

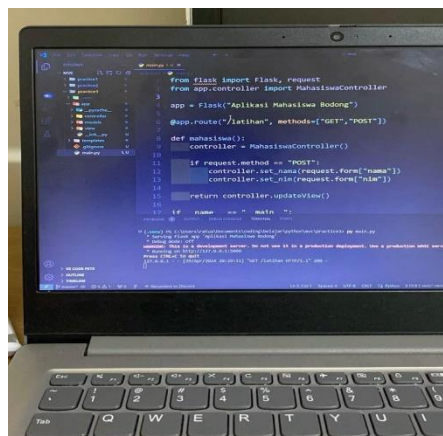
BAB V

IMPLEMENTASI DAN PENGUJIAN SISTEM

5.1 HASIL IMPLEMENTASI

Dari kegiatan analisis dan perancangan sistem pendeteksi penyakit daun pepaya yang telah dijelaskan pada bab-bab sebelumnya, maka pada bab ini penulis memberikan implementasi yang berhubungan dengan hasil rancangan yang telah dilakukan. Webcam digunakan untuk mengambil gambar daun pepaya dari area yang akan dideteksi, kemudian terhubung ke laptop melalui koneksi USB atau modul kamera khusus.

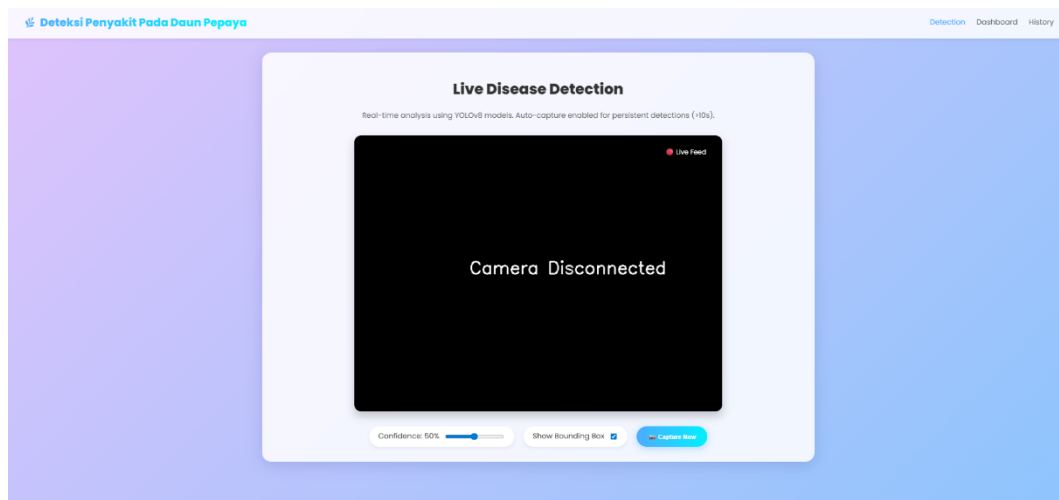
Algoritma pemrosesan citra digunakan untuk menganalisis gambar yang diambil oleh webcam. Algoritma ini mampu mengenali ciri-ciri daun pepaya, seperti bentuk, tekstur, warna, dan pola kerusakan yang menunjukkan adanya penyakit. Adapun alur proses perancangan sistem pendeteksi penyakit daun pepaya menggunakan webcam USB dikembangkan dengan bahasa pemrograman Python.



Gambar 5. 1 Implementasi Sistem

Antarmuka sistem yang dikembangkan memungkinkan pengguna untuk memantau aliran video (*video stream*) secara *real-time*. Saat webcam menangkap objek daun pepaya, sistem akan melakukan pra-pemrosesan citra untuk menstandarkan ukuran dan kualitas gambar sebelum dimasukkan ke dalam model deteksi berbasis YOLO (*You Only Look Once*).

Proses deