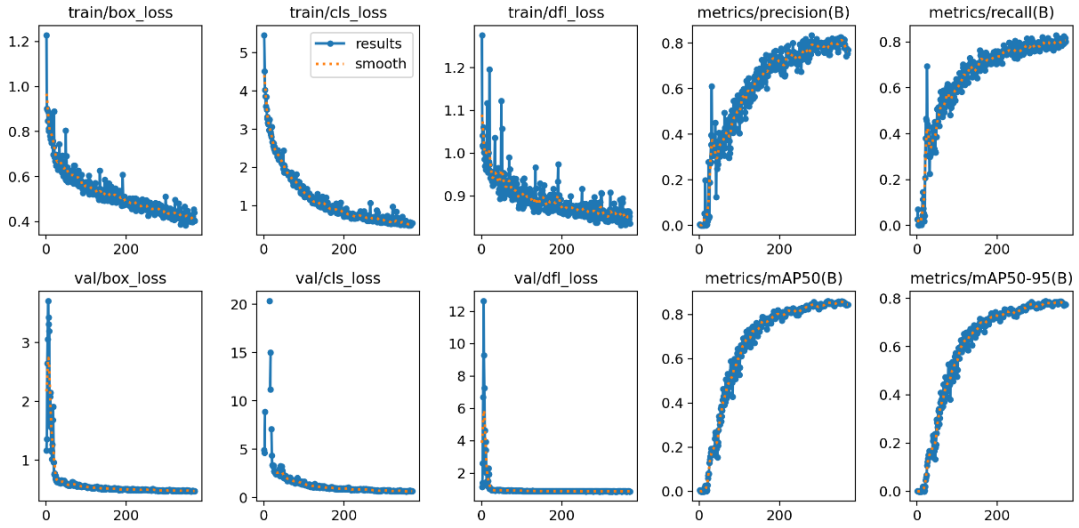
Model	Epocs	Box Loss	Cls Loss	DFL Loss	Precision	Recall	mAP50	mAP50-95	Inference Time	Waktu Training (Jam)
Yolov11	219	0.427	0.418	0.850	0.897	0.911	0.96	0.892	4.9ms	0.615
Yolov8	300	0.358	0.262	0.823	0.878	0.943	0.953	0.892	4.8ms	0.770

Hasil pelatihan model dengan pendekatan fully supervised menggunakan seluruh data berlabel menunjukkan performa yang sangat baik pada kedua arsitektur. YOLOv8 mencapai precision sebesar 0.878, recall 0.943, mAP50 0.953, dan mAP50-95 0.892 setelah 300 epoch pelatihan. Sementara itu, YOLOv11 menunjukkan precision 0.897, recall 0.911, mAP50 0.96, dan mAP50-95 0.892 setelah 219 epoch. Meskipun YOLOv11 memiliki precision yang lebih tinggi, YOLOv8 unggul pada recall dan mAP50, menandakan kemampuannya dalam mendeteksi objek secara lebih menyeluruh. Dari sisi efisiensi, inference time YOLOv11 sedikit lebih cepat (4.9 ms) dibanding YOLOv8 (4.8 ms), dan waktu pelatihannya juga lebih singkat (0.615 jam vs. 0.770 jam). Secara keseluruhan, kedua model menunjukkan performa yang sangat kompetitif dengan nilai mAP50-95 yang identik (0.892), menjadikannya acuan (baseline) untuk evaluasi model semi-supervised.



Gambar 4. 3 Teacher Model 20% Yolov11

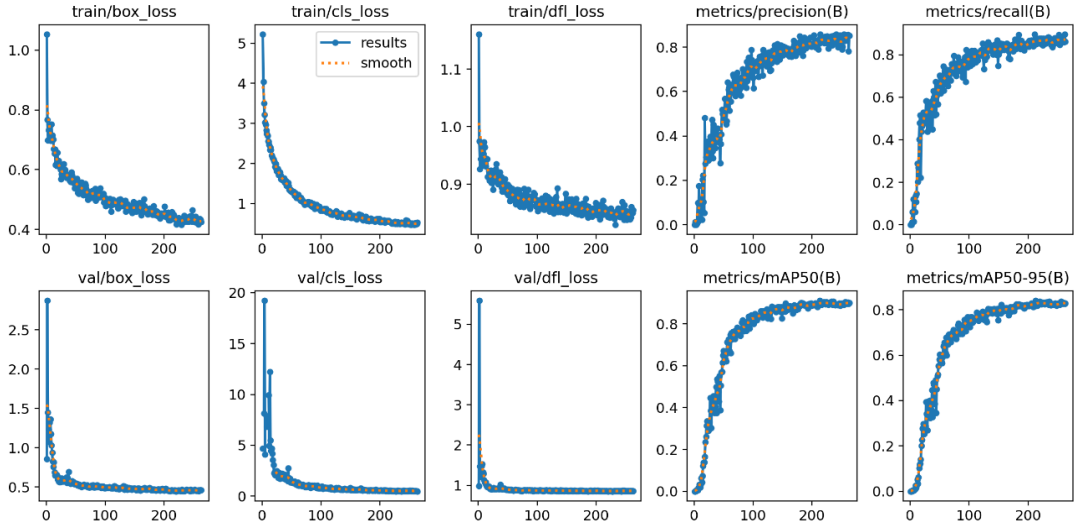
Pada eksperimen Teacher Model 20% dengan arsitektur YOLOv11, grafik training menunjukkan pola pembelajaran yang sangat stabil, ditandai dengan penurunan nilai loss yang berlangsung drastis pada fase awal pelatihan dan kemudian mulai melambat setelah melewati sekitar epoch ke-200, menandakan bahwa model mulai mencapai titik konvergensi. Tren ini selaras dengan peningkatan metrik evaluasi, di mana precision, recall, mAP50, dan mAP50-95 meningkat tajam pada awal proses training dan kemudian mengalami perlambatan peningkatan setelah epoch ke-200 ketika performa model telah mendekati stabil.



Gambar 4. 4 Teacher model 20% Yolov8

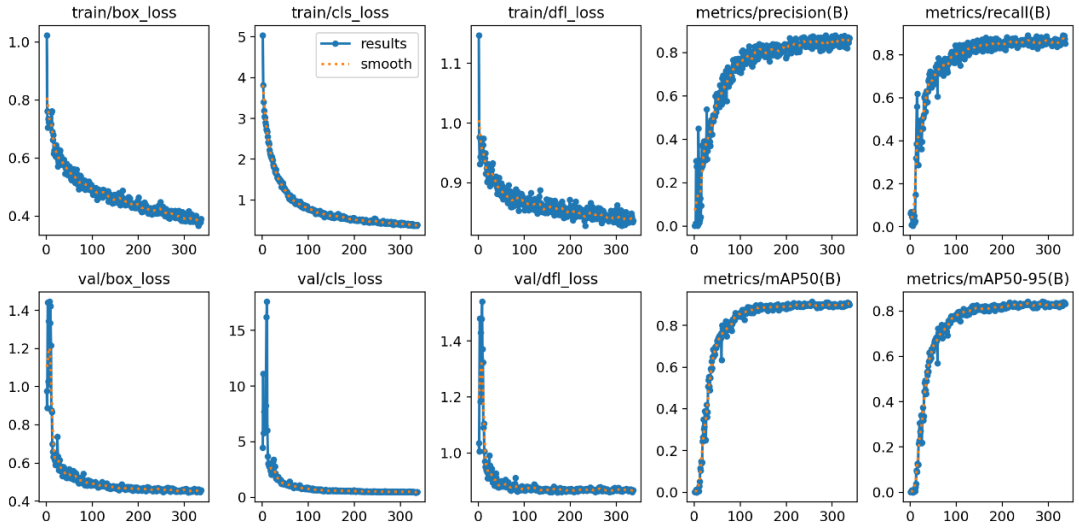
Pada eksperimen Teacher Model 20% dengan YOLOv8, grafik training memperlihatkan penurunan nilai loss yang cukup signifikan pada fase awal pelatihan hingga mendekati stabil setelah melewati sekitar epoch ke-200, menandakan proses konvergensi model yang baik.

Namun, berbeda dengan YOLOv11, grafik DFL Loss menunjukkan sedikit fluktuasi sepanjang pelatihan, meskipun masih berada dalam rentang yang wajar dan tidak mengganggu tren penurunan secara keseluruhan. Pada sisi metrik evaluasi, precision dan recall juga mengalami fluktuasi di awal pelatihan akibat proses penyesuaian model terhadap data, tetapi kemudian bergerak stabil setelah beberapa epoch awal terlewati. Secara keseluruhan, pola grafik menunjukkan bahwa walaupun terdapat dinamika pada DFL Loss dan metrik evaluasi di fase awal, model berhasil mencapai kestabilan performa setelah epoch ke-200 dan mampu beradaptasi dengan karakteristik data secara efektif.



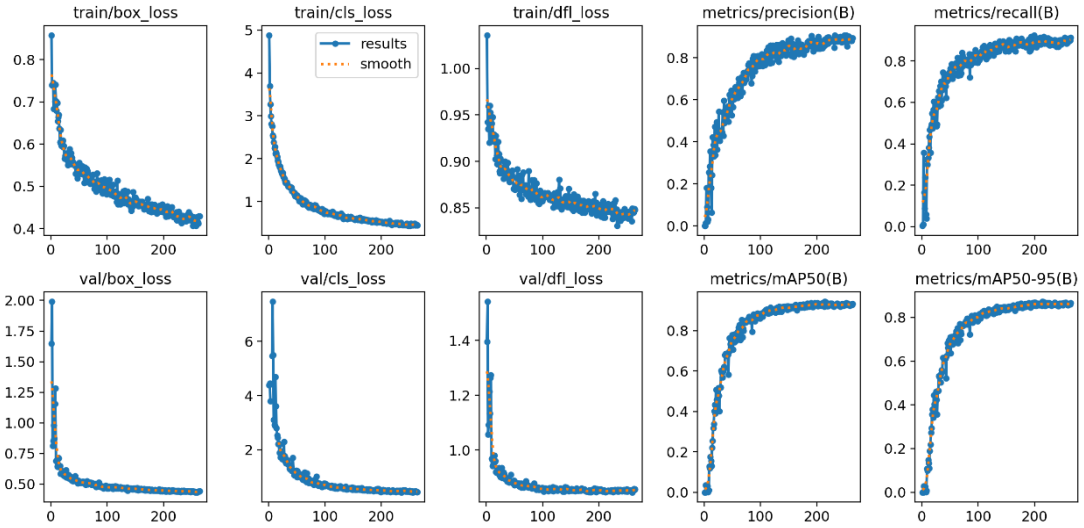
Gambar 4. 5 Teacher Model 40% Yolov11

Pada eksperimen Teacher Model 40% dengan YOLOv11, grafik training menunjukkan tren penurunan loss yang tajam pada fase awal pelatihan sebelum akhirnya mulai melambat dan mendekati titik konvergensi pada sekitar epoch ke-80 hingga 90, menandakan bahwa model mulai mencapai kestabilan lebih cepat dibanding eksperimen dengan proporsi data yang lebih rendah. Sejalan dengan hal tersebut, grafik metrik evaluasi seperti precision, recall, serta mAP50 dan mAP50-95 juga mengalami peningkatan pesat di awal pelatihan sebelum perubahan kenaikan semakin kecil setelah masuk ke epoch 80 - 90. Pola ini mengindikasikan bahwa model mampu belajar secara efisien dari data dalam jumlah yang lebih besar, sehingga adaptasi terhadap data berlangsung lebih cepat dan stabil tanpa menunjukkan fluktuasi berarti pada fase akhir pelatihan.



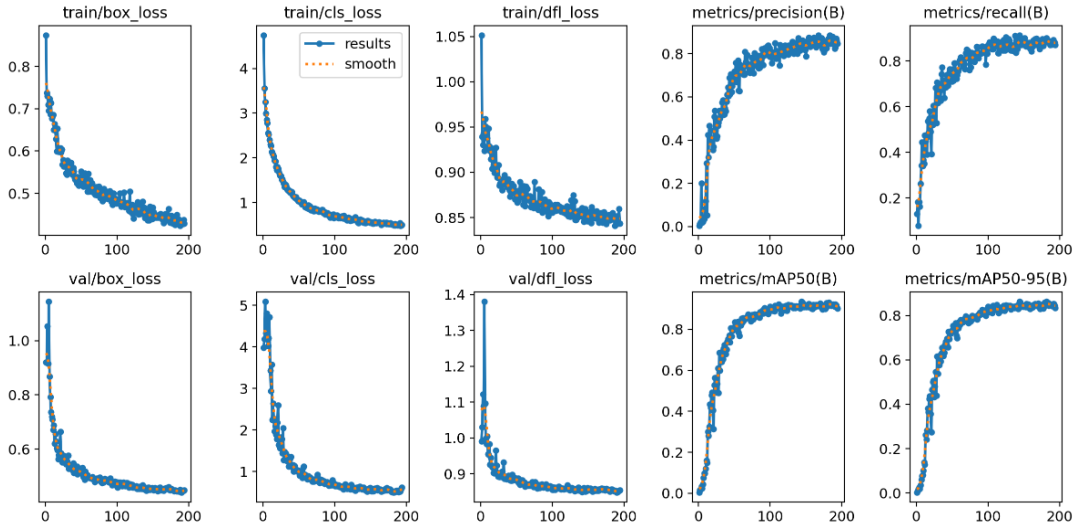
Gambar 4. 6 Teacher Model 40% Yolov8

Pada eksperimen Teacher Model 40% dengan YOLOv8, grafik training memperlihatkan penurunan loss yang cepat pada fase awal pelatihan sebelum akhirnya perlambatan mulai terjadi mendekati titik konvergensi setelah sekitar epoch ke-100, menunjukkan bahwa proses pembelajaran mulai stabil pada fase tersebut. Sejalan dengan pola loss, metrik precision, recall, serta mAP50 dan mAP50-95 juga meningkat secara signifikan pada awal pelatihan dan kemudian menunjukkan kenaikan yang lebih lambat setelah epoch ke-100, menandakan bahwa performa model sudah mulai mencapai plateau. Tidak tampak fluktuasi yang mengganggu pada fase akhir pelatihan, sehingga dapat disimpulkan bahwa model mampu memanfaatkan data dengan baik pada proporsi 40%, menghasilkan pembelajaran yang efisien dan stabil hingga akhir pelatihan.



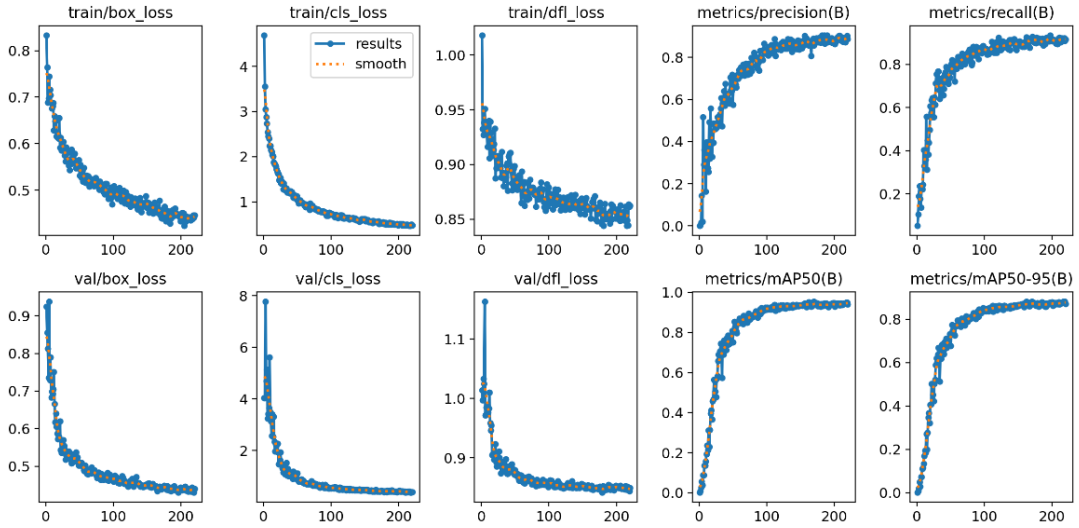
Gambar 4. 7 Teacher Model 60% Yolov11

Pada eksperimen Teacher Model 60% dengan YOLOv11, grafik training menunjukkan penurunan loss yang cukup tajam pada tahap awal pelatihan dan mulai memasuki fase perlambatan sekitar epoch ke-70 hingga ke-80, yang menandakan model telah mulai mencapai titik stabil dalam proses pembelajaran. Selaras dengan pola loss tersebut, metrik precision, recall, serta mAP50 dan mAP50-95 juga meningkat pesat pada fase awal sebelum kenaikannya melambat setelah memasuki epoch ke-70, di mana grafik metrik bergerak lebih landai dan mendekati plateau. Pola ini menggambarkan bahwa model sudah mampu mengekstraksi informasi dengan baik dari data pada proporsi 60% sehingga proses peningkatan performa berangsur stabil menjelang akhir pelatihan.



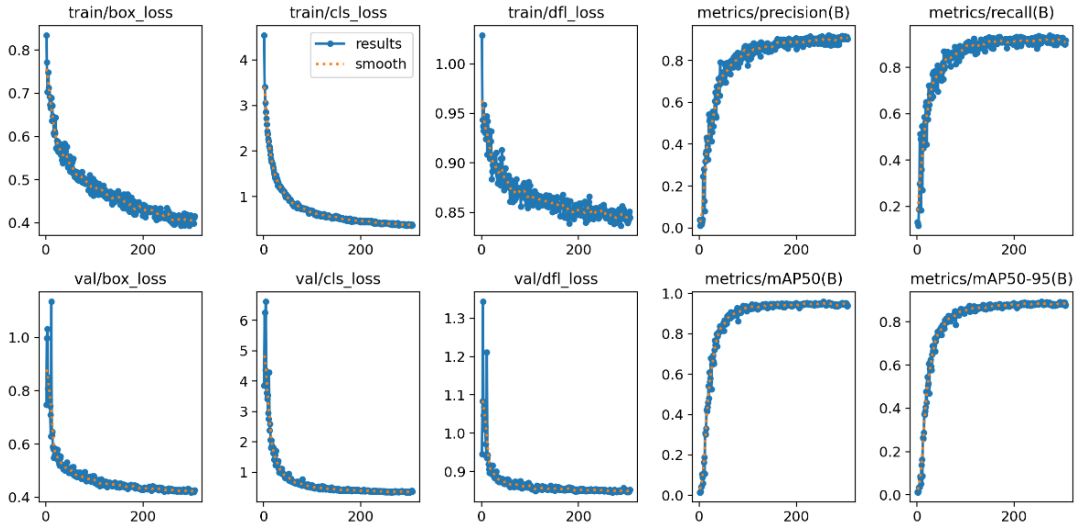
Gambar 4. 8 Teacher Model 60% Yolov8

Pada eksperimen Teacher Model 60% dengan YOLOv8, grafik training menunjukkan penurunan loss yang cukup agresif pada awal pelatihan, kemudian mulai melambat dan bergerak stabil di sekitar epoch ke-60 hingga ke-70, menandakan bahwa laju pembelajaran model mulai mencapai titik jenuh. Pola ini juga tercermin pada grafik precision, recall, serta mAP50 dan mAP50-95 yang meningkat tajam di fase awal sebelum perbaikannya bertahap dan cenderung landai setelah memasuki epoch ke-60. Tren ini mengindikasikan bahwa model telah berhasil memperoleh representasi fitur penting dari data pada proporsi 60%, sehingga proses optimisasi berikutnya hanya menghasilkan peningkatan kecil secara bertahap hingga akhir pelatihan.



Gambar 4. 9 Teacher Model 80% Yolov11

Pada eksperimen Teacher Model 80% dengan YOLOv11, grafik training memperlihatkan tren penurunan loss yang sangat cepat pada fase awal pelatihan sebelum kemudian melambat dan bergerak stabil pada kisaran epoch 70 hingga 80, menandakan bahwa proses pembelajaran memasuki tahap konvergensi. Meskipun secara umum loss turun secara konsisten, pada metrik DFL loss terlihat sedikit fluktuasi, namun pola penurunannya tetap tergolong mulus tanpa gangguan yang signifikan. Selaras dengan loss, metrik precision, recall, mAP50, dan mAP50-95 juga mengalami peningkatan pesat pada fase awal sebelum akhirnya meningkat secara bertahap setelah epoch ke-80. Pola tersebut menunjukkan bahwa model telah mampu mengekstraksi representasi fitur dengan baik dari data berproporsi 80%, sehingga perbaikan performa selanjutnya cenderung bersifat marginal mendekati batas kemampuan generalisasi model.



Gambar 4. 10 Teacher Model 80% Yolov8

Pada eksperimen Teacher Model 80% menggunakan YOLOv8, grafik training menunjukkan pola yang sangat mulus dengan tren penurunan loss yang stabil dan konsisten sejak awal pelatihan. Kurva pembelajaran mulai melambat pada sekitar epoch ke-100, menandakan bahwa model telah mencapai titik konvergensi dan pembaruan parameter selanjutnya hanya memberikan peningkatan yang kecil. Tidak terlihat fluktuasi berarti pada box loss, cls loss, maupun DFL loss, sehingga proses pelatihan berjalan optimal tanpa gejala overfitting ataupun ketidakstabilan gradien. Sejalan dengan hal tersebut, metrik precision, recall, mAP50, dan mAP50-95 juga meningkat secara progresif hingga mendekati epoch ke-100 sebelum akhirnya bergerak lebih landai. Pola ini menegaskan bahwa Teacher Model YOLOv8 berhasil memanfaatkan proporsi 80% data secara efektif untuk pembelajaran, menghasilkan kestabilan training yang baik dan performa mendekati puncak sejak sebelum pelatihan berakhir.

Tabel 4. 3 Hasil Training Teacher Model

Model	Epocs	Box Loss	Cls Loss	DFL Loss	Precision	Recall	mAP50	mAP50-95	Inference Time	Waktu Training (Jam)
Eksperimen 20%										
Yolov8	370	0.404	0.541	0.836	0.792	0.802	0.862	0.791	4.7ms	0.315
Yolov11	500	0.421	0.534	0.883	0.796	0.841	0.876	0.807	4.8ms	0.453
Eksperimen 40%										
Yolov8	337	0.389	0.388	0.836	0.869	0.854	0.91	0.841	5.0ms	0.407
Yolov11	263	0.425	0.527	0.855	0.839	0.877	0.911	0.841	4.4ms	0.357
Eksperimen 60%										
Yolov8	194	0.428	0.494	0.843	0.859	0.907	0.935	0.863	4.8ms	0.335
Yolov11	264	0.429	0.450	0.847	0.888	0.889	0.944	0.874	4.5ms	0.498
Eksperimen 80%										
Yolov8	306	0.415	0.367	0.844	0.901	0.921	0.96	0.893	4.7ms	0.655
Yolov11	220	0.446	0.478	0.862	0.858	0.916	0.954	0.882	4.6ms	0.518

Tabel tersebut menyajikan performa model YOLOv8 dan YOLOv11 pada empat skenario eksperimen dengan proporsi data berlabel sebesar 20%, 40%, 60%, dan 80%. Setiap model dievaluasi berdasarkan metrik Box Loss, Cls Loss, DFL Loss, Precision, Recall, mAP50, mAP50-95, waktu inferensi, serta durasi training. Pada proporsi 20% data berlabel, YOLOv11 menunjukkan performa yang sedikit lebih baik dibanding YOLOv8 pada metrik utama. YOLOv11 memperoleh mAP50 sebesar 0.876 dan mAP50-95 sebesar 0.807, lebih tinggi daripada YOLOv8 dengan mAP50 sebesar 0.862 dan mAP50-95 sebesar 0.791. Precision dan recall keduanya juga sedikit unggul pada YOLOv11. Namun, waktu pelatihan YOLOv11 lebih lama dibanding YOLOv8. Pada 40% data berlabel, terjadi peningkatan performa pada kedua model. YOLOv11 dan YOLOv8 sama-sama mencapai mAP50 sebesar 0.91, menunjukkan peningkatan dari eksperimen sebelumnya. Secara mAP50-95, kedua model memiliki hasil yang sama, yaitu 0.841. Precision tertinggi diperoleh YOLOv8 (0.869), sedangkan recall tertinggi dimiliki YOLOv11 (0.877). Dalam hal efisiensi, YOLOv11 memiliki waktu inferensi lebih cepat dan durasi training lebih singkat dibanding YOLOv8.

Dengan peningkatan proporsi data menjadi 60%, performa model meningkat secara konsisten. YOLOv11 mendapatkan mAP50 sebesar 0.944 dan mAP50-95 sebesar 0.874, sedikit lebih tinggi dibanding YOLOv8 dengan mAP50 sebesar 0.935 dan mAP50-95 sebesar 0.863. Precision lebih tinggi pada YOLOv11 (0.888), sementara recall tertinggi berada pada YOLOv8 (0.907). Pada eksperimen ini, waktu inferensi kedua model relatif mirip. Pada proporsi 80%, nilai performa model mencapai hasil terbaik pada keseluruhan eksperimen. YOLOv8 memperoleh mAP50 sebesar 0.960 dan mAP50-95 sebesar 0.893, sedikit lebih tinggi

dibanding YOLOv11 dengan mAP50 sebesar 0.954 dan mAP50-95 sebesar 0.882. Precision dan recall YOLOv8 juga berada pada nilai tertinggi, yaitu masing-masing 0.901 dan 0.921. Namun, YOLOv8 membutuhkan durasi training paling lama, sementara YOLOv11 tetap unggul pada efisiensi waktu.

4.3 Implementasi dan Hasil Pseudo-Labeling

Tahap ini mengimplementasikan proses pseudo-labeling pada data tidak berlabel dengan memanfaatkan teacher model yang telah dilatih pada tahap sebelumnya.

```
def generate_pseudo_labels(model_path, image_dir, output_dir, conf_threshold):
    model = YOLO(model_path)
    os.makedirs(output_dir, exist_ok=True)

    images = [f for f in os.listdir(image_dir) if f.endswith(('.jpg', '.jpeg', '.png', '.JPG', '.JPEG', '.PNG'))]
    total = len(images)
    labeled = 0
    skipped = 0

    for image in images:
        label_path = os.path.join(output_dir, image.rsplit('.', 1)[0] + ".txt")

        if os.path.exists(label_path):
            skipped += 1
            continue

        model.predict(
            source=os.path.join(image_dir, image),
            conf=conf_threshold,
            save=False,
            save_txt=True,
            save_conf=False,
            project=output_dir,
            name="",
            exist_ok=True
        )
        labeled += 1

if __name__ == "__main__":
    generate_pseudo_labels(
        model_path="/kaggle/working/runs/detect/train/weights/best.pt",
        image_dir="/kaggle/working/dataset/unlabeled/images",
        output_dir="/kaggle/working/dataset/unlabeled/labels",
        conf_threshold=0.7
    )
```

Gambar 4. 11 Kode Program Generate Pseudo-label

Berdasarkan Gambar 4.2 Kode tersebut digunakan untuk menghasilkan pseudo-labels dari kumpulan data gambar yang tidak berlabel dengan memanfaatkan model YOLO hasil training

teacher model. Fungsi generate_pseudo_labels menerima beberapa argumen penting, yaitu path model terlatih (model_path), direktori gambar yang belum berlabel (image_dir), direktori penyimpanan hasil pseudo-label (output_dir), serta nilai ambang batas kepercayaan prediksi (conf_threshold). Pertama, model YOLO dimuat dari bobot terbaik hasil training supervised, kemudian program memindai seluruh file gambar dalam direktori input dan menyiapkan folder output untuk menyimpan hasil label. Proses pseudo-labeling dilakukan dengan memanggil fungsi model.predict() pada setiap gambar, di mana bounding box hasil deteksi dengan confidence di atas ambang batas akan disimpan dalam format .txt sesuai standar YOLO. Mekanisme pemeriksaan juga ditambahkan: apabila file label untuk gambar tertentu sudah ada, gambar tersebut dilewati sehingga tidak diproses ulang. Setelah seluruh gambar diproses, sistem memberikan ringkasan jumlah total gambar, berapa yang berhasil dilabeli, dan berapa yang dilewati. Dengan alur ini, kode mampu secara otomatis memperbesar jumlah dataset berlabel melalui pseudo-label yang dihasilkan teacher model, yang kemudian dapat digunakan untuk tahap semi-supervised training pada student model.

4.3.1 Kualitas Pseudo-Label



(a)



(b)



(c)



(d)





(e) (f)

Gambar 4. 12 Perbandingan Visual Pseudo-label 20%
YOLOv11 (a, c, e) dan YOLOv8 (b, d, f),
label asli (hijau), pseudo-label (merah)

Pada eksperimen dengan proporsi 20% data berlabel, dilakukan perbandingan antara pseudo-label yang dihasilkan dengan ground truth pada beberapa gambar. Hasil perbandingan ditunjukkan pada Gambar a-f, di mana (a), (c), dan (e) merupakan hasil pseudo-labeling menggunakan YOLOv11, sedangkan (b), (d), dan (f) menggunakan YOLOv8.

Pada Gambar (a) dan (b) terlihat bahwa model mampu memberikan label pada seluruh objek yang sesuai dengan ground truth. Hal ini menunjukkan bahwa model dapat mengenali objek dengan baik pada kondisi visual yang relatif jelas dan teratur. Pada Gambar (c) dan (d), model memberikan label pada objek yang tidak dilabeli pada ground truth. Kondisi ini menghasilkan false positive, meskipun secara visual prediksi tersebut tampak benar. Temuan ini mengindikasikan bahwa terdapat objek yang sebenarnya ada, namun tidak tercatat pada ground truth, sehingga model dianggap melakukan kesalahan meskipun secara fungsional deteksinya tepat. Sementara itu, pada Gambar (e) dan (f), beberapa objek tidak berhasil dilabeli oleh model, sehingga menimbulkan false negative. Hal ini menunjukkan bahwa dalam kondisi tertentu seperti objek yang saling berhimpitan, pencahayaan kurang, atau ukuran objek yang terlalu kecil model masih memiliki keterbatasan dalam mendeteksi seluruh objek secara konsisten.



(a)



(b)



(c)



(d)



(e) (f)

Gambar 4. 13 Perbandingan Visual Pseudo-label 40%
YOLOv11 (a, c, e) dan YOLOv8 (b, d, f),
label asli (hijau), pseudo-label (merah)

Pada eksperimen dengan proporsi 40% data berlabel, perbandingan pseudo-label dengan ground truth ditampilkan pada Gambar a-f, di mana (a), (c), dan (e) merupakan hasil pseudo-labeling dengan YOLOv11, sedangkan (b), (d), dan (f) menggunakan YOLOv8. Pada Gambar (a) dan (b), model tampak memberikan label pada objek yang tidak tercantum dalam ground truth. Situasi ini menghasilkan adanya false positive, meskipun secara visual prediksi tersebut sebenarnya dapat dikatakan benar. Kejadian ini menunjukkan bahwa terdapat objek yang seharusnya masuk dalam ground truth tetapi tidak dilabeli, sehingga model dinilai melakukan kesalahan walaupun secara visual deteksinya relevan.

Pada Gambar (c) dan (d), terlihat bahwa model berhasil memberikan label pada seluruh objek yang sesuai dengan ground truth. Hal ini menandakan bahwa model mampu melakukan deteksi dengan baik pada situasi visual yang jelas serta objek yang tampil dengan bentuk yang mudah dikenali. Sementara itu, pada Gambar (e) dan (f), terdapat objek tertentu yang tidak berhasil terdeteksi oleh model, sehingga menghasilkan false negative. Kondisi tersebut menggambarkan bahwa pada situasi pencahayaan yang kurang ideal model masih mengalami kendala dalam melakukan deteksi secara menyeluruh dan konsisten.



(a) (b)



(c)



(d)



(e)



(f)

Gambar 4. 14 Perbandingan Visual Pseudo-label 60%
YOLOv11 (a, c, e) dan YOLOv8 (b, d, f),
label asli (hijau), pseudo-label (merah)

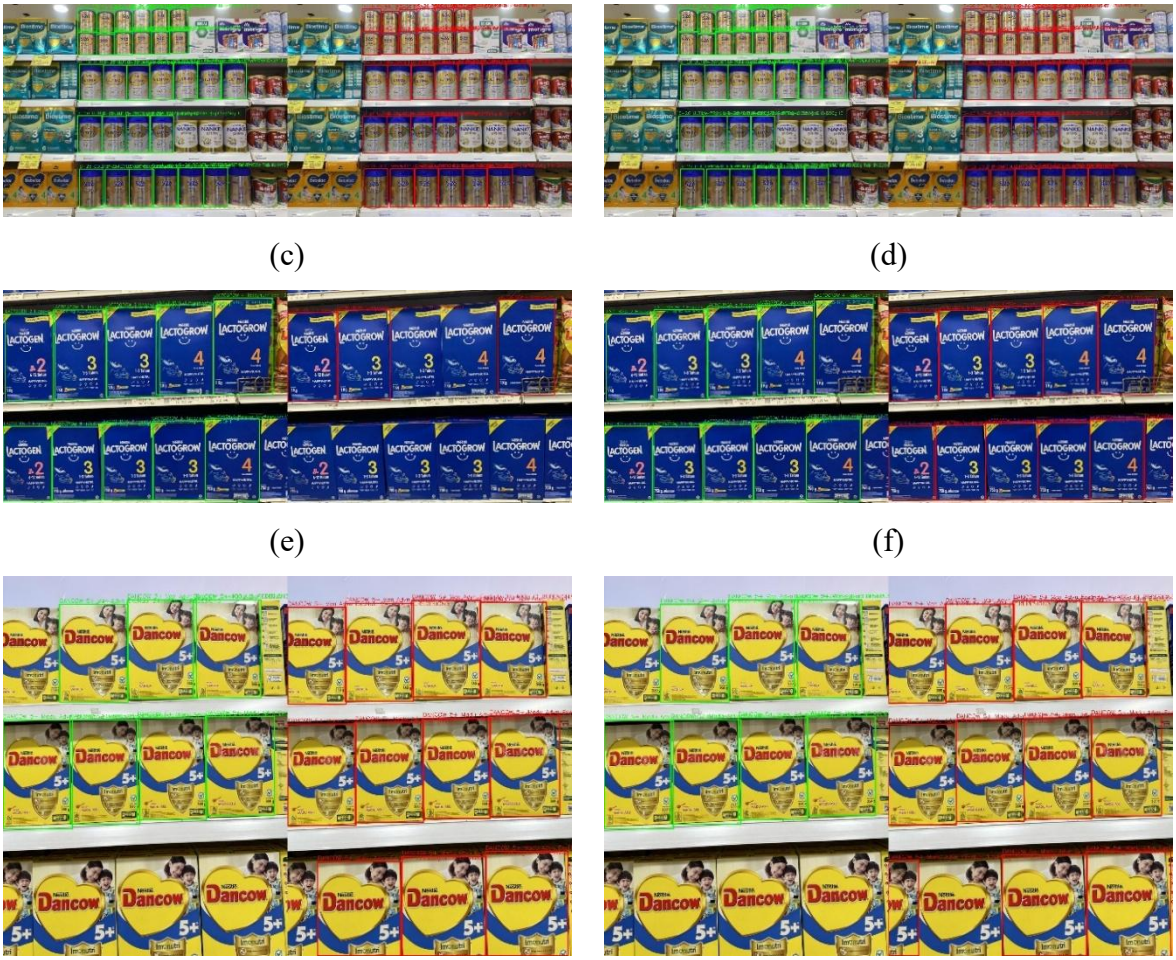
Pada eksperimen dengan proporsi 60% data berlabel, perbandingan pseudo-label dengan ground truth ditunjukkan pada Gambar a-f, dengan (a), (c), dan (e) merupakan hasil pseudo-labeling menggunakan YOLOv11, sedangkan (b), (d), dan (f) menggunakan YOLOv8. Pada Gambar (a) dan (b), pseudo-label mampu memberikan label pada seluruh objek yang telah dilabeli pada ground truth. Namun, terdapat kasus pada kelas susu Batita, di mana pseudo-label menambahkan anotasi meskipun pada ground truth objek tersebut tidak dilabeli. Walaupun hasil pseudo-label ini sebenarnya benar, hal tersebut tetap tercatat sebagai false positive akibat keterbatasan anotasi manual pada ground truth.

Pada Gambar (c) dan (d), pseudo-label dapat melabeli objek dengan hasil yang konsisten serta sesuai dengan ground truth. Hal ini menunjukkan peningkatan kualitas pseudo-label dibandingkan eksperimen dengan proporsi data berlabel yang lebih rendah. Sementara itu, pada Gambar (e) dan (f), pseudo-label berhasil mendeteksi objek, tetapi tidak seluruhnya. Hal ini terjadi karena objek pada gambar tersebut memiliki variasi yang cukup tinggi, sehingga model masih mengalami kesulitan dalam mempertahankan konsistensi pelabelan.





(a) (b)



(c) (d)

(e) (f)

(g) (h)

Gambar 4. 15 Perbandingan Visual Pseudo-label 80%
YOLOv11 (a, c, e, g) dan YOLOv8 (b, d, f, h),
label asli (hijau), pseudo-label (merah)

Pada eksperimen dengan proporsi 80% data berlabel, perbandingan pseudo-label dengan ground truth ditunjukkan pada Gambar a-h, dengan (a, c, e, g) merupakan hasil pseudo-labeling menggunakan YOLOv11, sedangkan (b, d, f, h) menggunakan YOLOv8. Pada Gambar (a) dan (b) model dapat melabeli semua objek, tetapi sama seperti sebelumnya terdapat pada kasus batita yang menimbulkan false walaupun hasilnya benar secara visual. Selanjutnya, pada Gambar (c) dan (d), pseudo-label berhasil memberikan anotasi yang akurat terhadap seluruh objek, sejalan dengan ground truth. Hal ini dimungkinkan karena objek pada gambar terlihat sangat jelas sehingga memudahkan model dalam melakukan deteksi. Pada Gambar (e) terdapat beberapa objek yang tidak berhasil dideteksi oleh model, sehingga menimbulkan false negatif, sedangkan pada gambar (f) model dapat melabeli semua objek yang ada digambar.

Terakhir, pada Gambar (g) dan (h), pseudo-label mampu melabeli seluruh objek yang telah dilabeli pada ground truth. Namun, terdapat tambahan objek yang tidak dilabeli pada ground truth tetapi berhasil dilabeli pada pseudo-label, dan anotasi tersebut sebenarnya benar. Kondisi ini kembali menimbulkan kasus false positive, yang lebih disebabkan oleh keterbatasan anotasi manual pada ground truth, bukan kesalahan dari model.

Tabel 4. 4 Metrik Evaluasi Pseudo-label

Model	Total TP	Total FP	Total FN	Precision	Recall	F1-Score	mAP@0.5	mAP 0.5:0.95	IoU
Eksperimen 20%									
Yolov8	3883	705	1633	0.803	0.677	0.708	0.571	0.506	0.835
Yolov11	4335	798	1172	0.814	0.755	0.763	0.634	0.563	0.859
Eksperimen 40%									
Yolov8	3145	663	708	0.798	0.789	0.77	0.63	0.563	0.87
Yolov11	3124	635	785	0.797	0.77	0.758	0.614	0.548	0.862
Eksperimen 60%									
Yolov8	1934	452	463	0.758	0.748	0.725	0.509	0.453	0.818
Yolov11	1898	383	467	0.771	0.761	0.744	0.526	0.473	0.826
Eksperimen 80%									
Yolov8	998	251	170	0.594	0.642	0.6	0.388	0.35	0.669
Yolov11	928	228	230	0.564	0.567	0.551	0.338	0.303	0.636

Tabel 4.4 menyajikan metrik evaluasi pseudo-label yang dihitung dengan membandingkan hasil pseudo-label terhadap ground truth. Evaluasi dilakukan pada empat proporsi data berlabel, yaitu 20%, 40%, 60%, dan 80% dan menghitung persentase dari objek yang ada. Pada eksperimen dengan proporsi 20% data berlabel, YOLOv8 menghasilkan total 3883 True Positive (62.42%), 705 False Positive (11.33%), dan 1633 False Negative (26.25%), dengan precision 0.803 dan recall 0.677. Hal ini menunjukkan bahwa sebagian besar pseudo-label yang dihasilkan sesuai dengan ground truth, meskipun masih terdapat kesalahan yang cukup tinggi pada objek yang tidak terdeteksi (FN). Sementara itu, YOLOv11 pada proporsi yang sama menunjukkan kinerja yang lebih baik, dengan 4335 TP (68.75%), 798 FP (12.66%), dan 1172 FN (18.59%). Precision dan recall lebih tinggi dibanding YOLOv8, yaitu 0.814 dan 0.755, sehingga menghasilkan F1-Score 0.763 dan mAP@0.5 sebesar 0.634. Secara keseluruhan, YOLOv11 menunjukkan pseudo-label yang lebih konsisten pada kondisi data terbatas.

Pada eksperimen 40%, kualitas pseudo-label meningkat pada kedua model. YOLOv8 mencatat 3145 TP (69.64%), 663 FP (14.68%), dan 708 FN (15.68%), dengan precision 0.798

dan recall 0.789. Hasil ini menunjukkan peningkatan keseimbangan antara TP dan FN dibanding eksperimen 20%. YOLOv11 juga menunjukkan pola serupa dengan 3124 TP (68.75%), 635 FP (13.97%), dan 785 FN (17.28%). Precision dan recall keduanya mendekati YOLOv8, dengan F1-Score 0.758 dan mAP@0.5 sebesar 0.614. Pada proporsi ini, kedua model menampilkan kualitas pseudo-label yang relatif stabil dan saling mendekati.

Pada eksperimen 60%, terjadi sedikit penurunan pada beberapa metrik. YOLOv8 menghasilkan 1934 TP (67.88%), 452 FP (15.87%), dan 463 FN (16.25%). Precision dan recall masing-masing sebesar 0.758 dan 0.748, menunjukkan keseimbangan yang cukup baik namun tidak setinggi proporsi 40%. YOLOv11 memberikan hasil yang serupa, dengan 1898 TP (69.07%), 383 FP (13.94%), dan 467 FN (16.99%). Precision dan recall sedikit lebih tinggi dibanding YOLOv8, menghasilkan F1-Score 0.744 serta mAP@0.5 sebesar 0.526. Secara umum, meskipun TP tetap tinggi, peningkatan FP menunjukkan bahwa model lebih sensitif namun kurang selektif pada proporsi ini.

Pada eksperimen 80%, kedua model menunjukkan penurunan performa pseudo-label paling signifikan. YOLOv8 mencatat 998 TP (70.33%), 251 FP (17.69%), dan 170 FN (11.98%), dengan precision 0.594 dan recall 0.642. Meskipun persentase TP masih tinggi, peningkatan FP menurunkan precision secara drastis. Sementara itu, YOLOv11 menghasilkan 928 TP (66.96%), 228 FP (16.45%), dan 230 FN (16.59%), dengan precision 0.564 dan recall 0.567. Nilai F1-Score dan mAP juga menunjukkan penurunan tajam. Pada proporsi ini, pseudo-label menjadi kurang stabil, kemungkinan akibat kompleksitas visual yang lebih luas yang belum sepenuhnya terwakili dalam data pelatihan sebelumnya.

Berdasarkan hasil evaluasi pseudo-label pada Tabel 4.4, terlihat bahwa setiap proporsi data berlabel menghasilkan jumlah false positive (FP) yang bervariasi. Untuk memahami karakteristik munculnya kesalahan tersebut, dilakukan proses klasterisasi false positive, yaitu pengelompokan kasus FP berdasarkan kesamaan pola visual dan kondisi kemunculannya. Hasil klasterisasi memperlihatkan bahwa false positive dalam pseudo-label dapat dikategorikan ke dalam empat klaster utama sebagai berikut:

a. False postive akibat ketidakkonsistenan Ground Truth

   Klaster ini mencakup objek yang sebenarnya ada pada gambar, tetapi tidak tercatat dalam ground truth karena keterbatasan anotasi manual. Kondisi ini banyak terjadi pada kelas dengan jumlah objek tinggi. Walaupun pseudo-label secara visual akurat, prediksi tetap tercatat sebagai FP karena tidak ada anotasi pembanding pada ground truth.

b. False positive akibat Kemiripan Visual Antar Produk

Beberapa produk memiliki warna, bentuk, atau pola kemasan yang sangat mirip, sehingga model memberi label pada objek dengan kelas yang salah. Kesalahan ini lebih sering terjadi pada eksperimen dengan proporsi data berlabel rendah, karena representasi visual pada proses pelatihan belum cukup beragam.

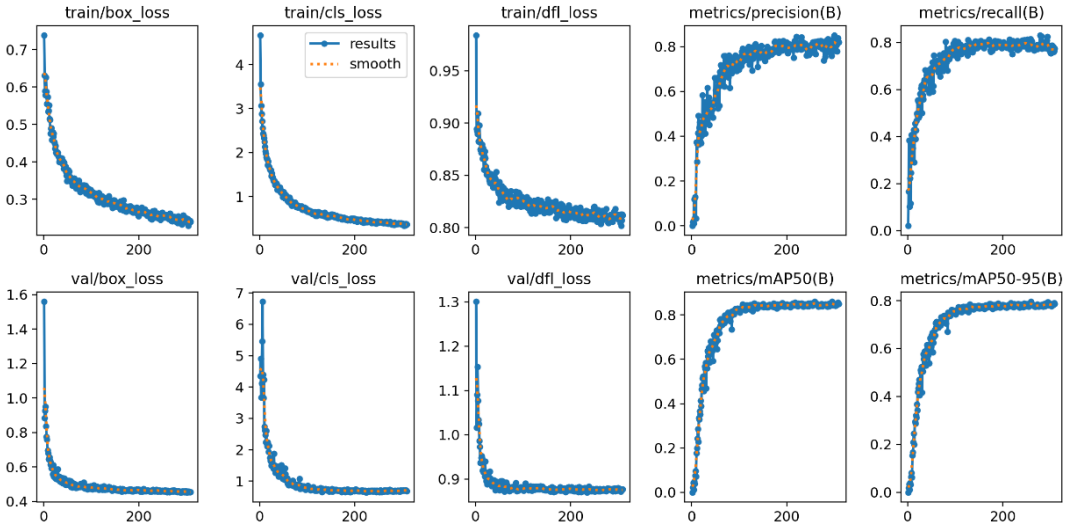
c. False positive akibat Noise Visual dan Lingkungan

Termasuk pantulan cahaya, bayangan rak, pola latar belakang, atau elemen visual lain yang menyerupai tekstur produk. Pada kondisi tertentu, model salah mengenali pola tersebut sebagai objek sehingga menghasilkan anotasi yang tidak sesuai. Meskipun tidak dominan, klaster ini masih muncul terutama pada gambar dengan pencahayaan yang tidak merata.

d. False positive pada Objek di Tepi Gambar

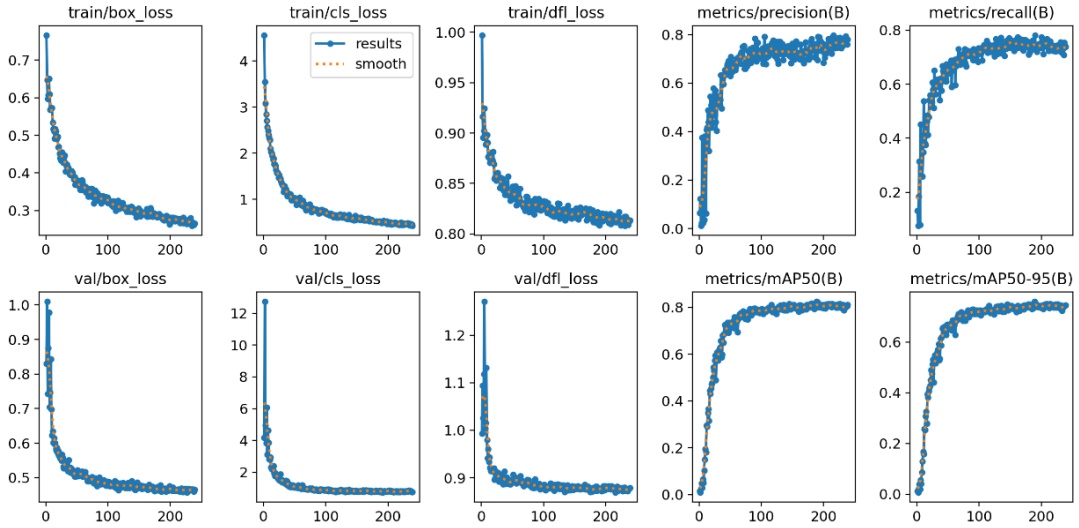
Klaster ini terdiri dari kasus di mana objek hanya terlihat sebagian karena berada di tepi frame, sehingga tidak terlihat secara utuh oleh manusia. Model tetap memberikan label karena masih mampu mengenali sebagian ciri visual objek, namun karena objek tersebut tidak dilabeli pada ground truth, prediksi tersebut tercatat sebagai FP. Dalam beberapa kasus, pseudo-label pada objek tepi gambar ini sebenarnya berpotensi benar, tetapi absennya anotasi menyebabkan model dianggap melakukan kesalahan.

4.3.2 Hasil Training Pseudo Model



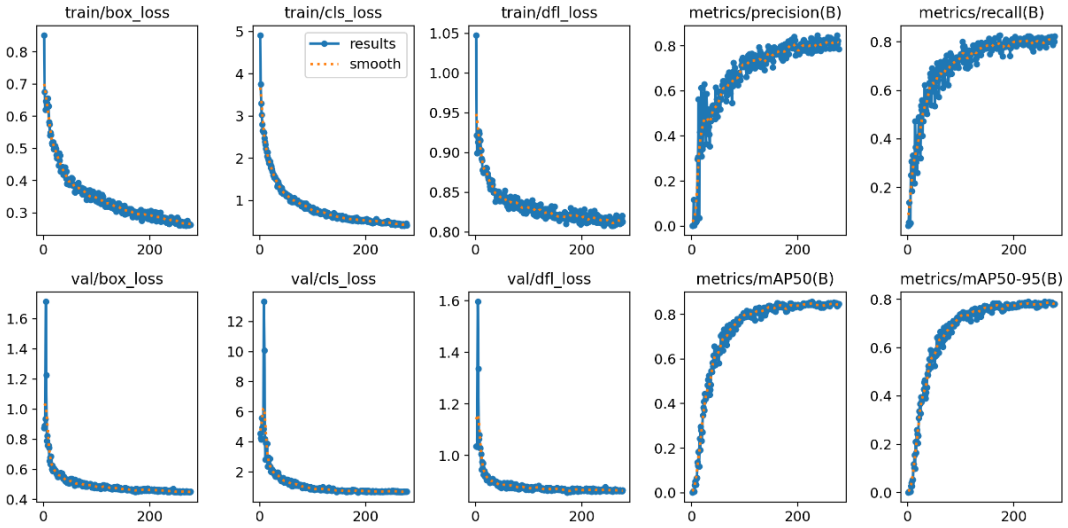
Gambar 4. 16 Pseudo Model 20% Yolov11

Pada eksperimen Pseudo Model 20% menggunakan YOLOv11, grafik training memperlihatkan penurunan nilai box loss, cls loss, dan DFL loss yang konsisten sejak awal pelatihan, dengan percepatan penurunan yang cukup tajam pada fase awal sebelum akhirnya melambat di sekitar epoch ke-100 sebagai tanda model mulai mencapai konvergensi. Seiring menurunnya seluruh komponen loss, metrik precision, recall, mAP50, dan mAP50-95 juga mengalami peningkatan signifikan pada awal proses training dan kemudian bergerak lebih landai setelah mendekati epoch ke-100. Tidak tampak fluktuasi besar pada kurva loss maupun metrik evaluasi, sehingga pelatihan berlangsung stabil meskipun model hanya dilatih menggunakan pseudo-label. Secara keseluruhan, grafik menunjukkan bahwa YOLOv11 masih mampu mengekstraksi informasi pembelajaran yang efektif dari 20% pseudo-label, walaupun saturasi performa mulai terlihat setelah epoch ke-100.



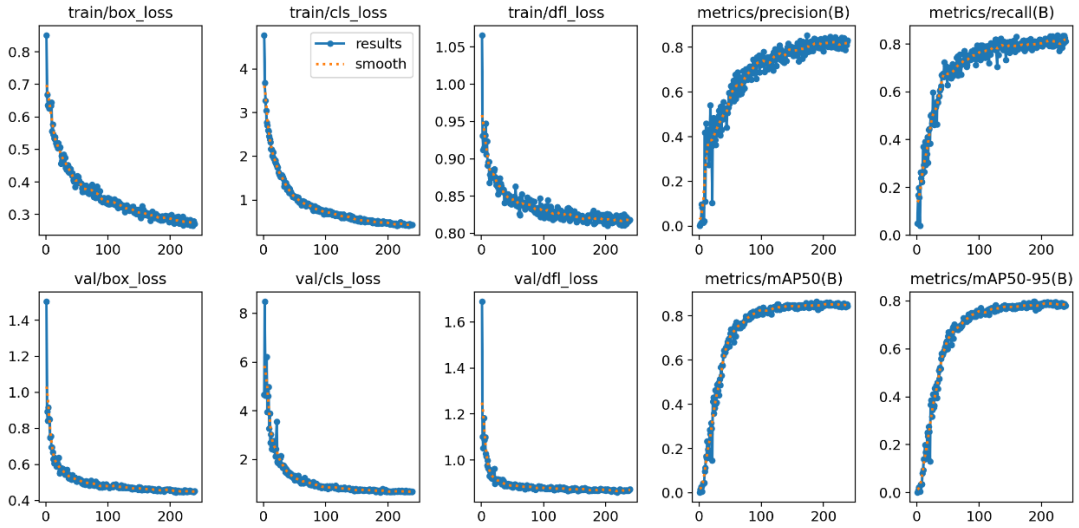
Gambar 4. 17 Pseudo Model 20% Yolov8

Pada eksperimen Pseudo Model 20% menggunakan YOLOv8, grafik training menunjukkan penurunan box loss, cls loss, dan DFL loss yang konsisten sejak awal pelatihan dengan tren penurunan yang cukup cepat pada fase awal, kemudian mulai melambat di sekitar epoch ke-80 sebagai indikasi bahwa model mulai mencapai konvergensi. Sejalan dengan menurunnya nilai loss, metrik precision, recall, mAP50, dan mAP50-95 mengalami peningkatan stabil dan progresif, tanpa fluktuasi signifikan yang mengganggu pola pembelajaran. Pola kurva yang relatif mulus mengindikasikan bahwa training berlangsung stabil dan model mampu memanfaatkan pseudo-label dengan baik meskipun kualitas label tidak seakurat label manual. Secara keseluruhan, grafik menggambarkan bahwa YOLOv8 masih dapat belajar efektif dari 20% pseudo-label, namun peningkatan performa mulai melambat saat mendekati epoch ke-80.



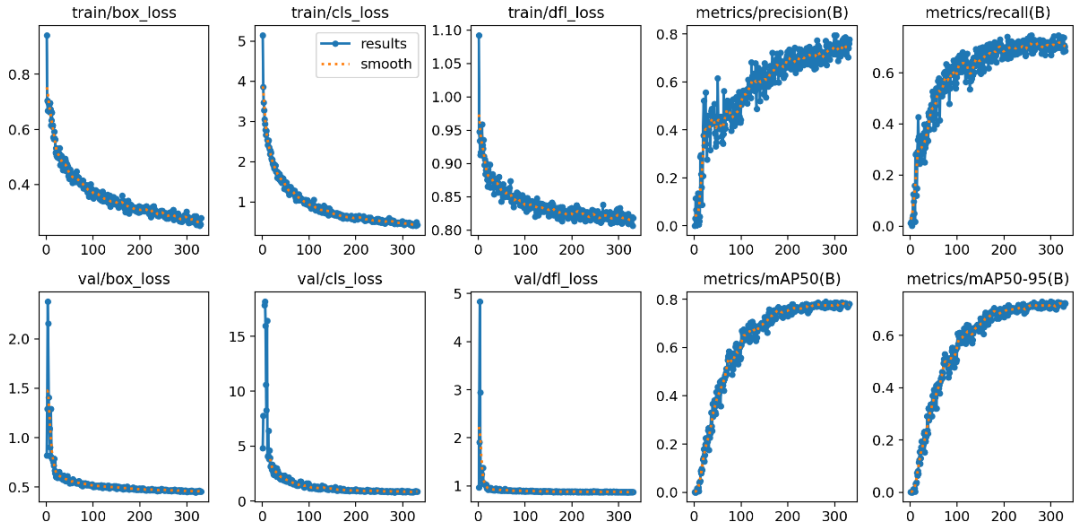
Gambar 4. 18 Pseudo Model 40% Yolov11

Pada grafik training Pseudo Model 40% menggunakan YOLOv11, kurva box loss, cls loss, dan DFL loss menunjukkan penurunan yang konsisten sejak awal dan mulai melambat setelah sekitar epoch ke-100, menandakan bahwa model telah memasuki fase konvergensi. Pada bagian metrik evaluasi, precision sempat mengalami fluktuasi pada fase awal training, mencerminkan proses adaptasi model terhadap pseudo-label yang bervariasi kualitasnya, namun setelah fase awal tersebut precision bergerak stabil dengan tren peningkatan yang mulus hingga mendekati akhir epoch. Metrik lain seperti recall, mAP50, dan mAP50-95 juga menunjukkan pola peningkatan progresif yang sejalan dengan penurunan loss, sehingga secara keseluruhan grafik mencerminkan proses pembelajaran yang stabil dengan performa yang terus menguat meskipun pelatihan hanya menggunakan pseudo-label.



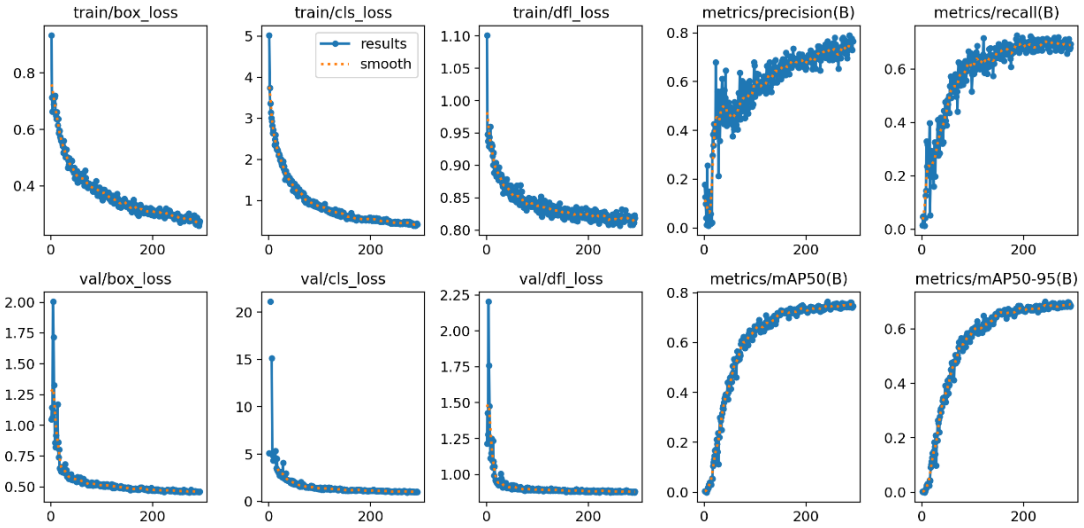
Gambar 4. 19 Pseudo Model 40% Yolov8

Pada grafik training Pseudo Model 40% menggunakan YOLOv8, kurva box loss, cls loss, dan DFL loss menunjukkan penurunan yang stabil sejak awal hingga akhirnya mulai melambat pada sekitar epoch ke-80, menandakan bahwa model telah mencapai fase konvergensi lebih cepat dibandingkan konfigurasi lain. Pada bagian metrik evaluasi, precision mengalami sedikit fluktuasi di awal training akibat proses penyesuaian model terhadap pseudo-label, namun setelah beberapa epoch awal precision bergerak stabil dengan tren peningkatan yang konsisten hingga akhir proses pelatihan. Sementara itu recall, mAP50, dan mAP50-95 juga memperlihatkan kenaikan progresif yang sejalan dengan penurunan loss, sehingga keseluruhan grafik mencerminkan pembelajaran yang tetap stabil meskipun hanya menggunakan pseudo-label sebagai data pelatihan.



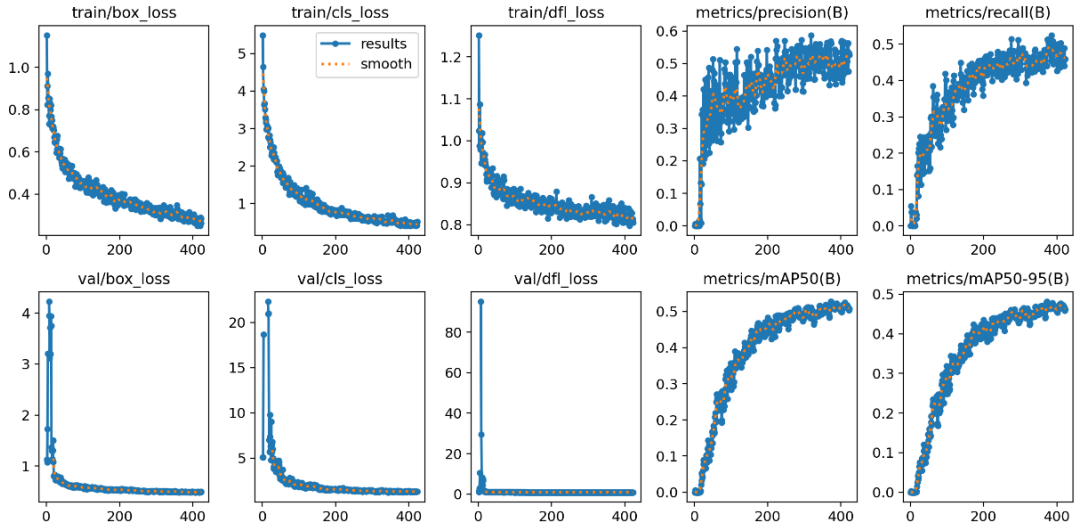
Gambar 4. 20 Pseudo model 60% Yolov11

Pada grafik training Pseudo Model 60% menggunakan YOLOv11, penurunan box loss, cls loss, dan DFL loss berlangsung stabil sejak awal hingga akhirnya mulai melambat pada sekitar epoch ke-200 sebagai tanda dimulainya fase konvergensi. Meskipun tren penurunan loss terlihat konsisten, metrik precision dan recall menunjukkan sedikit fluktuasi hingga akhir pelatihan akibat ketidakstabilan pseudo-label pada sebagian sampel, namun pola keduanya tetap meningkat secara umum seiring bertambahnya epoch. Selain itu, metrik mAP50 dan mAP50-95 terus mengalami peningkatan progresif tanpa penurunan tajam, sehingga secara keseluruhan grafik memperlihatkan kemampuan model untuk belajar secara efektif dari pseudo-label meskipun terdapat dinamika kecil pada metrik performa.



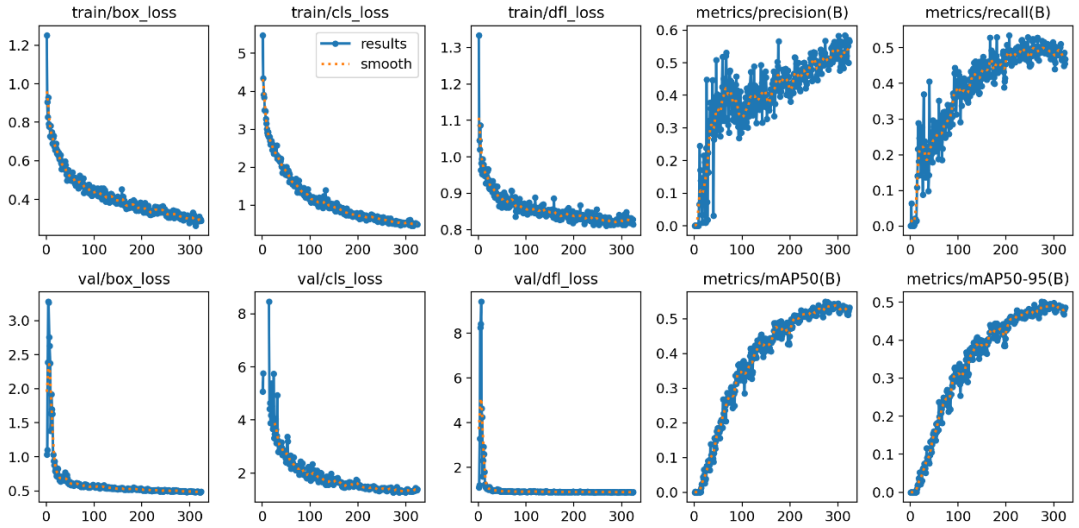
Gambar 4. 21 Pseudo Model 60% Yolov8

Pada grafik training Pseudo Model 60% menggunakan YOLOv8, kurva box loss, cls loss, dan DFL loss menurun secara konsisten sejak awal hingga mulai melambat sekitar epoch ke-150, menandakan bahwa model mulai memasuki fase konvergensi. Pada metrik precision, terlihat fluktuasi yang cukup signifikan pada awal pelatihan hingga sekitar epoch ke-50, kemungkinan karena model menyesuaikan diri terhadap pseudo-label yang memiliki variasi kualitas. Setelah epoch ke-50, precision bergerak stabil dengan tren peningkatan yang konsisten hingga akhir pelatihan. Metrik recall, mAP50, dan mAP50-95 juga meningkat secara progresif seiring penurunan loss, sehingga secara keseluruhan grafik menunjukkan bahwa YOLOv8 mampu mempelajari pseudo-label dengan efektif meskipun terdapat fluktuasi awal pada precision.



Gambar 4. 22 Pseudo model 80% Yolov11

Pada grafik training Pseudo Model 80% menggunakan YOLOv11, penurunan box loss, cls loss, dan DFL loss berlangsung cukup konsisten sejak awal dan mulai melambat sekitar epoch ke-100, menunjukkan model mulai mendekati fase konvergensi pada loss. Meskipun demikian, metrik precision dan recall mengalami fluktuasi yang cukup signifikan hingga akhir pelatihan, , namun secara keseluruhan kedua metrik tersebut tetap menunjukkan tren peningkatan. Metrik mAP50 dan mAP50-95 juga terus meningkat, menandakan bahwa model tetap mampu belajar dari pseudo-label meskipun terdapat dinamika fluktuasi pada precision dan recall.



Gambar 4. 23 Pseudo model 80% Yolov8

Pada grafik training Pseudo Model 80% menggunakan YOLOv8, penurunan box loss, cls loss, dan DFL loss terjadi cukup stabil sejak awal hingga mulai melambat pada sekitar epoch ke-100, menandakan fase awal konvergensi pada loss. Metrik mAP50 dan mAP50-95 menunjukkan tren peningkatan yang konsisten, namun mulai melambat sekitar epoch ke-200 seiring model mencapai kestabilan dalam performa deteksi. Sementara itu, precision dan recall mengalami fluktuasi yang cukup signifikan sepanjang pelatihan, kemungkinan disebabkan oleh variasi kualitas pseudo-label, tetapi secara keseluruhan pola keduanya tetap meningkat. Hal ini menunjukkan bahwa meskipun terdapat dinamika fluktuasi, model tetap mampu mempelajari pseudo-label dengan efektif dan meningkatkan kemampuan deteksinya.

Tabel 4. 5 Hasil Training Pseudo Model

Model	Epocs	Box Loss	Cls Loss	DFL Loss	Precision	Recall	mAP50	mAP50-95	Inference Time	Waktu Training (Jam)
Eksperimen 20%										
Yolov8	239	0.265	0.426	0.813	0.762	0.744	0.825	0.76	4.7ms	0.505
Yolov11	309	0.241	0.368	0.811	0.8	0.798	0.856	0.794	4.8ms	0.699
Eksperimen 40%										
Yolov8	239	0.271	0.429	0.818	0.808	0.813	0.864	0.797	4.7ms	0.403
Yolov11	277	0.263	0.410	0.813	0.833	0.789	0.857	0.791	4.6ms	0.501
Eksperimen 60%										
Yolov8	292	0.275	0.422	0.818	0.766	0.699	0.764	0.7	4.8ms	0.374
Yolov11	331	0.278	0.428	0.818	0.771	0.689	0.791	0.728	5.5ms	0.465
Eksperimen 80%										
Yolov8	324	0.289	0.490	0.814	0.567	0.489	0.55	0.502	4.5ms	0.276
Yolov11	423	0.284	0.520	0.805	0.561	0.44	0.527	0.48	4.7ms	0.387

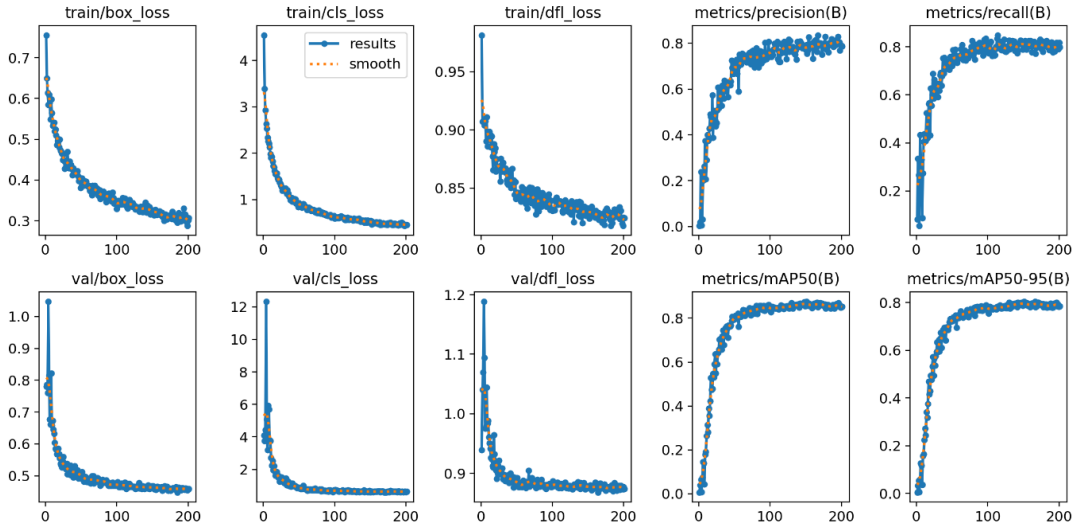
Pada eksperimen 20% dengan 587 gambar pseudo-label, model YOLOv11 menunjukkan performa paling optimal dengan mAP50 sebesar 0.856 dan mAP50-95 sebesar 0.794. Precision dan recall yang sama-tinggi (0.800 dan 0.798) mencerminkan kemampuan model dalam mengidentifikasi objek sekaligus mempertahankan tingkat kesalahan prediksi yang rendah. Meskipun hanya menggunakan pseudo-label tanpa data berlabel asli, kuantitas data pada proporsi 20% masih mampu memberikan informasi yang cukup bagi model untuk belajar secara efektif. Pada eksperimen 40% dengan 440 gambar pseudo-label, performa terbaik dicapai oleh model YOLOv8 dengan nilai mAP50 0.864 dan mAP50-95 0.797. Precision dan recall yang relatif seimbang (0.808 dan 0.813) menunjukkan model mampu mendeteksi objek dengan stabil. Jumlah data pseudo-label pada proporsi ini masih cukup untuk memperkuat proses pembelajaran tanpa menyebabkan dominasi pseudo-label yang berkualitas rendah.

Eksperimen 60% dengan 294 pseudo-label menunjukkan penurunan performa dibanding proporsi sebelumnya. YOLOv11 memperoleh mAP50 0.791 dan mAP50-95 0.728, precision 0.771, dan recall 0.689. Penurunan recall mengindikasikan bahwa model mulai kesulitan dalam menangkap seluruh objek pada data uji. Hal ini dapat dikaitkan dengan semakin berkurangnya variasi data pseudo-label yang tersedia bagi model selama pelatihan. Pada eksperimen 80% dengan 147 pseudo-label, performa model menurun secara signifikan. YOLOv8 mendapatkan mAP50 0.550 dan mAP50-95 0.502, sedangkan YOLOv11 mencatat mAP50 0.527 dan mAP50-95 0.480. Precision dan recall pada kedua model berada di bawah 0.60, mengindikasikan model kesulitan melakukan generalisasi. Jumlah pseudo-label yang semakin sedikit menyebabkan informasi yang dipelajari model menjadi terbatas sehingga berdampak pada hasil pengujian. Secara keseluruhan, hasil ini menunjukkan bahwa pseudo-label dapat membantu performa model pada proporsi menengah kebawah (20% - 40%), tetapi cenderung menurunkan performa ketika proporsinya bertambah besar karena jumlah gambar yang tambah sedikit.

4.4 Implementasi dan Hasil Student Model

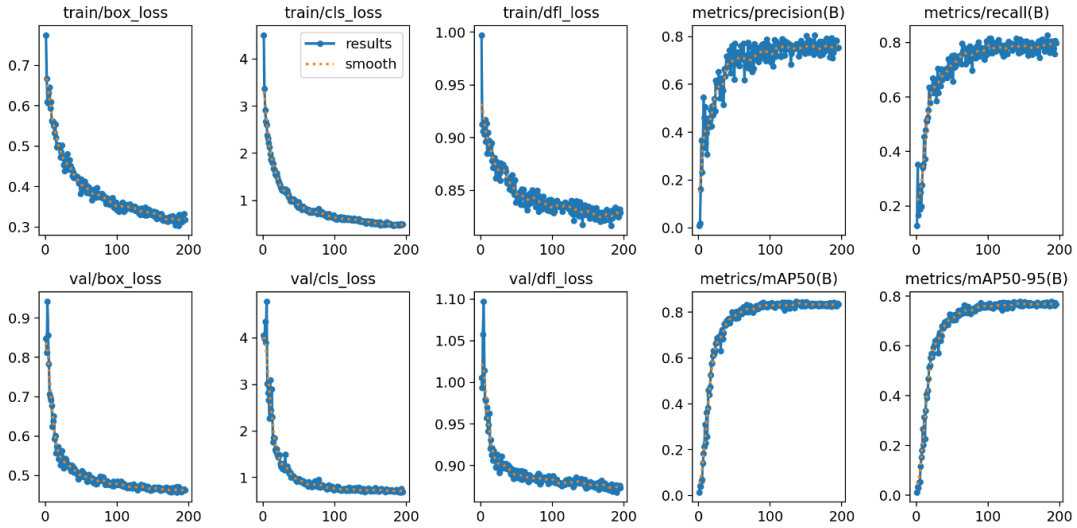
Tahap ketiga dalam penelitian ini adalah implementasi pelatihan model menggunakan pendekatan semi-supervised learning. Pada tahap ini, proses pelatihan dilakukan dengan memanfaatkan kombinasi antara data berlabel dan data tidak berlabel yang telah diberikan pseudo-label oleh model teacher hasil dari pelatihan supervised pada tahap sebelumnya.

Proses pelatihan semi-supervised ini menggunakan parameter dan konfigurasi yang sama seperti pada skenario supervised training yang ada di Gambar 4. 2 pada bab 4.2.



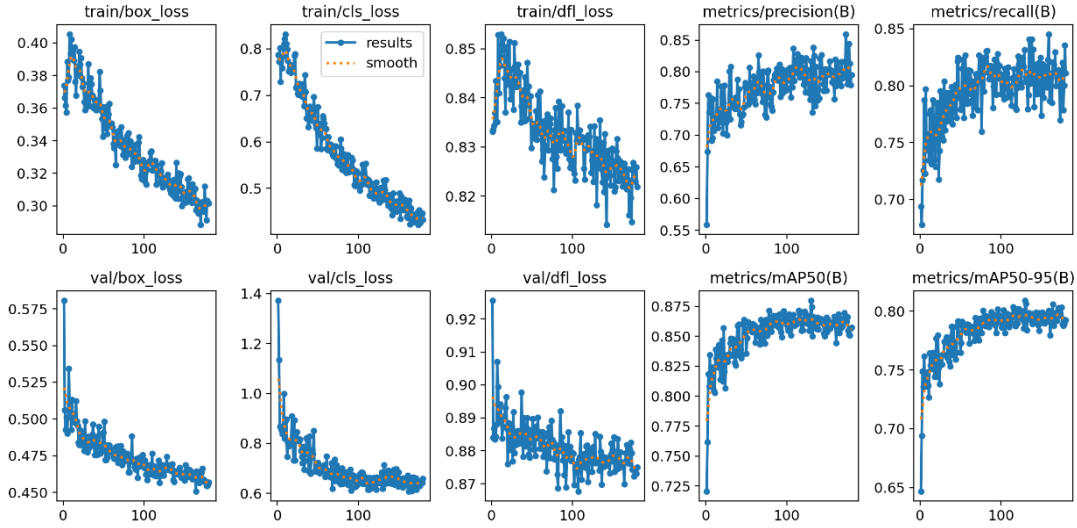
Gambar 4. 24 Student Model 20% dengan Pretrained Model Yolov11

Pada grafik training Student Model 20% yang ditrain menggunakan Pretrained YOLOv11, penurunan box loss, cls loss, dan DFL loss terlihat cukup konsisten sejak awal hingga mulai melambat sekitar epoch ke-80, menandakan model mulai memasuki fase konvergensi. Metrik precision, recall, mAP50, dan mAP50-95 meningkat secara signifikan hingga epoch ke-80 dan kemudian menunjukkan perlambatan pertumbuhan, menggambarkan bahwa model mulai mencapai kestabilan dalam performa deteksi. Secara keseluruhan, grafik menunjukkan bahwa Pretrained YOLOv11 mampu memanfaatkan data 20% yang tersedia secara efektif untuk meningkatkan akurasi deteksi student model.



Gambar 4. 25 Student Model 20% dengan Pretrained Model Yolov8

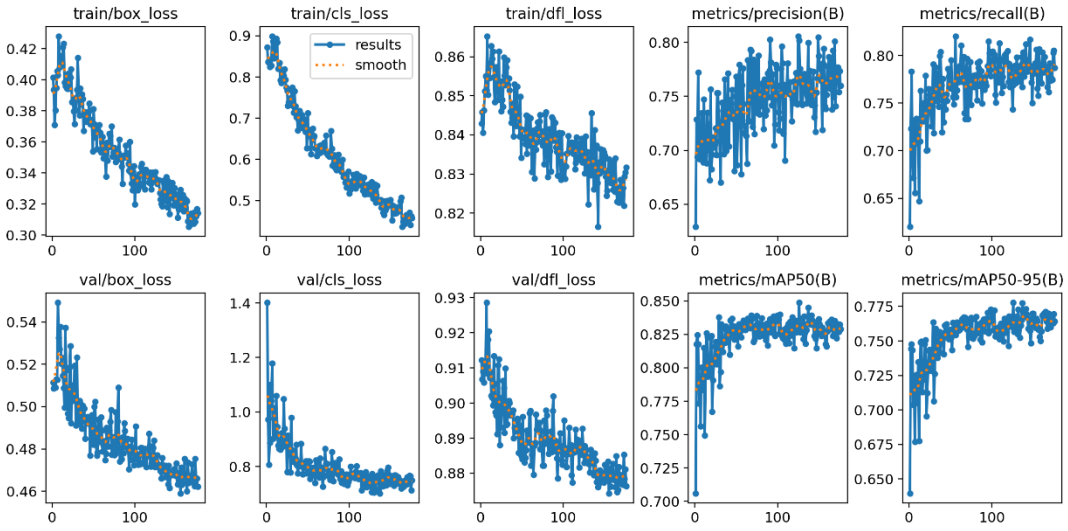
Pada grafik training Student Model 20% yang ditrain menggunakan Pretrained YOLOv8, penurunan box loss, cls loss, dan DFL loss berlangsung cukup stabil sejak awal hingga mulai melambat sekitar epoch ke-70, menandakan model memasuki fase konvergensi. Metrik precision, recall, mAP50, dan mAP50-95 meningkat secara konsisten hingga epoch ke-70 dan kemudian mulai melambat, menunjukkan bahwa model mulai mencapai kestabilan dalam performa deteksi. Grafik ini menandakan bahwa Pretrained YOLOv8 mampu memanfaatkan data 20% secara efektif untuk meningkatkan akurasi student model meskipun jumlah data terbatas.



Gambar 4. 26 Student Model 20% dengan Teacher Model Yolov11

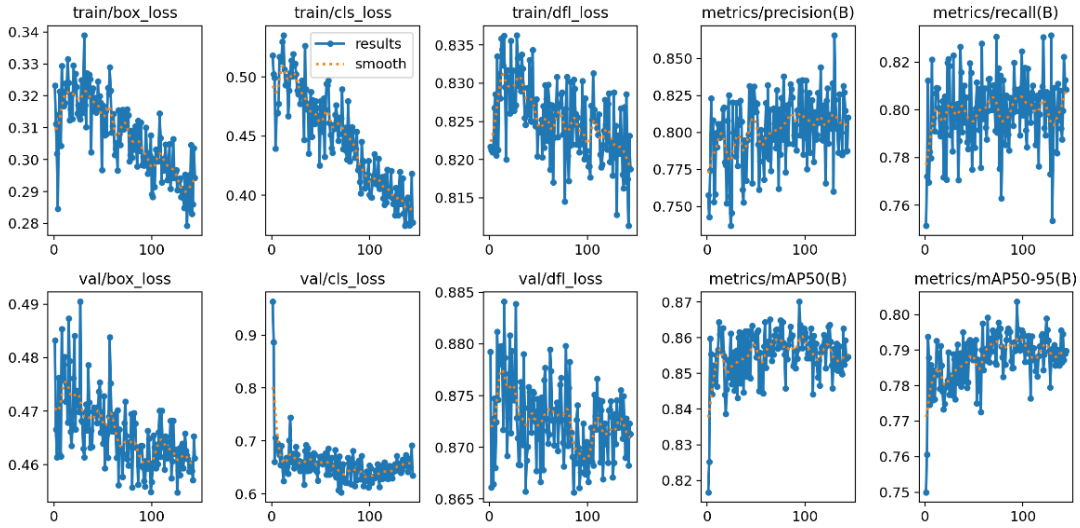
Pada grafik training Student Model 20% yang ditrain menggunakan Teacher Model YOLOv11, terjadi lonjakan awal pada loss di epoch-epoch pertama, kemudian penurunan dan

peningkatan loss terjadi secara perlahan dengan fluktuasi yang cukup signifikan sepanjang pelatihan. Metrik precision dan recall menunjukkan tren naik meskipun terdapat fluktuasi cukup parah, sementara box loss, cls loss, dan DFL loss juga mengalami fluktuasi tetapi secara keseluruhan cenderung menurun. Pola ini menunjukkan bahwa meskipun data 20% terbatas, student model tetap mampu mempelajari informasi dari Teacher YOLOv11, dengan performa deteksi yang meningkat seiring pelatihan meskipun tidak semulus model yang menggunakan pretrained weights.



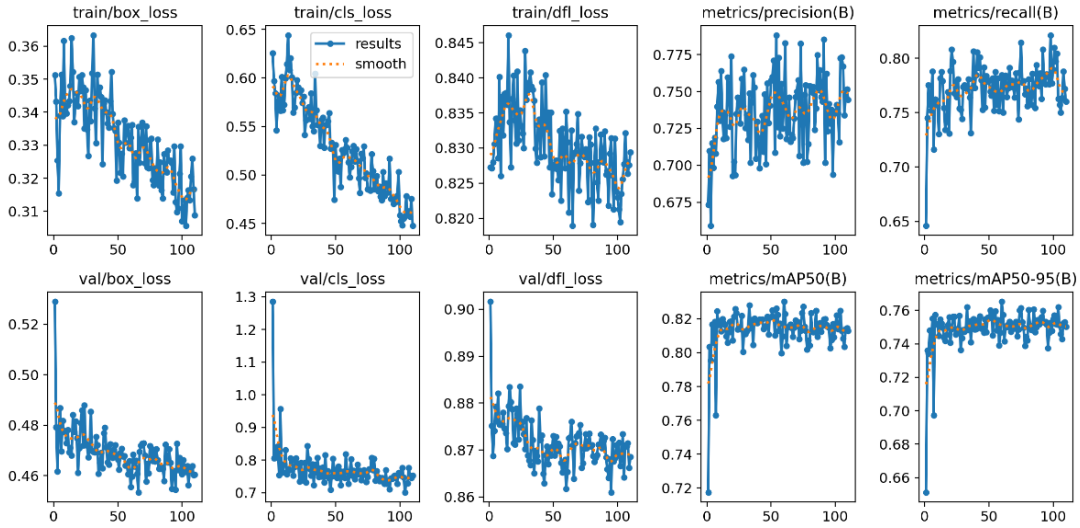
Gambar 4. 27 Student Model 20% dengan Teacher Model Yolov8

Pada grafik training Student Model 20% yang ditrain menggunakan Teacher Model YOLOv8, terlihat lonjakan awal pada loss di epoch-epoch pertama, diikuti dengan penurunan dan peningkatan loss yang terjadi secara perlahan dengan fluktuasi yang lebih parah dibandingkan Teacher YOLOv11. Metrik precision dan recall juga menunjukkan tren naik meskipun mengalami fluktuasi cukup signifikan, sedangkan box loss, cls loss, dan DFL loss tetap menurun secara umum walaupun tidak stabil. Kondisi ini menunjukkan bahwa meskipun data yang tersedia hanya 20%, student model masih mampu mempelajari informasi dari Teacher YOLOv8, namun kestabilan pelatihan lebih rendah dibandingkan menggunakan Teacher YOLOv11.



Gambar 4. 28 Student Model 20% dengan Pseudo Model Yolov11

Pada grafik training Student Model 20% yang ditrain menggunakan Pseudo Model YOLOv11, metrik precision dan recall langsung meningkat tajam hingga sekitar 0,800 pada awal pelatihan. Setelah itu, kedua metrik mengalami fluktuasi cukup signifikan di rentang 0,750 - 0,825 hingga akhir epoch. Pola yang sama juga terlihat pada mAP@50, yang berada di rentang 0,84 - 0,87, serta mAP@50-95 yang berada di kisaran 0,77 - 0,805. Fluktuasi ini menunjukkan bahwa meskipun model belajar dengan cepat dari pseudo-label, kestabilan pelatihan kurang konsisten akibat variasi kualitas pseudo-label dan jumlah data yang terbatas.



Gambar 4. 29 Student Model 20% dengan Pseudo Model Yolov8

Pada grafik training Student Model 20% yang ditrain menggunakan Pseudo Model YOLOv8, metrik precision dan recall menunjukkan fluktuasi yang cukup parah sepanjang epoch, mirip dengan Student Model 20% Pseudo YOLOv11, namun kestabilan mAP sedikit lebih baik. Precision dan recall bergerak naik-turun secara signifikan, sementara mAP@50 berada di rentang yang lebih konsisten dibandingkan versi YOLOv11, dan mAP@50-95 juga tetap menunjukkan pola fluktuatif. Hal ini mengindikasikan bahwa meskipun pseudo-label digunakan sebagai data utama, model masih mampu belajar, tetapi pelatihan kurang stabil karena keterbatasan data dan variasi kualitas pseudo-label.

Tabel 4. 6 Hasil Training Student model 20%

Eksperimen 20%										
Model	Epocs	Box Loss	Cls Loss	DFL Loss	Precision	Recall	mAP50	mAP50-95	Inference Time	Waktu Training (Jam)
Pre Trained Yolov8	195	0.318	0.498	0.829	0.736	0.812	0.845	0.778	4.6ms	0.506
Pre Trained Yolov11	201	0.306	0.458	0.824	0.793	0.815	0.875	0.804	5.9ms	0.571
Teacher Model Yolov8	157	0.314	0.457	0.831	0.802	0.769	0.849	0.779	5.9ms	0.459
Teacher Model Yolov11	180	0.301	0.432	0.821	0.821	0.804	0.88	0.809	5.1ms	0.507
Pseudo Model Yolov8	102	0.308	0.44	0.829	0.741	0.784	0.83	0.765	5.0ms	0.287
Pseudo Model Yolov11	144	0.294	0.377	0.818	0.835	0.806	0.87	0.803	4.7ms	0.408



(a) Pretrained Model Yolov8



(b) Pretrained Model Yolov11



(c) Teacher Model Yolov8



(d) Teacher Model Yolov11



(e) Pseudo Model Yolov8



(f) Pseudo Model Yolov11



(g) Fully Supervised Yolov8



(h) Fully Supervised Yolov11



(i) Ground Truth

Gambar 4. 30 Visual Test Semi Supervised 20%

Tabel 4. 7 Hasil Visual Semi Supervised 20%

Lokasi	Ground Truth	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
A1	Lactogen 1 Happynutri 24x350g	TP	TP	TP	TP	FN - noBB	TP	TP	TP
A2	Lactogen 2 Happynutri 24x350g	TP	TP	TP	TP	TP	TP	TP	TP
A3	Lactogrow 3 Happynutri Van 24x350g	TP	TP	FP – Lactogrow 3 Happynutri 24x350g	TP	FP – Lactogrow 3 Happynutri 24x350g	TP	TP	TP
A4	Lactogen 1 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
A5	Lactogen 2 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
A6	Lactogrow 3 Happynutri Honey 12x750g	TP	TP	TP	TP	TP	TP	TP	TP

B2	Lactogrow 3 Happynutri Honey 24x350g	TP	TP	TP	TP	TP	TP	TP	TP
B3	Lactogrow 3 Happynutri 12x750g	TP	TP	TP	TP	FN-noBB	TP	TP	TP
B4	Lactogrow 3 Happynutri Van 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
B5	Lactogrow 4 Happynutri Honey 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
B6	Lactogrow 4 Happynutri Van 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
C1	Lactogen 1 Happynutri 40x180g	TP	TP	FN - NoBB	TP	FN - noBB	TP	TP	TP
C2	Lactogen 2 Happynutri 40x180g	TP	TP	TP	TP	TP	FN - noBB	TP	TP

Lokasi objek dinyatakan berdasarkan posisi rak dan urutan dari kiri ke kanan, dengan A menunjukkan rak atas, B rak tengah, dan C rak bawah. Angka menunjukkan urutan objek pada rak (misalnya B5 adalah objek ke-5 dari kiri pada rak tengah). Keterangan noBB menunjukkan tidak terdapat bounding box pada objek tersebut, sedangkan FP – … menunjukkan objek yang terdeteksi secara keliru oleh model.

Berdasarkan hasil pelatihan student model dengan proporsi 20% data berlabel, performa keenam model menunjukkan variasi yang dipengaruhi oleh keterbatasan jumlah sampel berlabel. Untuk menganalisis performa visual model secara komprehensif, dilakukan pengujian pada gambar yang sama untuk seluruh strategi, dan hasilnya dibandingkan dengan ground truth

serta hasil fully supervised. Perbandingan visual disajikan pada Gambar 4.30, dengan analisis detail deteksi per objek pada Tabel 4. 7.

Berdasarkan Tabel 4. 7, terlihat bahwa performa deteksi bervariasi antar model:

a. Model (a) Pretrained YOLOv8 berhasil mendeteksi seluruh 13 objek yang terdapat pada ground truth dengan benar. Namun, masih terdapat beberapa false positive. Pada objek Lactogrow 3 Happynutri 24×350 g (B1), objek tersebut tidak diberi label maupun bounding box pada ground truth, tetapi terdeteksi secara keliru sebagai Lactogrow 3 Happynutri Honey 40×180 g. Pola kesalahan yang sama juga terjadi pada produk Vidoran yang berada di sudut kiri atas, yang tidak dilabeli pada ground truth dan terdeteksi sebagai Dancow Baby 2 Box 12×550 g. Selain itu, terdapat false positive pada objek Dancow 1+ (C5), Dancow 3+ (C6), dan SGM di sudut kanan atas, yang tidak memiliki label dan bounding box pada ground truth, tetapi tetap terdeteksi oleh model dengan kelas yang benar. Objek-objek tersebut tidak dilabeli pada ground truth karena hanya terlihat sebagian atau berada pada sudut pengambilan gambar yang menyebabkan informasi visual tidak cukup jelas untuk proses anotasi. Oleh karena itu, meskipun secara visual deteksi model tampak benar, objek-objek tersebut tetap dikategorikan sebagai false positive karena tidak tercantum dalam anotasi ground truth yang digunakan sebagai acuan evaluasi.
b. Model (b) *Pretrained YOLOv11* menunjukkan performa terbaik dengan berhasil mendeteksi seluruh 13 objek pada *ground truth*. Meskipun demikian, masih terdapat beberapa false positive pada objek Lactogrow 3 Happynutri Honey 40×180 g (B1), SGM di sudut kanan atas, dan Dancow 1+ Van Advn ExcNutr 24×400 g (C5). Objek-objek tersebut secara visual terdeteksi dengan benar, namun tidak dilabeli dan tidak memiliki *bounding box* pada *ground truth* karena hanya terlihat sebagian, sehingga tetap dikategorikan sebagai *false positive* berdasarkan acuan evaluasi yang digunakan.
c. Model (c) Teacher YOLOv8 berhasil mendeteksi 12 objek dari ground truth, dengan satu false negative pada produk Lactogen 1 Happynutri 40×180 g (C1). Selain itu, masih terdapat beberapa false positive. Pada objek Lactogrow 3 Happynutri Van 24×350 g (A3), objek tersebut tidak diberi label maupun bounding box pada ground truth dan terdeteksi secara keliru sebagai Lactogrow 3 Happynutri 24×350 g. Dan terjadi juga false positive pada objek Dancow 1+ Van Advn ExcNutr 24×400 g (C5) dan Lactogrow 3 Happynutri Honey 40×180 g (B1), yang tidak diberi label dan

bounding box pada ground truth. Meskipun secara visual deteksi pada objek-objek tersebut sudah benar, baik dari sisi kelas maupun posisi, objek-objek tersebut tetap dikategorikan sebagai false positive karena tidak tercantum dalam anotasi ground truth akibat berada pada sudut pengambilan gambar yang menyebabkan informasi visual tidak cukup jelas untuk proses anotasi.

d. Model (d) Teacher YOLOv11 berhasil mendeteksi 13 objek pada ground truth. Terdapat satu false positive, yaitu pada objek dan Dancow 1+ Van Advn ExcNutr 24x400g (B5). Objek tersebut dideteksi dengan kelas dan posisi yang benar oleh model, namun tidak diberi label maupun bounding box pada ground truth karena keterbatasan visibilitas objek.
e. Model (e) *Pseudo YOLOv8* hanya mampu mendeteksi 10 objek pada *ground truth*, dengan tiga false negative pada objek Lactogen 1 Happynutri 24×350 g (A1)**,** Lactogrow 3 Happynutri 12×750 g (B3), dan Lactogen 1 Happynutri 40×180 g (C1). Selain itu, masih terdapat beberapa false positive**.** Pada objek Lactogrow 3 Happynutri Van 24×350 g (A3), objek tersebut terdeteksi sebagai Lactogrow 3 Happynutri 24×350 g. False positive juga terjadi pada objek Lactogrow 3 Happynutri Honey 40×180 g (B1) dan Dancow 3+ Madu Advn ExcNutr 24×400 g (C6), objek tersebut tidak diberi label dan bounding box pada ground truth tetapi dapat terdeteksi dengan baik. Objek tersebut tidak dilabeli karena informasi yang terlihat pada objek terbatas seperti pada tulisan gramasi produk. Oleh karena itu, objek-objek tersebut tetap dikategorikan sebagai *false positive* berdasarkan acuan evaluasi yang digunakan.
f. Model (f) Pseudo YOLOv11 berhasil mendeteksi 12 objek pada ground truth, dengan satu false negative pada produk Lactogen 2 Happynutri 40x180g (C2). Selain itu, terdapat false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 1+ Van Advn ExcNutr 24x400g (C5), dan Dancow 3+ (C6), yang tidak dilabeli dan tidak diberi bounding box karena terdapat keterbatasan informasi seperti tulisan gramasi yang tidak terlihat semua tetapi dapat dideteksi dengan baik oleh model.

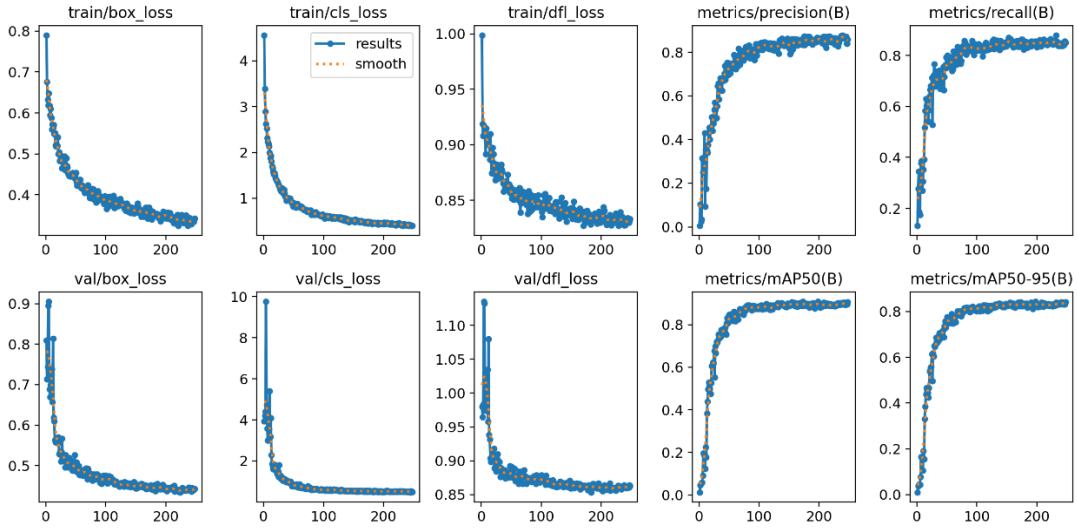
Pada eksperimen semi-supervised learning dengan 20% data berlabel, evaluasi performa model dilakukan menggunakan metrik precision, recall, mAP50, mAP50–95, inference time, serta training time, yang kemudian dianalisis secara kualitatif melalui hasil visual deteksi pada Gambar 4.30(a–h). Model Pretrained YOLOv11 menunjukkan performa yang cukup baik

dengan nilai precision 0,793, recall 0,815, serta mAP50 dan mAP50–95 masing-masing sebesar 0,875 dan 0,804, sejalan dengan hasil visual pada Gambar 4.30(b) yang mampu mendeteksi seluruh 13 objek ground truth secara benar. Namun, masih ditemukan false positive pada beberapa produk yang secara visual terdeteksi dengan benar tetapi tidak dilabeli pada ground truth karena objek tidak terlihat sepenuhnya. Kondisi serupa juga terlihat pada model Pretrained YOLOv8 (Gambar 4.30(a)) yang memiliki nilai precision dan recall lebih rendah (0,736 dan 0,812) akibat kesalahan klasifikasi pada objek yang terpotong atau tertutup sebagian.

Pada model semi-supervised yang di inisiasi dengan teacher model, baik YOLOv8 maupun YOLOv11 menunjukkan peningkatan kestabilan deteksi, yang tercermin dari nilai precision di atas 0,80 dan mAP50 hingga 0,88 pada Teacher YOLOv11. Secara visual pada Gambar 4.30(d), Teacher YOLOv11 mampu mendeteksi 13 dari 13 objek ground truth, dengan false positive yang masih didominasi oleh objek yang tidak sepenuhnya terlihat sehingga tidak tercantum dalam ground truth. Hal ini mengindikasikan bahwa kesalahan model lebih banyak disebabkan oleh keterbatasan anotasi daripada kegagalan deteksi.

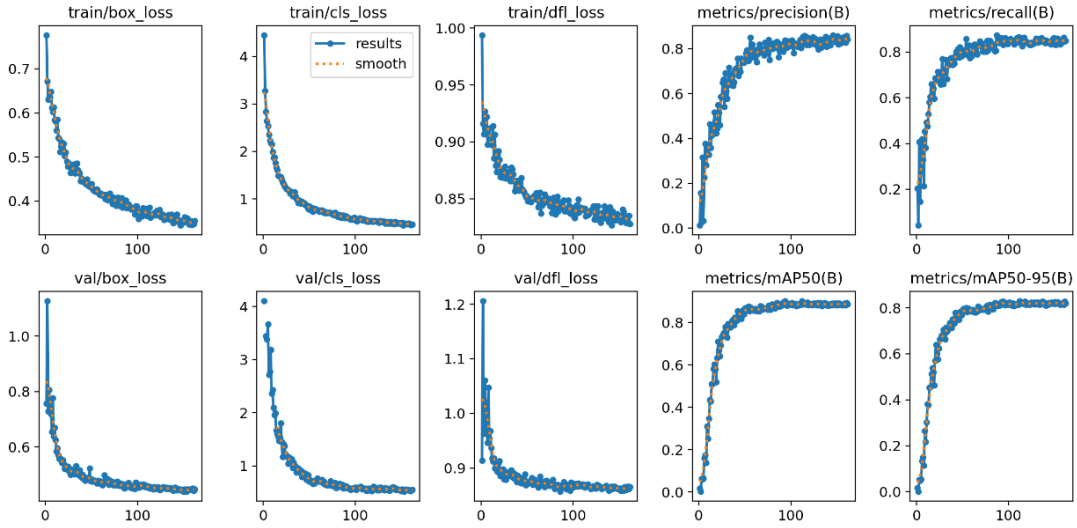
Sementara itu, pada model semi-supervised yang di inisiasi dengan Pseudo model, performa Pseudo YOLOv11 relatif lebih baik dibandingkan Pseudo YOLOv8, dengan precision 0,835, recall 0,806, serta mAP50 dan mAP50–95 sebesar 0,87 dan 0,803. Hasil visual pada Gambar 4.30(f) menunjukkan bahwa Pseudo YOLOv11 mampu mendeteksi 12 objek ground truth, dengan hanya satu false negative, sedangkan Pseudo YOLOv8 pada Gambar 4.30(e) mengalami penurunan performa dengan hanya 10 objek terdeteksi, yang tercermin dari nilai mAP50–95 terendah (0,765).

Sebagai pembanding, model fully supervised YOLOv8 dan YOLOv11 pada Gambar 4.30(g) dan 4.30(h) mampu mendeteksi seluruh 13 objek ground truth, namun tetap menghasilkan false positive pada objek yang tidak dilabeli akibat keterbatasan visibilitas. Secara keseluruhan, hasil ini menunjukkan bahwa pada skenario 20% data berlabel memberikan performa yang cukup baik dengan menghemat effort pelabelan sebanyak 80%.



Gambar 4. 31 Student Model 40% dengan Pretrained Model Yolov11

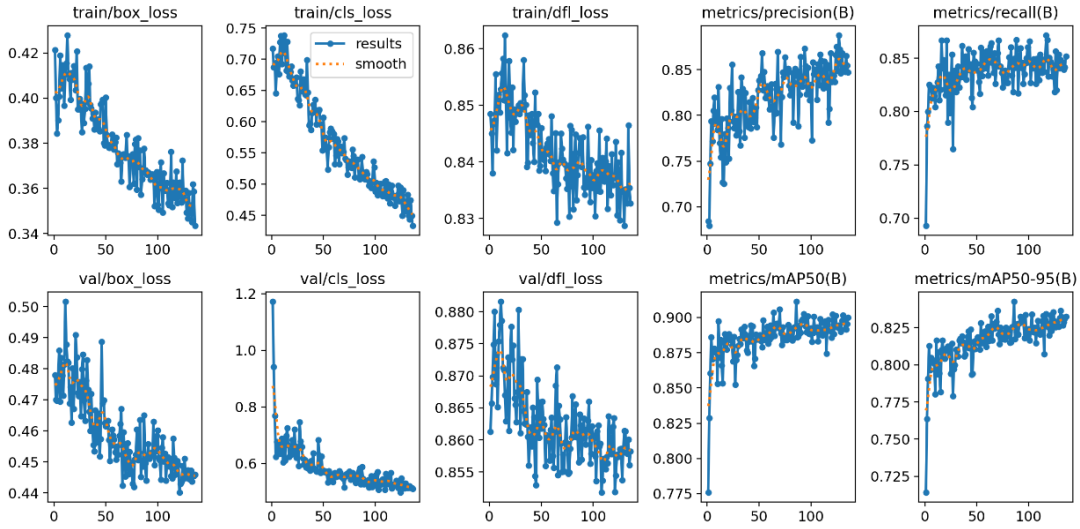
Grafik training Student Model 40% yang ditrain menggunakan Pretrained YOLOv11 menunjukkan pola yang cukup mulus sepanjang pelatihan. Metrik loss menurun secara konsisten, sedangkan precision, recall, mAP@50, dan mAP@50-95 meningkat secara stabil. Proses pelatihan mulai melambat sekitar epoch ke-80, menandakan bahwa model mulai mendekati konvergensi. Pola grafik yang mulus ini menunjukkan bahwa penambahan proporsi data berlabel 40% membantu model belajar lebih efisien dan stabil dibandingkan eksperimen dengan 20% data berlabel.



Gambar 4. 32 Student Model 40% dengan Pretrained Model Yolov8

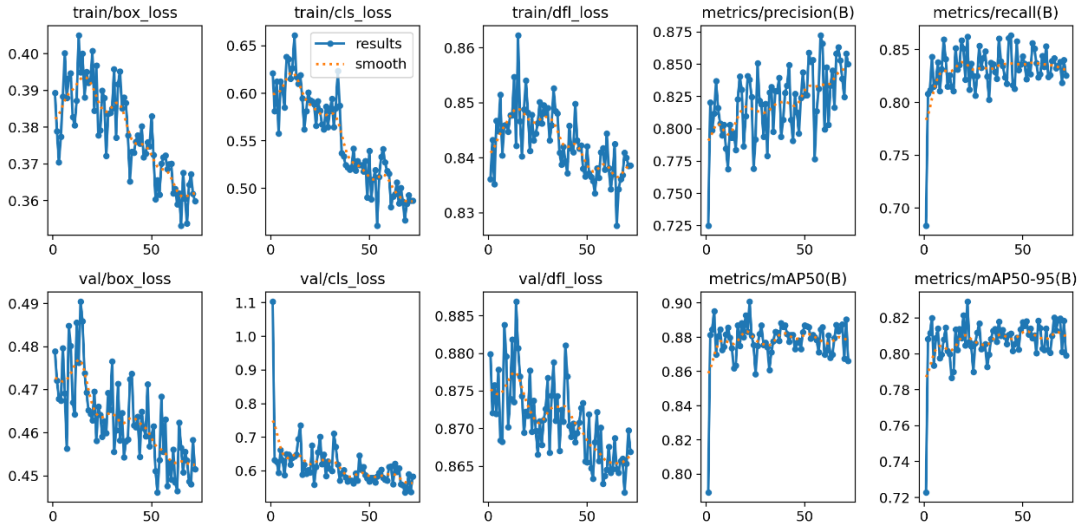
Grafik training Student Model 40% yang ditrain menggunakan Pretrained YOLOv8 menunjukkan penurunan loss yang mulus dan peningkatan metrik precision, recall, mAP@50, serta mAP@50-95 yang stabil sepanjang pelatihan. Proses pelatihan mulai melambat pada

sekitar epoch ke-60 hingga 70, menandakan model mulai mendekati konvergensi. Pola grafik yang mulus ini mengindikasikan bahwa model mampu belajar dengan efisien dari proporsi data berlabel 40%, menghasilkan pelatihan yang stabil dan performa yang konsisten.



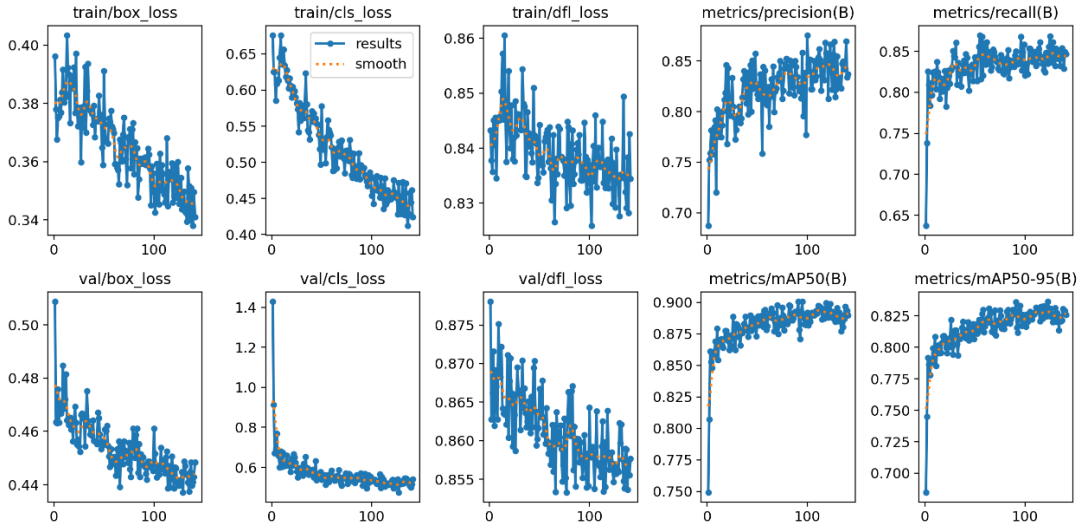
Gambar 4. 33 Student Model 40% dengan Teacher Model Yolov11

Grafik training Student Model 40% yang ditrain menggunakan Teacher Model YOLOv11 menunjukkan bahwa metrik precision dan recall langsung mencapai nilai sekitar 0,8 pada awal pelatihan, kemudian mengalami fluktuasi cukup parah namun secara umum tetap menunjukkan tren meningkat. DFL loss juga mengalami fluktuasi yang signifikan, menunjukkan ketidakstabilan pada beberapa update awal, sedangkan metrik mAP@50 dan mAP@50-95 juga mengalami fluktuasi tetapi tidak separah precision, recall, dan DFL loss. Pola ini mengindikasikan bahwa walaupun terdapat ketidakstabilan, model tetap belajar dan mengalami peningkatan performa secara keseluruhan.



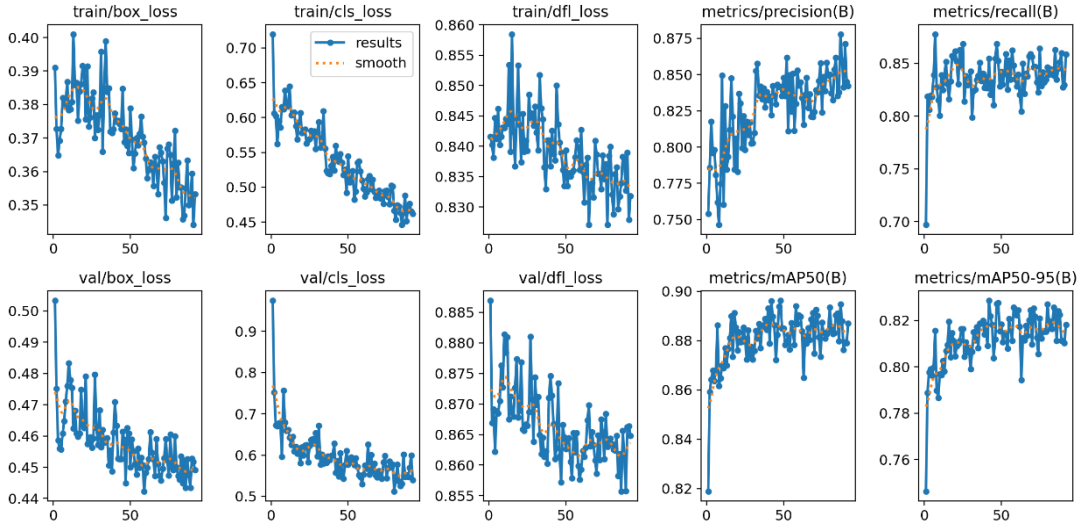
Gambar 4. 34 Student Model 40% dengan Teacher Model Yolov8

Grafik training Student Model 40% yang ditrain menggunakan Teacher Model YOLOv8 menunjukkan pola serupa dengan Teacher Model YOLOv11, di mana metrik precision dan recall langsung mencapai nilai tinggi di awal pelatihan dan kemudian mengalami fluktuasi cukup parah, namun tetap menunjukkan tren meningkat secara keseluruhan. DFL loss juga mengalami fluktuasi yang signifikan, sementara mAP@50 dan mAP@50-95 fluktuatif tetapi lebih stabil dibanding precision dan recall. Pelatihan terhenti pada epoch 72 karena tidak terjadi peningkatan selama 50 epoch berturut-turut, menandakan model telah mencapai titik konvergensi meskipun masih terdapat fluktuasi pada beberapa metrik.



Gambar 4. 35 Student Model 40% dengan Pseudo Model Yolov11

Grafik training Student Model 40% yang ditrain menggunakan Pseudo Model YOLOv11 menunjukkan peningkatan metrik yang relatif stabil dibandingkan dengan Teacher Model, meskipun tetap terdapat fluktuasi pada precision, recall, dan DFL loss. Fluktuasi tersebut tidak separah pada eksperimen sebelumnya, dan tren metrik tetap meningkat secara keseluruhan sepanjang pelatihan. Hal ini menunjukkan bahwa model mampu belajar dengan baik dari pseudo-label, meskipun jumlah data masih terbatas pada proporsi 40%.



Gambar 4. 36 Student Model 40% dengan Pseudo Model Yolov8

Grafik training Student Model 40% yang ditrain menggunakan Pseudo Model YOLOv8 menunjukkan lonjakan awal pada metrik, kemudian meningkat dengan fluktuasi yang sedang.

Meskipun DFL loss terus menurun, terdapat fluktuasi yang cukup signifikan sepanjang pelatihan. Secara keseluruhan, tren metrik tetap meningkat, menunjukkan bahwa model mampu belajar dari pseudo-label meskipun terjadi ketidakstabilan pada beberapa metrik.

Tabel 4. 8 Hasil Training Student model 40%

Eksperimen 40%										
Model	Epocs	Box Loss	Cls Loss	DFL Loss	Precision	Recall	mAP50	mAP50-95	Inference Time	Waktu Training (Jam)
Pre Train Yolov8	162	0.355	0.466	0.827	0.823	0.862	0.9	0.83	4.6ms	0.425
Pre Train Yolov11	248	0.342	0.402	0.833	0.857	0.857	0.908	0.843	4.7ms	0.698
Teacher Model Yolov8	72	0.359	0.487	0.838	0.838	0.862	0.901	0.829	4.6ms	0.187
Teacher Model Yolov11	136	0.343	0.433	0.823	0.843	0.85	0.912	0.843	4.6ms	0.383
Pseudo Model Yolov8	92	0.353	0.461	0.831	0.822	0.855	0.896	0.829	5.0ms	0.240
Pseudo Model Yolov11	141	0.341	0.424	0.834	0.866	0.835	0.899	0.837	4.7ms	0.399



(a) Pretrained Model Yolov8



(b) Pretrained Model Yolov11



(c) Teacher Model Yolov8



(d) Teacher Model Yolov11



(e) Pseudo Model Yolov8



(f) Pseudo Model Yolov11



(g) Fully Supervised Yolov8



(h) Fully Supervised Yolov11



(i) Ground Truth

Gambar 4. 37 Visual Test Semi Supervised 40%

Tabel 4. 9 Hasil Visual Semi Supervised 40%

Lokasi	Ground Truth	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
A1	Lactogen 1 Happynutri 24x350g	TP	TP	TP	FN - noBB	TP	TP	TP	TP
A2	Lactogen 2 Happynutri 24x350g	TP	TP	TP	TP	TP	TP	TP	TP
A3	Lactogrow 3 Happynutri 24x350g	TP	TP	TP	FN - noBB	TP	TP	TP	TP
A4	Lactogen 1 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
A5	Lactogen 2 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
A6	Lactogrow 3 Happynutri Honey 12x750g	TP	TP	TP	FN - noBB	TP	FN - noBB	TP	TP
B2	Lactogrow 3 Happynutri Honey 24x350g	TP	TP	TP	FN - noBB	TP	TP	TP	TP
B3	Lactogrow 3 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP

B4	Lactogrow 3 Happynutri Van 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
B5	Lactogrow 4 Happynutri Honey 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
B6	Lactogrow 4 Happynutri Van 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
C1	Lactogen 1 Happynutri 40x180g	TP	TP	TP	TP	TP	TP	TP	TP
C2	Lactogen 2 Happynutri 40x180g	FN - noBB	FN - noBB	TP	FN - noBB	TP	TP	TP	TP

Lokasi objek dinyatakan berdasarkan posisi rak dan urutan dari kiri ke kanan, dengan A menunjukkan rak atas, B rak tengah, dan C rak bawah. Angka menunjukkan urutan objek pada rak (misalnya B5 adalah objek ke-5 dari kiri pada rak tengah). Keterangan noBB menunjukkan tidak terdapat bounding box pada objek tersebut, sedangkan FP – … menunjukkan objek yang terdeteksi secara keliru oleh model.

Pelatihan student model dengan proporsi 40% data berlabel menunjukkan peningkatan performa secara umum dibandingkan eksperimen sebelumnya, tercermin dari nilai mAP50 seluruh model yang kini berada pada kisaran 0.896-0.912. Analisis detail per objek disajikan pada Tabel 4. 9, dengan perbandingan visual pada Gambar 4.37.

Berdasarkan Tabel 4. 9, performa deteksi visual pada eksperimen 40% sebagai berikut:

a. Model (a) Pretrained YOLOv8 berhasil mendeteksi 12 objek pada ground truth, dengan satu false negative pada objek Lactogen 2 Happynutri 40x180g (C2). Selain itu, terdapat satu false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), yang tidak diberi label dan bounding box pada ground truth, tetapi dapat terdeteksi dengan kelas yang benar oleh model.
b. Model (b) Pretrained YOLOv11 mendeteksi 12 objek pada ground truth, dengan satu false negative pada Lactogen 2 Happynutri 40x180g (C2). Terdapat dua false positive, yaitu pada objek Lactogrow 3 Happynutri Honey 40x180g (B2), Dancow 1+ Van Advn ExcNutr 24x400g (C5), dan Dancow 3+ (C6) yang tidak diberi label

dan bounding box pada ground truth, tetapi dapat terdeteksi dengan kelas yang benar oleh model, objek tersebut tidak dilabeli ataupun tidak diberi bounding box karena terdapat informasi yang tidak cukup jelas untuk proses anotasi seperti tulisan gramasinya. Meskipun secara visual deteksi model tampak benar, objek-objek tersebut tetap dikategorikan sebagai false positive karena tidak tercantum dalam anotasi ground truth yang digunakan sebagai acuan evaluasi.

c. Model (c) Teacher YOLOv8 berhasil mendeteksi seluruh 13 objek pada ground truth. Namun, masih terdapat tiga false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 1+ (C5), dan Dancow 3+ (C6) yang tidak diberi label maupun bounding box pada ground truth, meskipun terdeteksi oleh model dengan kelas yang sesuai. Hal ini disebabkan oleh keterbatasan informasi visual pada objek, seperti tulisan gramasi yang tidak terlihat secara jelas, sehingga objek tidak memenuhi kriteria anotasi. Oleh karena itu, deteksi tersebut tetap dikategorikan sebagai false positive karena tidak tercantum dalam ground truth yang digunakan sebagai acuan evaluasi.
d. Model (d) Teacher YOLOv11 menunjukkan penurunan performa visual dengan hanya 8 objek berhasil terdeteksi. Terdapat lima false negative, yaitu pada Lactogen 1 Happynutri 24x350g (A1), Lactogrow 3 Happynutri Van 24x350g (A3), Lactogrow 3 Happynutri Honey 12x750g (A6), Lactogrow 3 Happynutri Honey 24x350g (B2), dan Lactogen 2 Happynutri 40x180g (C2). Selain itu, terdapat dua false positive pada objek Dancow 1+ (C5) dan Dancow 3+ (C6), yang tidak diberi label dan bounding box pada ground truth, tetapi dapat dideteksi dengan kelas dan posisi yang benar oleh model. Objek tersebut tidak dilabeli karena keterbatasan informasi visual pada objek, seperti tulisan gramasi yang tidak terlihat secara jelas. Oleh karena itu, deteksi tersebut tetap dikategorikan sebagai false positve karena tidak tercantum dalam ground truth yang digunakan sebagai acuan evaluasi.
e. Model (e) Pseudo YOLOv8 berhasil mendeteksi seluruh 13 objek pada ground truth. Meskipun demikian, terdapat tiga false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 1+ (C5), dan Dancow 3+ (C6). Meskipun model mampu mendeteksi objek dengan kelas yang benar, objek tersebut tidak dilabeli dan tidak memiliki bounding box pada ground truth akibat informasi visual yang tidak cukup jelas, seperti tulisan gramasi yang tidak terbaca. Dengan demikian, deteksi

tersebut tetap dihitung sebagai false positive berdasarkan acuan evaluasi ground truth.

f. Model (f) Pseudo YOLOv11 mendeteksi 12 dari 13 objek pada ground truth, dengan satu false negative pada Lactogrow 3 Honey 12x750g (A6). Selain itu, terdapat dua false positive pada objek Dancow 1+ (C5) dan Dancow 3+ (C6), deteksi pada kedua objek yang dihasilkan model menunjukkan kesesuaian kelas secara visual. Namun, objek tersebut tidak dianotasi pada ground truth karena informasi visual yang tidak memadai untuk proses anotasi, seperti tulisan gramasi yang tidak terlihat secara utuh. Oleh karena itu, deteksi ini dikategorikan sebagai false positive sesuai dengan skema evaluasi berbasis ground truth.

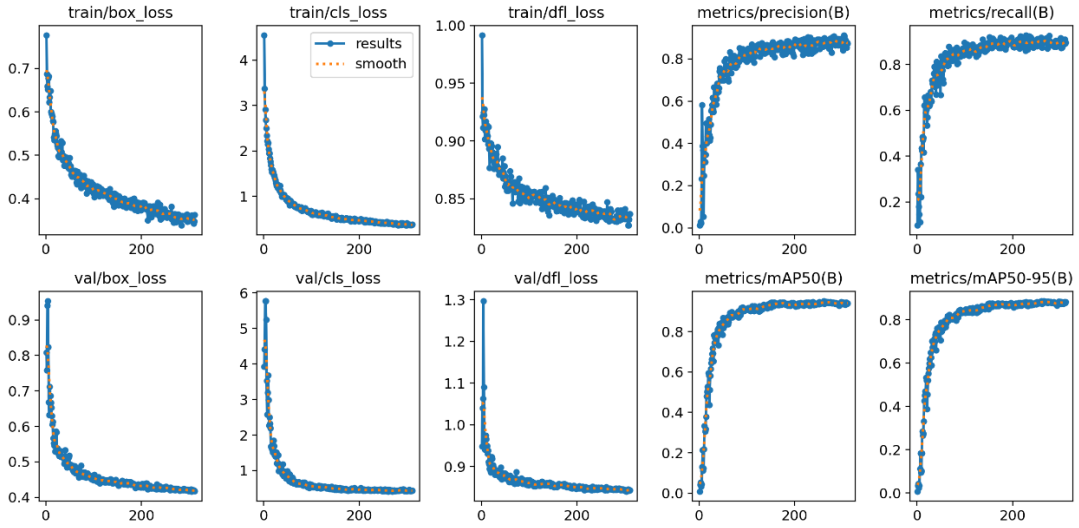
Pada eksperimen 40% data berlabel, peningkatan proporsi data beranotasi memberikan dampak positif terhadap performa model, baik secara kuantitatif maupun visual sebagaimana ditunjukkan pada Gambar 4.37(a–h). Model Pretrained YOLOv8 dan Pretrained YOLOv11 masing-masing mampu mendeteksi 12 dari 13 objek ground truth, dengan satu false negative pada produk Lactogen 2 Happynutri 40x180g. Nilai precision dan recall yang relatif tinggi (YOLOv8: 0,823 dan 0,862; YOLOv11: 0,857 dan 0,857) mencerminkan kemampuan model dalam menyeimbangkan ketepatan dan kelengkapan deteksi, meskipun masih ditemukan false positive pada objek yang secara visual terdeteksi dengan benar namun tidak dilabeli pada ground truth karena tidak terlihat sepenuhnya.

Pada model semi-supervised yang di inisiasi dengan teacher model, Teacher YOLOv8 menunjukkan performa visual yang sangat baik dengan berhasil mendeteksi seluruh 13 objek ground truth, sejalan dengan nilai recall 0,862 dan mAP50 sebesar 0,901. False positive yang muncul didominasi oleh produk Lactogrow 3 Happynutri Honey 40x180g, Dancow 1+, dan Dancow 3+, yang tidak tercantum pada ground truth akibat keterbatasan visibilitas objek. Sebaliknya, Teacher YOLOv11 mengalami penurunan performa visual dengan hanya 8 objek terdeteksi, yang ditandai oleh 5 false negative, meskipun secara kuantitatif masih mempertahankan nilai mAP50 tertinggi sebesar 0,912. Hal ini menunjukkan bahwa metrik agregat dapat tetap tinggi meskipun terjadi kegagalan deteksi pada kasus tertentu, terutama pada objek kecil atau tertutup sebagian.

Pada model semi-supervised dengan pseudo model, baik Pseudo YOLOv8 maupun Pseudo YOLOv11 menunjukkan performa yang stabil. Pseudo YOLOv8 mampu mendeteksi seluruh 13 objek ground truth, sedangkan Pseudo YOLOv11 mendeteksi 12 dari 13 objek dengan satu

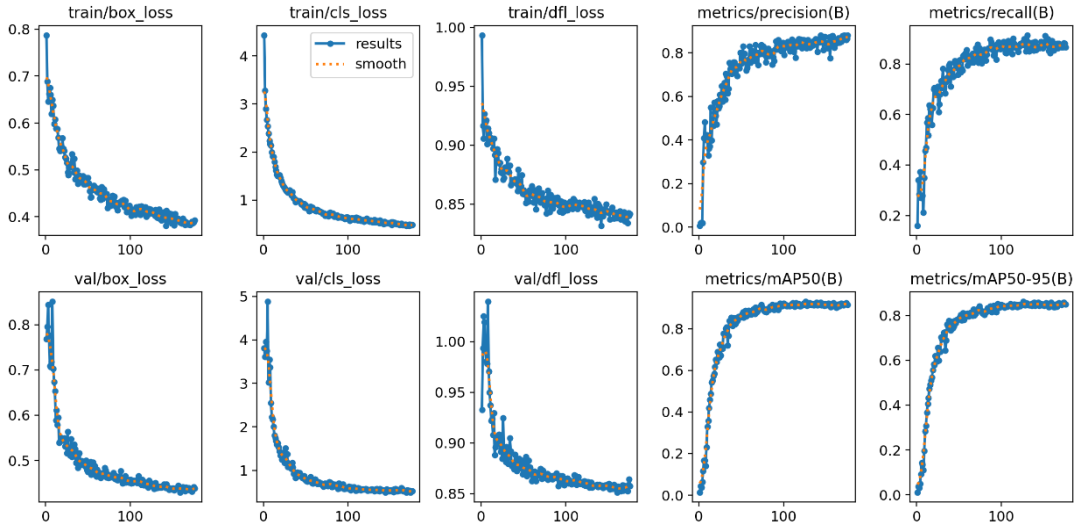
false negative pada Lactogrow 3 Honey 12x750g. Nilai precision dan recall pada kedua model berada di atas 0,82, dengan Pseudo YOLOv11 mencatat precision tertinggi sebesar 0,866.

Sebagai pembanding, model fully supervised YOLOv8 dan YOLOv11 pada Gambar 4.37(g) dan 4.37(h) mampu mendeteksi seluruh 13 objek ground truth, namun tetap menghasilkan false positive pada beberapa produk yang tidak dilabeli akibat objek yang tidak terlihat sepenuhnya. Secara keseluruhan, hasil eksperimen ini menunjukkan bahwa pada skenario 40% data berlabel, pendekatan semi-supervised khususnya pada Teacher YOLOv8 dan Pseudo YOLOv11 mampu mendekati performa fully supervised, dengan kesalahan deteksi yang sebagian besar disebabkan oleh keterbatasan anotasi ground truth, bukan kegagalan model dalam mengenali objek.



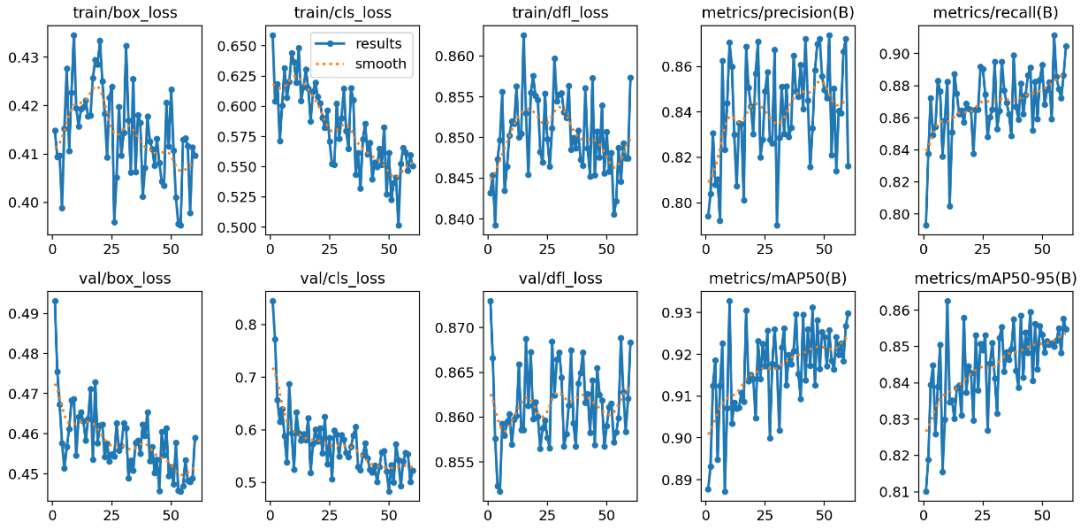
Gambar 4. 38 Student Model 60% dengan Pretrained Model Yolov11

Grafik training Student Model 60% yang ditrain menggunakan Pretrained Model YOLOv11 terlihat mulus dengan peningkatan metrik yang konsisten. Terlihat bahwa laju peningkatan mulai melambat sekitar epoch 100, menandakan model mendekati konvergensi. Secara keseluruhan, tren metrik seperti precision, recall, mAP50, dan mAP50-95 menunjukkan performa yang stabil dan peningkatan yang terus berlangsung hingga mendekati akhir pelatihan.



Gambar 4. 39 Student Model 60% dengan Pretrained Model Yolov8

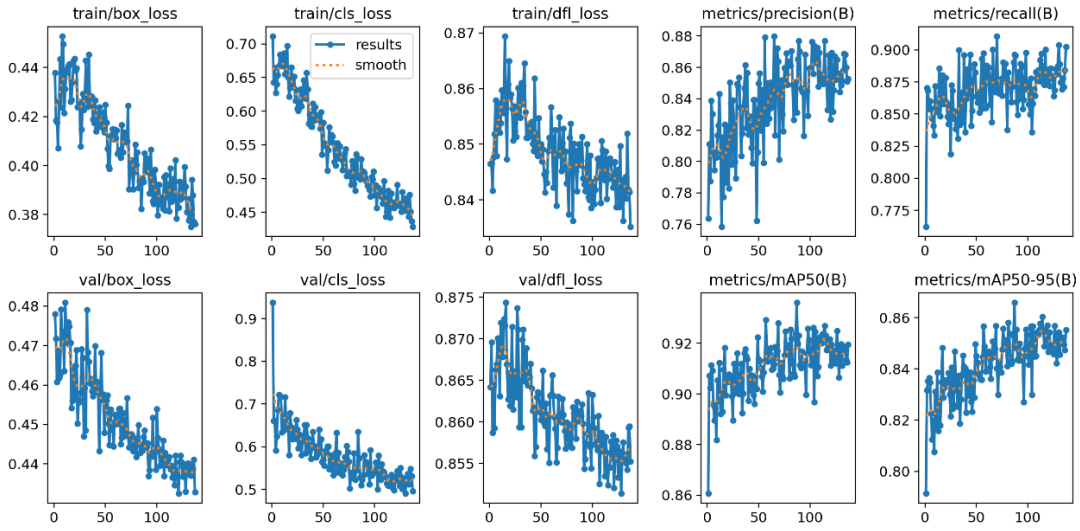
Grafik training Student Model 60% yang ditrain menggunakan Pretrained Model YOLOv8 menunjukkan tren yang mulus dengan peningkatan metrik yang konsisten di awal pelatihan. Laju peningkatan mulai melambat sekitar epoch 60, di mana hanya terjadi sedikit kenaikan metrik selanjutnya, menandakan model mulai mendekati titik konvergensi. Secara keseluruhan, metrik seperti precision, recall, mAP50, dan mAP50-95 tetap menunjukkan performa yang stabil hingga akhir pelatihan.



Gambar 4. 40 Student Model 60% dengan Teacher Model Yolov11

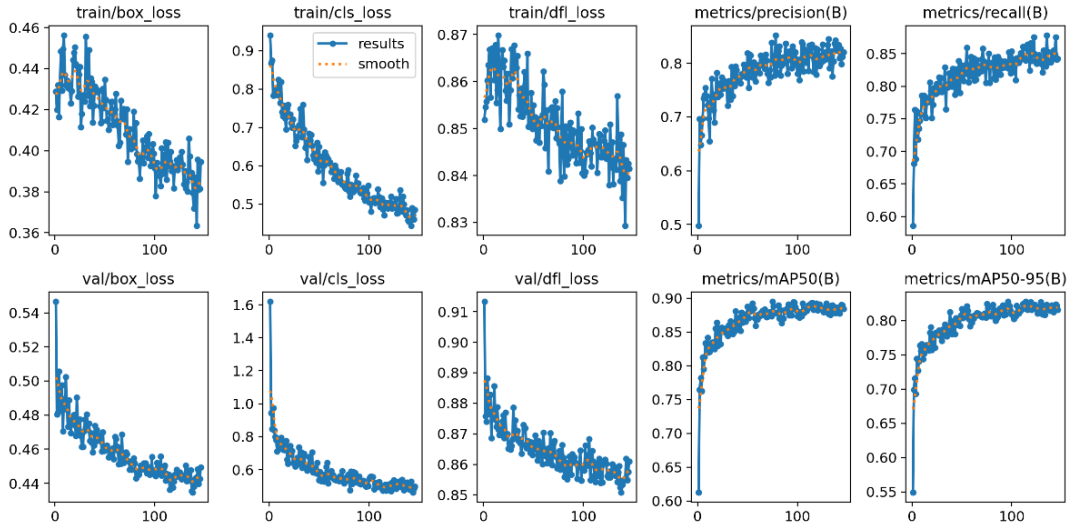
Grafik training Student Model 60% yang ditrain menggunakan Teacher Model YOLOv11 menunjukkan bahwa metrik awal langsung berada pada nilai tinggi, mencerminkan performa awal yang baik. Setelah itu, terjadi fluktuasi dengan pola naik-turun sekitar rentang 0,06 pada

metrik seperti precision, recall, dan mAP, menunjukkan ketidakstabilan sementara dalam proses konvergensi. Pelatihan berhenti pada epoch 60 karena tidak terjadi peningkatan selama 50 epoch berturut-turut, menandakan model telah mencapai kondisi stagnan.



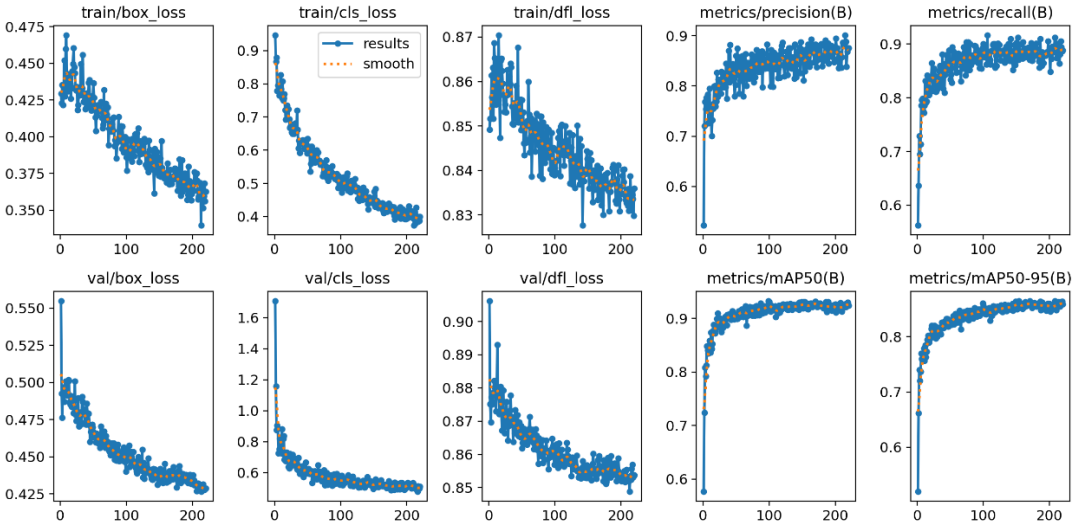
Gambar 4. 41 Student Model 60% dengan Teacher Model Yolov8

Grafik training Student Model 60% yang ditrain menggunakan Teacher Model YOLOv8 menunjukkan kenaikan metrik yang cepat di awal, mencerminkan performa awal yang baik. Setelah itu, metrik mengalami fluktuasi cukup parah, menunjukkan ketidakstabilan selama proses training. Untuk DFL loss, terdapat kenaikan di awal epoch sebelum akhirnya menurun, namun tetap disertai fluktuasi yang signifikan sepanjang pelatihan. Pola ini menggambarkan bahwa meskipun model dapat belajar dengan cepat, konsistensi pelatihan masih terpengaruh oleh variasi data.



Gambar 4. 42 Student Model 60% dengan Pseudo Model Yolov11

Grafik training Student Model 60% yang ditrain menggunakan Pseudo Model YOLOv11 menunjukkan pola yang relatif mulus dengan hanya sedikit fluktuasi pada metrik precision, recall, dan mAP. Fluktuasi yang cukup terlihat terjadi pada DFL loss, meskipun polanya tetap menunjukkan penurunan secara umum. Hal ini mengindikasikan bahwa model belajar dengan stabil dari pseudo-label, dengan performa yang meningkat secara konsisten meskipun ada variasi pada loss tertentu.



Gambar 4. 43 Student Model 60% dengan Pseudo Model Yolov8

Grafik training Student Model 60% yang ditrain menggunakan Pseudo Model YOLOv8 menampilkan pola yang mirip dengan versi YOLOv11, namun sedikit lebih mulus. Metrik precision, recall, dan mAP menunjukkan peningkatan yang konsisten, dengan fluktuasi

minimal pada awal epoch. Setelah epoch 60, grafik cenderung stabil hingga akhir, menandakan bahwa model berhasil belajar dari pseudo-label dengan stabil dan konvergen.

Tabel 4. 10 Hasil Training Student model 60%

Model	Epocs	Box Loss	Cls Loss	DFL Loss	Precision	Recall	mAP50	mAP50-95	Inference Time	Waktu Training (Jam)
Pre Trained Yolov8	176	0.392	0.480	0.841	0.848	0.876	0.931	0.863	4.8ms	0.477
Pre Trained Yolov11	311	0.363	0.379	0.836	0.866	0.904	0.951	0.888	4.8ms	0.910
Teacher Model Yolov8	137	0.376	0.428	0.835	0.831	0.897	0.936	0.865	5.6ms	0.370
Teacher Model Yolov11	60	0.409	0.550	0.857	0.871	0.882	0.933	0.862	4.7ms	0.176
Pseudo Model Yolov8	220	0.362	0.399	0.836	0.883	0.87	0.931	0.865	5.5ms	0.598
Pseudo Model Yolov11	146	0.394	0.485	0.841	0.823	0.828	0.895	0.829	5.7ms	0.426



(a) Pretrained Model Yolov8



(b) Pretrained Model Yolov11



(c) Teacher Model Yolov8



(d) Teacher Model Yolov11



(e) Pseudo Model Yolov8



(f) Pseudo Model Yolov11



(g) Fully Supervised Yolov8



(h) Fully Supervised Yolov11



(i) Ground Truth

Gambar 4. 44 Visual Test Semi Supervised 60%

Tabel 4. 11 Hasil Visual Semi Supervised 60%

Lokasi	Ground Truth	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
A1	Lactogen 1 Happynutri 24x350g	TP	TP	TP	TP	TP	TP	TP	TP
A2	Lactogen 2 Happynutri 24x350g	TP	TP	TP	TP	TP	TP	TP	TP
A3	Lactogrow 3 Happynutri 24x350g	TP	TP	TP	TP	TP	TP	TP	TP
A4	Lactogen 1 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
A5	Lactogen 2 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
A6	Lactogrow 3 Happynutri Honey 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
B2	Lactogrow 3 Happynutri Honey 24x350g	TP	FP – Lactogrow 3 Happynutri Honey 40x180g	TP	TP	TP	TP	TP	TP
B3	Lactogrow 3 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP

B4	Lactogrow 3 Happynutri Van 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
B5	Lactogrow 4 Happynutri Honey 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
B6	Lactogrow 4 Happynutri Van 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
C1	Lactogen 1 Happynutri 40x180g	TP	TP	TP	TP	TP	TP	TP	TP
C2	Lactogen 2 Happynutri 40x180g	TP	TP	TP	TP	TP	TP	TP	TP

Lokasi objek dinyatakan berdasarkan posisi rak dan urutan dari kiri ke kanan, dengan A menunjukkan rak atas, B rak tengah, dan C rak bawah. Angka menunjukkan urutan objek pada rak (misalnya B5 adalah objek ke-5 dari kiri pada rak tengah). Keterangan noBB menunjukkan tidak terdapat bounding box pada objek tersebut, sedangkan FP – … menunjukkan objek yang terdeteksi secara keliru oleh model.

Pelatihan student model dengan menggunakan 60% data berlabel menunjukkan peningkatan performa yang lebih signifikan dibanding eksperimen sebelumnya. Seluruh model memperoleh nilai mAP50 di atas 0.895, dengan dua model Pretrained YOLOv11 dan Teacher YOLOv8 mencapai 0.951 dan 0.936. Analisis detail per objek disajikan pada Tabel 4.11 dengan visualisasi pada Gambar 4.44.

Berdasarkan Tabel 4.11, performa deteksi menunjukkan peningkatan:

a. Model (a) Pretrained YOLOv8 menunjukkan kinerja deteksi yang baik dengan berhasil mendeteksi seluruh 13 objek pada ground truth. Meskipun demikian, terdapat tiga false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 1+ (C5), dan Dancow 3+ (C6), yang tidak diberi label dan bounding box pada ground truth, tetapi dapat terdeteksi dengan kelas dan posisi yang benar oleh model. Ketiga objek tersebut tidak dilabeli maupun tidak diberi bounding box karena terdapat informasi yang tidak cukup jelas untuk proses anotasi seperti varian gramasi yang tidak terlalu terlihat. Meskipun secara visual deteksi model tampak benar,

objek-objek tersebut tetap dikategorikan sebagai false positive karena tidak tercantum dalam anotasi di ground truth yang digunakan sebagai acuan evaluasi model.

b. Model (b) Pretrained YOLOv11 berhasil mendeteksi 12 objek pada ground truth. Namun, terdapat lima false positive, yang meliputi kesalahan klasifikasi pada Lactogrow 3 Honey 40x180g yang seharusnya terdeteksi sebagai Lactogrow 3 Happynutri Honey 24x350g (B2), serta false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 1+ (C5), dan dua objek SGM yang berada di tepi gambar sebelah kanan. Objek-objek tersebut tidak diberi label maupun bounding box pada ground truth, meskipun terdeteksi oleh model dengan kelas yang sesuai. Hal ini disebabkan oleh keterbatasan informasi visual pada objek, seperti tulisan gramasi yang tidak terlihat secara jelas dan objek yang hanya terlihat sebagian seperti pada produk SGM, sehingga objek tidak memenuhi kriteria anotasi. Oleh karena itu, deteksi tersebut tetap dikategorikan sebagai false positive karena tidak tercantum dalam ground truth yang digunakan sebagai acuan evaluasi.
c. Model (c) Teacher YOLOv8 mencapai deteksi sempurna dengan berhasil mendeteksi seluruh 13 objek pada ground truth. Namun, masih terdapat tiga false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 1+ (C5), dan SGM. Meskipun model mampu mendeteksi objek dengan kelas yang benar, objek tersebut tidak dilabeli dan tidak memiliki bounding box pada ground truth akibat informasi visual yang tidak cukup jelas, seperti tulisan gramasi yang tidak terbaca. Dengan demikian, deteksi tersebut tetap dihitung sebagai false positive berdasarkan acuan evaluasi ground truth.
d. Model (d) Teacher YOLOv11 juga mencapai deteksi sempurna dengan mendeteksi 13 objek pada ground truth. Terdapat tiga false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 1+ (C5), dan SGM, yang tidak dilabeli dan tidak diberi bounding box karena terdapat keterbatasan informasi seperti tulisan gramasi yang tidak terlihat dengan jelas dan objek yang hanya terlihat sebagian seperti pada objek SGM tetapi dapat dideteksi dengan kelas dan posisi benar oleh model. Oleh karena itu, objek-objek tersebut tetap dikategorikan sebagai false positive berdasarkan acuan evaluasi yang digunakan.
e. Model (e) Pseudo YOLOv8 berhasil mendeteksi seluruh 13 objek pada ground truth. Meskipun demikian, terdapat empat false positive pada objek Lactogrow 3

Happynutri Honey 40x180g (B1), Dancow 1+ (C5), Dancow 3+ (C6), dan SGM, objek-objek tersebut tidak dilabeli dan tidak memiliki bounding box pada ground truth akibat keterbatasan informasi visual, seperti tulisan gramasi yang tidak terbaca dan objek yang hanya terlihat sebagian (misalnya pada objek SGM). Walaupun model mendeteksi objek dengan kelas dan posisi yang benar, deteksi tersebut tetap dikategorikan sebagai false positive sesuai acuan evaluasi.

f. Model (f) Pseudo YOLOv11 juga berhasil mendeteksi seluruh 13 objek pada ground truth. Namun, masih terdapat tiga false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 3+ (C6), dan SGM, yang tidak dilabeli dan tidak memiliki bounding box pada ground truth akibat keterbatasan informasi visual, seperti tulisan gramasi yang tidak terbaca dan objek yang hanya terlihat sebagian (misalnya pada objek SGM). Walaupun model mendeteksi objek dengan kelas dan posisi yang benar, deteksi tersebut tetap dikategorikan sebagai false positive sesuai acuan evaluasi.

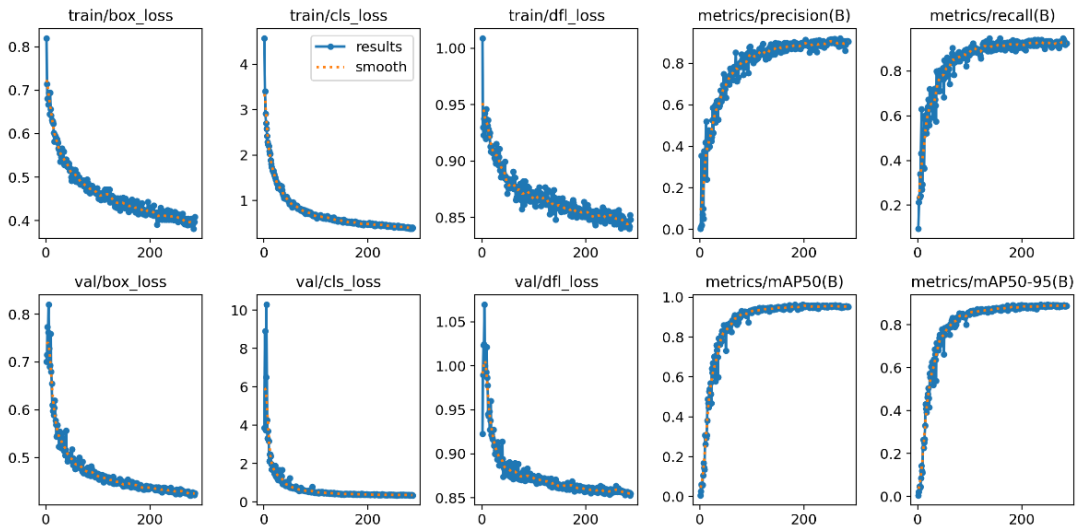
Pada eksperimen dengan proporsi data berlabel yang lebih besar, seluruh model menunjukkan peningkatan performa yang konsisten baik secara kuantitatif maupun kualitatif, sebagaimana ditunjukkan pada Gambar 4.44(a–h). Model Pretrained YOLOv8 dan Pretrained YOLOv11 mampu mencapai tingkat deteksi yang sangat baik, dengan YOLOv8 berhasil mendeteksi seluruh 13 objek ground truth dan YOLOv11 mendeteksi 12 dari 13 objek, yang tercermin dari nilai recall tinggi masing-masing 0,876 dan 0,904. Meskipun demikian, masih ditemukan false positive pada beberapa produk yang secara visual terdeteksi dengan benar namun tidak dilabeli pada ground truth karena objek tidak terlihat sepenuhnya, serta kesalahan klasifikasi pada variasi kemasan produk yang memiliki kemiripan visual.

Pada model semi-supervised dengan teacher model, baik Teacher YOLOv8 maupun Teacher YOLOv11 mencapai deteksi sempurna 13/13 objek ground truth, yang sejalan dengan nilai mAP50 di atas 0,93 dan mAP50–95 di atas 0,86. False positive yang muncul pada kedua model didominasi oleh objek yang tidak dilabeli pada ground truth akibat keterbatasan visibilitas, bukan karena kegagalan model dalam mengenali kelas objek.

Pada model semi-supervised dengan pseudo model, Pseudo YOLOv8 dan Pseudo YOLOv11 juga menunjukkan kinerja yang kuat dengan seluruh 13 objek ground truth berhasil terdeteksi. Pseudo YOLOv8 mencatat precision tertinggi sebesar 0,883, menunjukkan ketepatan prediksi yang sangat baik, sedangkan Pseudo YOLOv11 mengalami sedikit penurunan performa dengan precision 0,823 dan recall 0,828, yang mengindikasikan bahwa

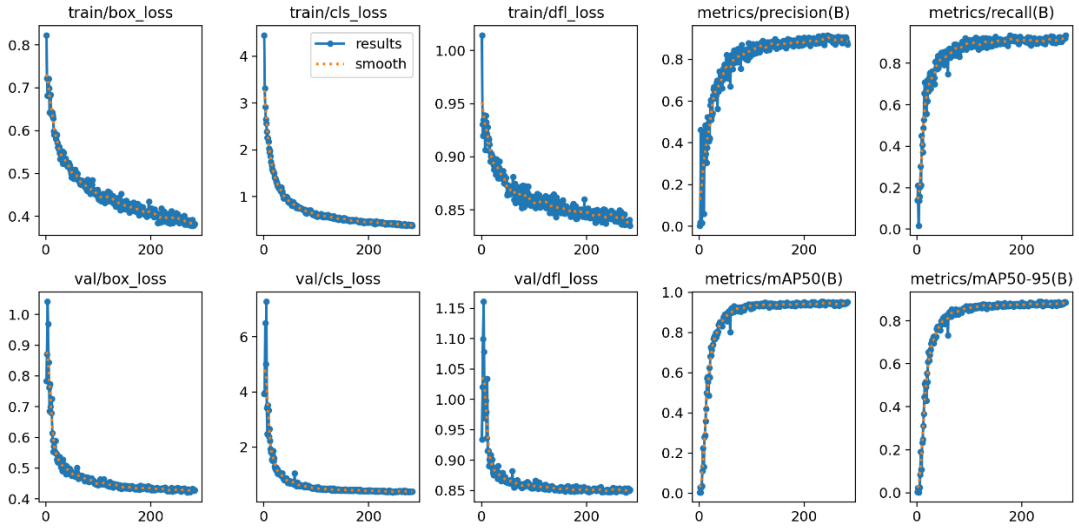
penambahan pseudo-label pada proporsi data berlabel yang lebih tinggi tidak selalu memberikan peningkatan signifikan pada semua arsitektur.

Sebagai pembanding, model fully supervised YOLOv8 dan YOLOv11 pada Gambar 4.44(g) dan 4.44(h) juga mampu mendeteksi seluruh 13 objek ground truth, namun tetap menghasilkan false positive pada beberapa produk yang tidak dilabeli akibat objek yang tidak terlihat sepenuhnya. Secara keseluruhan, hasil ini menunjukkan bahwa pada proporsi data berlabel yang lebih tinggi, pendekatan semi-supervised learning khususnya Teacher dan Pseudo YOLOv8 mampu menyamai bahkan mendekati performa fully supervised, dengan kesalahan deteksi yang sebagian besar dipengaruhi oleh keterbatasan anotasi ground truth, bukan oleh kelemahan model.



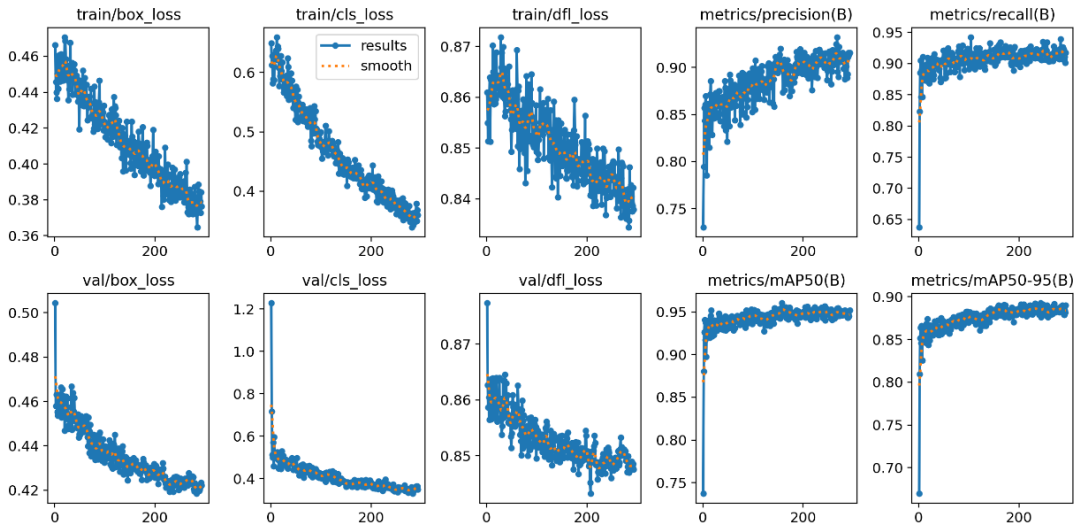
Gambar 4. 45 Student Model 80% dengan Pretrained Model Yolov11

Grafik training Student Model 80% yang menggunakan Pretrained Model YOLOv11 terlihat mulus sepanjang proses, dengan penurunan loss yang konsisten dan peningkatan metrik precision, recall, serta mAP yang stabil. Model mulai menunjukkan tanda-tanda konvergensi pada sekitar epoch 100, menandakan bahwa pembelajaran dari data 80% berlabel telah mencapai kestabilan dan model mampu memanfaatkan data berlabel secara efektif.



Gambar 4. 46 Student Model 80% dengan Pretrained Model Yolov8

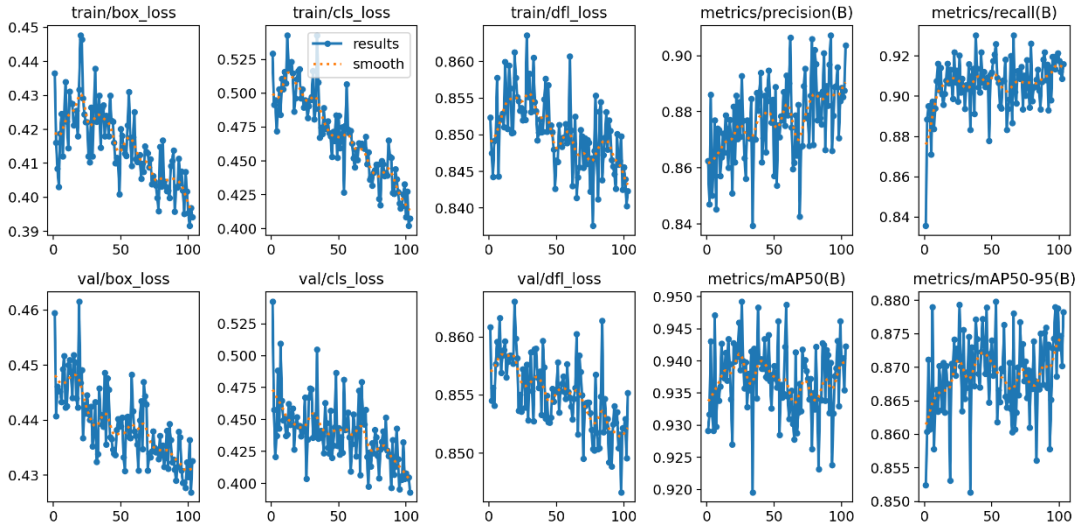
Grafik training Student Model 80% yang menggunakan Pretrained Model YOLOv8 menunjukkan pola yang mulus dan konsisten, dengan penurunan loss yang stabil serta peningkatan metrik precision, recall, dan mAP yang teratur. Model mulai menunjukkan tanda-tanda konvergensi pada sekitar epoch 80, menunjukkan bahwa pelatihan dengan 80% data berlabel cukup efektif untuk mencapai kestabilan performa model.



Gambar 4. 47 Student Model 80% dengan Teacher Model Yolov11

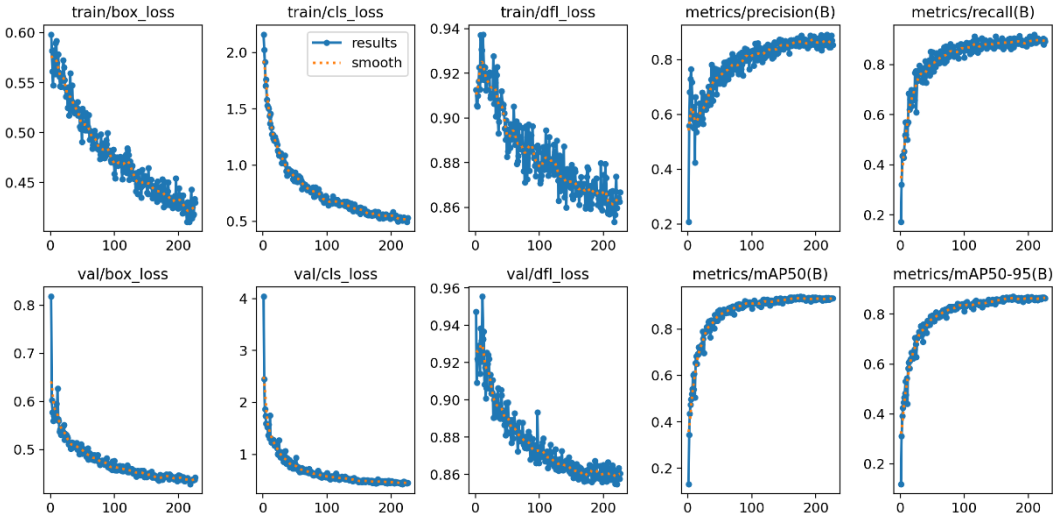
Grafik training Student Model 80% yang menggunakan Teacher Model YOLOv11 menunjukkan adanya fluktuasi pada metrik training, namun amplitudonya relatif kecil dibandingkan dengan eksperimen sebelumnya (20%, 40%, dan 60%). Meskipun masih terdapat naik-turun pada precision, recall, dan mAP, pola keseluruhan tetap menunjukkan tren

peningkatan, menandakan bahwa model mampu mempelajari representasi dari pseudo-label dengan lebih stabil ketika proporsi data berlabel yang digunakan lebih besar.



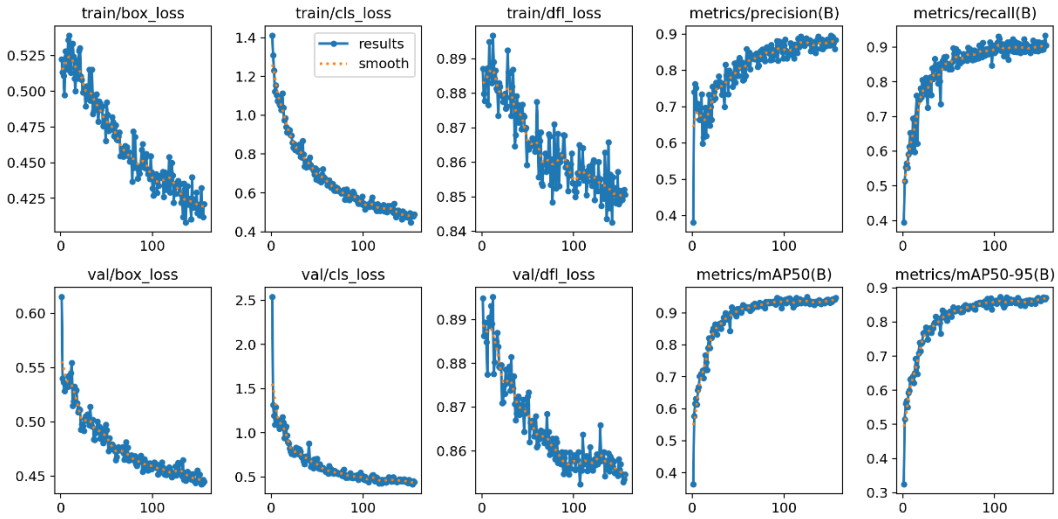
Gambar 4. 48 Student Model 80% dengan Teacher Model Yolov8

Grafik training Student Model 80% dengan Teacher Model YOLOv8 menunjukkan bahwa metrik precision, recall, dan mAP cepat mencapai nilai tinggi sekitar 0,90 sebelum epoch 30. Setelah itu, terlihat fluktuasi yang cukup parah pada hampir semua metrik, meskipun tren keseluruhan tetap meningkat secara perlahan. Hal ini menunjukkan bahwa meskipun model belajar dengan cepat pada awal training, beberapa variasi pada pseudo-label menyebabkan ketidakstabilan dalam pelatihan di sisa epoch.



Gambar 4. 49 Student Model 80% dengan Pseudo Model Yolov11

Grafik training Student Model 80% yang dilatih menggunakan Pseudo Model YOLOv11 menunjukkan pola yang cukup mulus, dengan fluktuasi hanya terjadi pada DFL loss. Precision, recall, dan mAP meningkat secara konsisten, dan grafik mulai menunjukkan konvergensi pada sekitar epoch 80. Hal ini menandakan bahwa model mampu memanfaatkan pseudo-label dengan baik untuk belajar, meskipun terdapat sedikit variasi pada perhitungan DFL loss.



Gambar 4. 50 Student Model 80% dengan Pseudo Model Yolov8

Grafik training Student Model 80% yang dilatih menggunakan Pseudo Model YOLOv8 relatif mulus, dengan konvergensi mulai terlihat pada epoch 60 - 70. Fluktuasi yang cukup parah hanya terjadi pada DFL loss, sedangkan metrik precision, recall, dan mAP meningkat secara konsisten, menunjukkan bahwa model mampu belajar dari pseudo-label dengan baik meskipun terdapat variasi pada perhitungan DFL loss.

Tabel 4. 12 Hasil Training Student model 80%

Eksperimen 80%										
Model	Epocs	Box Loss	Cls Loss	DFL Loss	Precision	Recall	mAP50	mAP50-95	Inference Time	Waktu Training (Jam)
Pre Trained Yolov8	283	0.381	0.378	0.834	0.904	0.92	0.954	0.891	4.6ms	0.748
Pre Trained Yolov11	286	0.408	0.405	0.847	0.904	0.932	0.959	0.894	4.6ms	0.822
Teacher Model Yolov8	103	0.394	0.407	0.842	0.875	0.926	0.946	0.88	4.9ms	0.275
Teacher Model Yolov11	293	0.376	0.359	0.837	0.906	0.913	0.957	0.892	5.4ms	0.842
Pseudo Model Yolov8	156	0.421	0.488	0.850	0.843	0.909	0.945	0.873	4.5ms	0.420
Pseudo Model Yolov11	226	0.429	0.532	0.866	0.862	0.902	0.941	0.87	4.5ms	0.226



(a) Pretrained Model Yolov8



(b) Pretrained Model Yolov11



(c) Teacher Model Yolov8



(d) Teacher Model Yolov11



(e) Pseudo Model Yolov8



(f) Pseudo Model Yolov11



(g) Fully Supervised Yolov8



(h) Fully Supervised Yolov11



(i) Ground Truth

Gambar 4. 51 Visual Test Semi Supervised 80%

Tabel 4. 13 Hasil Visual Semi Supervised 80%

Lokasi	Ground Truth	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
A1	Lactogen 1 Happynutri 24x350g	TP	TP	TP	TP	TP	TP	TP	TP
A2	Lactogen 2 Happynutri 24x350g	TP	TP	TP	TP	TP	TP	TP	TP
A3	Lactogrow 3 Happynutri 24x350g	TP	TP	TP	TP	TP	TP	TP	TP
A4	Lactogen 1 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
A5	Lactogen 2 Happynutri 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
A6	Lactogrow 3 Happynutri Honey 12x750g	TP	TP	TP	TP	TP	TP	TP	TP

B2	Lactogrow 3 Happynutri Honey 24x350g	TP	TP	TP	FP – Lactogrow 3 Happynutri Honey 40x180g	TP	TP	TP	TP
B3	Lactogrow 3 Happynutri 12x750g	TP	TP	TP	TP	TP	FP – Lactogrow 3 Happynutri 24x350g	TP	TP
B4	Lactogrow 3 Happynutri Van 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
B5	Lactogrow 4 Happynutri Honey 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
B6	Lactogrow 4 Happynutri Van 12x750g	TP	TP	TP	TP	TP	TP	TP	TP
C1	Lactogen 1 Happynutri 40x180g	TP	TP	TP	TP	TP	TP	TP	TP
C2	Lactogen 2 Happynutri 40x180g	TP	TP	TP	TP	TP	TP	TP	TP

Lokasi objek dinyatakan berdasarkan posisi rak dan urutan dari kiri ke kanan, dengan A menunjukkan rak atas, B rak tengah, dan C rak bawah. Angka menunjukkan urutan objek pada rak (misalnya B5 adalah objek ke-5 dari kiri pada rak tengah). Keterangan noBB menunjukkan tidak terdapat bounding box pada objek tersebut, sedangkan FP – … menunjukkan objek yang terdeteksi secara keliru oleh model.

Eksperimen dengan 80% data berlabel menghasilkan performa terbaik dari seluruh pengujian student model. Seluruh model mencapai mAP50 di atas 0.94, dengan dua model tertinggi Pretrained YOLOv11 dan Teacher YOLOv11 masing-masing mencapai 0.959 dan

0.957. Analisis detail per objek disajikan pada Tabel 4.13 dengan visualisasi pada Gambar 4.51. Berdasarkan Tabel 4.13, performa deteksi menunjukkan hasil yang sangat konsisten:

a. Model (a) Pretrained YOLOv8 mencapai deteksi sempurna dengan berhasil mendeteksi 13 dari 13 objek yang terdapat pada ground truth. Namun, terdapat 4 false positive pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 1+ (C5), Dancow 3+ (C6), dan SGM. Keempat objek tidak dianotasi pada ground truth akibat keterbatasan visibilitas, seperti tulisan gramasi yang tidak terlihat secara jelas serta objek yang hanya tampak sebagian (contohnya pada produk SGM). Walaupun deteksi model menunjukkan kesesuaian kelas dan posisi, hasil tersebut tetap diklasifikasikan sebagai false positive berdasarkan skema evaluasi yang digunakan.
b. Model (b) Pretrained YOLOv11 berhasil mendeteksi 13 dari 13 objek pada ground truth dengan baik. Meskipun demikian, terdapat 4 false positive, yaitu pada objek Lactogrow 3 Happynutri Honey 40x180g (B1), Dancow 1+ (C5), dan Dancow 3+ (C6) yang tidak memiliki label dan bounding box pada ground truth, tetapi tetap terdeteksi oleh model dengan kelas yang benar. Objek-objek tersebut tidak dilabeli pada ground truth karena hanya terlihat sebagian atau berada pada sudut pengambilan gambar yang menyebabkan informasi visual tidak cukup jelas untuk proses anotasi. Selain itu, terdapat kesalahan klasifikasi pada objek Vidoran (sudut kiri atas), objek tersebut tidak diberi label maupun bounding box pada ground truth, tetapi terdeteksi secara keliru sebagai Dancow Baby 2 Box 12x550 g. Objek-objek tersebut tetap dikategorikan sebagai false positive karena tidak tercantum dalam anotasi ground truth yang digunakan sebagai acuan evaluasi.
c. Model (c) Teacher YOLOv8 mencapai deteksi sempurna dengan mendeteksi 13 dari 13 objek pada ground truth. Namun, terdapat 4 false positive, yaitu pada objek Lactogrow 3 Happynutri Honey 40x180 g (B1), Dancow 1+ (C5), dan Dancow 3+ (C6) yang tidak memiliki label dan bounding box pada ground truth, tetapi tetap terdeteksi oleh model dengan kelas yang benar. Objek-objek tersebut tidak dilabeli pada ground truth karena hanya terlihat sebagian atau berada pada sudut pengambilan gambar yang menyebabkan informasi visual tidak cukup jelas untuk proses anotasi. Selain itu, terdapat 1 objek SGM yang tidak memiliki label dan bounding box juga tetapi dideteksi salah oleh model, diklasifikasikan sebagai Dancow 1+ Var Advn ExcNutr 40x200 g (Sudut kanan bawah). Objek-objek tersebut tetap dikategorikan

sebagai false positive karena tidak tercantum dalam anotasi ground truth yang digunakan sebagai acuan evaluasi.

d. Model (d) Teacher YOLOv11 mendeteksi 12 dari 13 objek pada ground truth dengan benar. Terdapat 6 false positive, yaitu kesalahan klasifikasi pada objek Lactogrow 3 Honey 24x350 g (B2) yang terdeteksi sebagai Lactogrow 3 Happynutri Honey 40x180 g (B2). Selain itu, terdapat false positive pada objek Lactogrow 3 Happynutri Honey 40x180 g (B1), Dancow 1+ (C5), Dancow 3+ (C6), serta 2 objek SGM yang tidak diberi label maupun bounding box pada ground truth, meskipun terdeteksi oleh model dengan kelas dan posisi yang sesuai. Hal ini disebabkan oleh keterbatasan informasi visual pada objek, seperti tulisan gramasi yang tidak terlihat secara jelas dan objek yang hanya terlihat sebagian seperti pada produk SGM, sehingga objek tidak memenuhi kriteria anotasi. Oleh karena itu, deteksi tersebut tetap dikategorikan sebagai false positive karena tidak tercantum dalam ground truth yang digunakan sebagai acuan evaluasi.
e. Model (e) Pseudo YOLOv8 mencapai deteksi yang baik dengan mendeteksi 13 dari 13 objek pada ground truth. Namun, terdapat 2 false positive, yaitu pada kelas Lactogrow 3 Happynutri Honey 40×180 g (B1) dan Dancow 1+ (C5). deteksi pada kedua objek yang dihasilkan model menunjukkan kesesuaian kelas secara visual. Namun, objek tersebut tidak dianotasi pada ground truth karena informasi visual yang tidak memadai untuk proses anotasi, seperti tulisan gramasi yang tidak terlihat secara utuh. Oleh karena itu, deteksi ini dikategorikan sebagai false positive sesuai dengan skema evaluasi berbasis ground truth.
f. Model (f) Pseudo YOLOv11 berhasil mendeteksi 12 dari 13 objek pada ground truth dengan benar. Terdapat 3 false positive, yaitu kesalahan klasifikasi pada objek Lactogrow 3 12×750 g yang terdeteksi sebagai Lactogrow 3 Happynutri 24×350 g (B3), serta objek Lactogrow 3 Happynutri Honey 40×180 g (B1), Dancow 1+ (C5), dan SGM yang tidak diberi label maupun bounding box pada ground truth, meskipun terdeteksi oleh model dengan kelas yang sesuai. Hal ini disebabkan oleh keterbatasan informasi visual pada objek, seperti tulisan gramasi yang tidak terlihat secara jelas, sehingga objek tidak memenuhi kriteria anotasi. Oleh karena itu, deteksi tersebut tetap dikategorikan sebagai false positive karena tidak tercantum dalam ground truth yang digunakan sebagai acuan evaluasi.

Pada eksperimen dengan 80% data berlabel, seluruh model menunjukkan performa yang sangat tinggi baik secara kuantitatif maupun kualitatif, sebagaimana ditunjukkan pada Gambar 4.51(a–h). Pretrained YOLOv8 dan Pretrained YOLOv11 mencapai deteksi sempurna 13/13 objek ground truth, yang selaras dengan nilai precision 0,904 dan recall masing-masing 0,920 dan 0,932, serta mAP50 di atas 0,95. False positive yang muncul pada kedua model didominasi oleh produk yang secara visual benar namun tidak dilabeli pada ground truth karena keterbatasan visibilitas objek, serta beberapa kesalahan klasifikasi pada kemasan produk yang memiliki kemiripan visual.

Pada model semi-supervised dengan teacher model, Teacher YOLOv8 mampu mendeteksi seluruh objek ground truth dengan recall 0,926, sedangkan Teacher YOLOv11 mendeteksi 12 dari 13 objek dengan nilai precision tertinggi sebesar 0,906. Meskipun demikian, Teacher YOLOv11 menghasilkan jumlah false positive yang relatif lebih tinggi, terutama pada objek dengan variasi ukuran dan kemasan yang mirip, seperti produk Lactogrow dan SGM, yang secara visual benar namun tidak tercantum pada ground truth. Hal ini menunjukkan bahwa pada proporsi data berlabel yang besar, kesalahan model lebih banyak dipengaruhi oleh ambiguitas anotasi daripada keterbatasan pembelajaran model.

Pada model semi-supervised dengan pseudo model, Pseudo YOLOv8 dan Pseudo YOLOv11 juga menunjukkan performa yang stabil dengan recall di atas 0,90 dan mAP50–95 mendekati 0,87. Pseudo YOLOv8 mampu mendeteksi seluruh 13 objek ground truth dengan false positive yang relatif lebih sedikit, sedangkan Pseudo YOLOv11 mendeteksi 12 dari 13 objek dengan beberapa kesalahan klasifikasi pada produk dengan kemiripan visual tinggi. Nilai inference time pada seluruh model relatif konsisten di kisaran 4,5–5,4 ms, menunjukkan bahwa peningkatan akurasi tidak berdampak signifikan terhadap waktu inferensi.

Sebagai pembanding, model fully supervised YOLOv8 dan YOLOv11 pada Gambar 4.51(g) dan 4.51(h) juga berhasil mendeteksi seluruh 13 objek ground truth, namun tetap menghasilkan false positive pada produk yang tidak dilabeli akibat objek yang tidak terlihat sepenuhnya. Secara keseluruhan, hasil eksperimen ini menunjukkan bahwa pada skenario 80% data berlabel, pendekatan semi-supervised learning khususnya Pretrained dan Teacher YOLOv11 telah mencapai performa yang sangat mendekati fully supervised, dengan kesalahan deteksi yang lebih disebabkan oleh keterbatasan dan inkonsistensi anotasi ground truth dibandingkan oleh kemampuan model itu sendiri.

4.5 Analisis dan Pembahasan Performa

Berdasarkan hasil eksperimen yang telah dilakukan pada empat proporsi data berlabel (20%, 40%, 60%, dan 80%) dengan enam strategi inisialisasi yang berbeda, dapat dilakukan analisis komprehensif mengenai efektivitas pendekatan semi-supervised learning dalam konteks deteksi objek produk retail.

4.5.1 Pengaruh Proporsi Data Berlabel terhadap Performa Model

Hasil eksperimen menunjukkan bahwa proporsi data berlabel memiliki pengaruh yang signifikan terhadap performa model deteksi objek. Pada proporsi 20%, baik YOLOv8 maupun YOLOv11 menghasilkan performa yang masih terbatas dengan mAP50 berkisar antara 0.83-0.88. Peningkatan proporsi data berlabel menjadi 40% memberikan dampak yang baik terhadap performa model. Nilai mAP50 meningkat menjadi rentang 0.896 - 0.912, disertai dengan peningkatan kemampuan deteksi pada sebagian besar kelas. Pada proporsi ini, pseudo-labeling mulai memberikan kontribusi yang terlihat jelas, terutama dalam membantu model mengenali variasi visual produk yang lebih beragam.

Eksperimen dengan 60% data berlabel menghasilkan performa terbaik dari segi efisiensi, dengan mAP50 mencapai 0.931-0.951. Pada proporsi ini, model sudah mampu mendeteksi hampir seluruh objek pada kelas mayoritas dengan konsisten. Yang menarik, peningkatan performa dari 60% ke 80% tidak terlalu signifikan, menunjukkan bahwa model telah mendekati titik saturasi dalam kemampuan pembelajarannya. Pada proporsi 80%, meskipun mAP50 mencapai nilai tertinggi (0.941-0.959). Hal ini mengindikasikan bahwa proporsi 60% dapat dianggap sebagai titik optimal yang menyeimbangkan antara performa deteksi dan effort dalam anotasi.

4.5.2 Perbandingan Strategi Inisialisasi Model

Dari enam strategi inisialisasi yang diuji (pretrained, teacher model, dan pseudo model untuk masing-masing arsitektur), hasil menunjukkan pola yang konsisten. Strategi pretrained model secara konsisten memberikan performa terbaik pada semua proporsi data berlabel. Pada eksperimen 80%, pretrained YOLOv11 mencapai mAP50 tertinggi sebesar 0.959, diikuti oleh pretrained YOLOv8 dengan 0.954.

Teacher model yang diinisialisasi dengan bobot hasil supervised training juga menunjukkan performa yang sangat baik, terutama pada proporsi data tinggi. Pada eksperimen 60%, teacher model YOLOv8 mencapai mAP50 sebesar 0.936, hampir setara dengan

pretrained model (0.931). Keunggulan teacher model terletak pada waktu pelatihan yang lebih singkat, karena model sudah memiliki representasi fitur yang relevan dari data berlabel.

Pseudo model yang dilatih hanya menggunakan pseudo-label menunjukkan performa yang bervariasi. Pseudo model mampu memberikan hasil yang kompetitif, tetapi tidak melampaui dua model yang lain, karena sangat bergantung kepada kualitas pseudo model yang dihasilkan.

4.5.3 Kualitas Pseudo-Label dan Dampaknya terhadap Performa

Evaluasi kualitas pseudo-label mengungkapkan pola yang menarik. Pada proporsi 20% dan 40%, kualitas pseudo-label masih cukup baik dengan precision mencapai 0.797 - 0.814 dan recall 0.755 - 0.789. Namun, pada proporsi 80%, terjadi penurunan kualitas pseudo-label yang ditunjukkan dengan nilai precision yang turun menjadi 0.564-0.594. Analisis klasterisasi false positive mengidentifikasi empat penyebab utama: ketidakkonsistenan ground truth, kemiripan visual antar produk, noise visual lingkungan, dan objek parsial di tepi gambar. Penyebab pertama merupakan kontributor terbesar, dimana pseudo-label sebenarnya benar secara visual tetapi tercatat sebagai false positive karena keterbatasan anotasi manual.

Fenomena penurunan kualitas pada proporsi 80% dapat dijelaskan dengan berkurangnya jumlah data untuk pseudo-labeling (147 gambar) yang menyebabkan pseudo-label memiliki variasi yang lebih terbatas. Hal ini mengonfirmasi bahwa pseudo-labeling paling efektif ketika terdapat keseimbangan antara data berlabel dan tidak berlabel.

4.5.4 Perbandingan Visual Testing Semi Supervised Learning dengan Fully Supervised

Perbandingan visual antara pendekatan semi-supervised learning dan fully supervised learning dilakukan untuk mengevaluasi kemampuan model dalam mendeteksi objek secara kualitatif, khususnya pada kondisi rak produk retail yang padat, objek saling berdekatan, serta sebagian objek tidak terlihat sepenuhnya. Analisis visual ini mengacu pada hasil pengujian yang ditampilkan pada Gambar 4.30(a–f) untuk model semi-supervised dan Gambar 4.30(g–h) untuk model fully supervised.

Berdasarkan hasil visual, model fully supervised YOLOv8 dan YOLOv11 mampu mendeteksi seluruh 13 objek ground truth secara konsisten. Namun, masih ditemukan false positive pada beberapa produk seperti Lactogrow 3 Happynutri Honey 40x180g, Dancow 1+, Dancow 3+, dan produk SGM. False positive tersebut terjadi karena objek-objek tersebut terlihat secara visual tetapi tidak dilabeli pada ground truth akibat kondisi objek yang

terpotong, tertutup sebagian, atau berada di tepi gambar. Hal ini menunjukkan bahwa meskipun fully supervised memiliki akurasi tinggi, performa visualnya tetap dipengaruhi oleh keterbatasan dan inkonsistensi anotasi ground truth.

Pada sisi lain, model semi-supervised learning baik dengan inisiasi pretrained, teacher, maupun pseudo model menunjukkan kemampuan visual yang mendekati bahkan setara dengan fully supervised, terutama pada eksperimen dengan proporsi data berlabel 60% dan 80%. Pada skenario tersebut, sebagian besar model semi-supervised mampu mendeteksi seluruh objek ground truth dengan baik, dengan kesalahan yang didominasi oleh false positive pada objek yang secara visual benar tetapi tidak tercantum dalam ground truth. Hal ini menandakan bahwa model semi-supervised memiliki kemampuan generalisasi yang kuat dalam mengenali pola visual produk retail, meskipun dilatih dengan jumlah data berlabel yang lebih sedikit.

4.5.5 Efektivitas Semi-Supervised Learning

Perbandingan langsung antara model semi-supervised dan fully supervised menunjukkan efektivitas pendekatan yang diusulkan. Model semi-supervised dengan 60% data berlabel (pretrained YOLOv11) mencapai mAP50 sebesar 0.951, yang mendekati performa fully supervised mAP50 0.953 dan 0.96. Dengan kata lain, model dapat menghemat 40% effort dalam anotasi manual sambil mempertahankan performa yang hampir setara.

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, maka diperoleh beberapa kesimpulan sebagai berikut:

a. Performa berdasarkan metrik evaluasi antara model semi-supervised dengan fully supervised menunjukkan hasil yang kompetitif. Hal ini menegaskan bahwa metode pseudo-labeling dapat menjadi solusi praktis jika ketersediaan data berlabel yang terbatas.
b. Rasio optimal data berlabel dan tidak berlabel untuk skenario semi supervised terdapat pada proporsi 60%. Pada kondisi ini, model sudah memiliki performa yang stabil, dengan mAP50 mendekati fully supervised, namun tetap menghemat sebagian beban anotasi.

Secara keseluruhan, penelitian ini berhasil mengembangkan model deteksi produk retail yang lebih efisien dengan pendekatan semi-supervised learning, menganalisis keunggulannya dibandingkan supervised learning, serta menentukan rasio optimal data berlabel dan tak berlabel untuk mendukung proses pelabelan produk retail yang lebih efektif.

5.2 Saran

Berdasarkan temuan penelitian, beberapa saran yang dapat diajukan untuk pengembangan lebih lanjut adalah sebagai berikut:

a. Peningkatan kualitas pseudo-label

   Perlu dilakukan penelitian lebih lanjut terkait mekanisme seleksi pseudo-label, misalnya dengan threshold confidence yang lebih tinggi dan menambah iteratif pseudo-labeling, agar kualitas pseudo-label lebih terjamin dan tidak menurunkan performa model.
b. Perbandingan dengan metode SSL lain

   Penelitian lanjutan dapat membandingkan pseudo-labeling dengan metode semi-supervised learning berbasis consistency regularization atau generative approaches, sehingga dapat diperoleh gambaran yang lebih komprehensif terkait pendekatan paling efektif untuk deteksi produk retail.

DAFTAR PUSTAKA

Cai, Yuanqiang, Longyin Wen, Libo Zhang, Dawei Du, and Weiqiang Wang. 2021. "Rethinking Object Detection in Retail Stores." Pp. 947–54 in *35th AAAI Conference on Artificial Intelligence, AAAI 2021*. Vol. 2A.

Chauhan, Jaydeep, Srikrishna Varadarajan, and Muktabh Mayank Srivastava. 2021. "Semi-Supervised Learning for Dense Object Detection in Retail Scenes." *ArXiv Preprint ArXiv:2107.02114* 1–4.

Digo Saputra, Raihan, and Dhomas Hatta Fudholi. 2023. "Model Mobile Untuk Deteksi Objek Pada On-Shelf Availability Produk Retail." *AUTOMATA* 4(2).

Fudholi, Dhomas Hatta, Arrie Kurniawardhani, Gabriel Imam Andaru, Ahmad Azzam Alhanafi, and Nabil Najmudin. 2024. "YOLO-Based Small-Scaled Model for On-Shelf Availability in Retail." *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)* 8(2):265–71. doi:10.29207/resti.v8i2.5600.

Girshick, Ross. 2015. "Fast R-CNN." *Proceedings of the IEEE International Conference on Computer Vision* 2015:1440–48. doi:10.1109/ICCV.2015.169.

Guimarães, Vânia, Jéssica Nascimento, Paula Viana, and Pedro Carvalho. 2023. "A Review of Recent Advances and Challenges in Grocery Label Detection and Recognition." *Applied Sciences (Switzerland)* 13(5). doi:10.3390/app13052871.

Hidayatullah, Priyanto, Nurjannah Syakrani, Muhammad Rizqi Sholahuddin, Trisna Gelar, and Refdinal Tubagus. 2025. "YOLOv8 to YOLO11: A Comprehensive Architecture In-Depth Comparative Review."

Jha, Dipendra, Ata Mahjoubfar, and Anupama Joshi. 2022. *Designing an Efficient End-to-End Machine Learning Pipeline for Real-Time Empty-Shelf Detection*. Association for Computing Machinery.

Jiang, Peiyuan, Daji Ergu, Fangyao Liu, Ying Cai, and Bo Ma. 2021. "A Review of Yolo Algorithm Developments." *Procedia Computer Science* 199:1066–73. doi:10.1016/j.procs.2022.01.135.

Khanam, Rahima, and Muhammad Hussain. 2024. "YOLOv11: An Overview of the Key Architectural Enhancements." 2024:1–9.

Li, Gang, Xiang Li, Yujie Wang, Yichao Wu, Ding Liang, and Shanshan Zhang. 2022. "PseCo: Pseudo Labeling and Consistency Training for Semi-Supervised Object Detection." Pp. 457–72 in *Computer Vision -- ECCV 2022*. Vol. 13669 LNCS. Cham.

Li, Hengduo, Zuxuan Wu, Abhinav Shrivastava, and Larry S. Davis. 2022. “Rethinking Pseudo Labels for Semi-Supervised Object Detection.” *Proceedings of the 36th AAAI Conference on Artificial Intelligence, AAAI 2022* 36:1314–22. doi:10.1609/aaai.v36i2.20019.

Liu, Yen-cheng, Chih-yao Ma, Zijian He, Chia-wen Kuo, Kan Chen, Peizhao Zhang, Bichen Wu, Zsolt Kira, and Peter Vajda. 2021. “UNBIASED TEACHER FOR SEMI-SUPERVISED OBJECT DETECTION.” *ICLR 2021 - 9th International Conference on Learning Representations* 1–17.

Mey, Alexander, and Marco Loog. 2023. “Improved Generalization in Semi-Supervised Learning: A Survey of Theoretical Results.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45(4):4747–67. doi:10.1109/TPAMI.2022.3198175.

Milella, Annalisa, Antonio Petitti, Roberto Marani, Grazia Cicirelli, and Tiziana D’Orazio. 2020. “Towards Intelligent Retail: Automated on-Shelf Availability Estimation Using a Depth Camera.” *IEEE Access* 8:19353–63. doi:10.1109/ACCESS.2020.2968175.

Moussaoui, Issam, Brent D. Williams, Christian Hofer, John A. Aloysius, and Matthew A. Waller. 2016. “Drivers of Retail On-Shelf Availability: Systematic Review, Critical Assessment, and Reflections on the Road Ahead.” *International Journal of Physical Distribution and Logistics Management* 46(5):516–35. doi:10.1108/IJPDLM-11-2014-0284.

Ouali, Yassine, Céline Hudelot, and Myriam Tami. 2020. “An Overview of Deep Semi-Supervised Learning.” *ArXiv Preprint ArXiv:2006.05278* (Jul. 2020):1–43.

Qi, Jianchun, Minh Nguyen, and Wei Qi Yan. 2024. “CISO: Co-Iteration Semi-Supervised Learning for Visual Object Detection.” *Multimedia Tools and Applications* 83(11):33941–57. doi:10.1007/s11042-023-16915-4.

Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. “You Only Look Once: Unified, Real-Time Object Detection.” Pp. 779–88 in *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Rizve, Mamshad Nayeem, Kevin Duarte, Yogesh S. Rawat, and Mubarak Shah. 2021. “In Defense of Pseudo-Labeling: An Uncertainty-Aware Pseudo-Label Selection Framework for Semi-Supervised Learning.” *ICLR 2021 - 9th International Conference on Learning Representations* 1–20.

Rizzotti, Alessandro, and Prof Ivan Russo. 2025. “Doctoral Program in ACCOUNTING AND MANAGEMENT In Agreement with Università Degli Studi Di Verona XXXVII Cycle Challenges in Retail Supply Chains : Strategies for Optimizing Logistics and Enhancing

On-Shelf Availability Doctoral Candidate Supervisor Federi."

Santos Júnior, Eder Silva dos, Thuanne Paixão, and Ana Beatriz Alvarez. 2025. "Comparative Performance of YOLOv8, YOLOv9, YOLOv10, and YOLOv11 for Layout Analysis of Historical Documents Images." *Applied Sciences (Switzerland)* 15(6). doi:10.3390/app15063164.

Shehzadi, Tahira, Ifza, Didier Stricker, and Muhammad Zeshan Afzal. 2024. "Semi-Supervised Object Detection: A Survey on Progress from CNN to Transformer." 1–21.

Smith, John. 2025. "Advances in Semi-Supervised Learning Techniques for Real-World Applications." *American Journal of Machine Learning* 6(1):9–20.

Sohan, Mupparaju, Thotakura Sai Ram, and Ch. Venkata Rami Reddy. 2024. "A Review on YOLOv8 and Its Advancements." (May):529–45. doi:10.1007/978-981-99-7962-2_39.

Sohn, Kihyuk, David Berthelot, Chun-Liang Li, Zizhao Zhang, Nicholas Carlini, Ekin D. Cubuk, Alex Kurakin, Han Zhang, and Collin Raffel. 2022. "FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence." *IEEE Transactions on Industrial Informatics* 37(10):1575–85.

Sohn, Kihyuk, David Berthelot, Chun-Liang Li, Zizhao Zhang, Nicholas Carlini, Ekin Dogus Cubuk, Alexey Kurakin, Han Zhang, Colin A. Collin Raffel, Zizhao Zhang, Han Zhang, Colin A. Collin Raffel, Ekin Dogus Cubuk, Alexey Kurakin, and Chun-Liang Li. 2020. "FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence." Pp. 596–608 in *Advances in Neural Information Processing Systems*. Vol. 33, edited by H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin. Curran Associates, Inc.

Sohn, Kihyuk, Zizhao Zhang, Chun-Liang Li, Han Zhang, Chen-Yu Lee, and Tomas Pfister. 2020. "A Simple Semi-Supervised Learning Framework for Object Detection."

Tzutalin. 2015."LabelImg." https://github.com/tzutalin/labelImg.

Wei, Yuchen, Son Tran, Shuxiang Xu, Byeong Kang, and Matthew Springer. 2020. "Deep Learning for Retail Product Recognition: Challenges and Techniques." *Computational Intelligence and Neuroscience* 2020(Nov. 2020):23. doi:10.1155/2020/8875910.

Wu, Xiongwei, Doyen Sahoo, and Steven C. H. Hoi. 2020. "Recent Advances in Deep Learning for Object Detection." *Neurocomputing* 396(d):39–64. doi:10.1016/j.neucom.2020.01.085.

Yang, Xiangli, Zixing Song, Irwin King, and Zenglin Xu. 2023. "A Survey on Deep Semi-Supervised Learning." *IEEE Transactions on Knowledge and Data Engineering* 35(9):8934–54. doi:10.1109/TKDE.2022.3220219.

Yilmazer, Ramiz, and Derya Birant. 2021. “Shelf Auditing Based on Image Classification Using Semi-Supervised Deep Learning to Increase On-Shelf.” *Sensors Article* 21(2).

LAMPIRAN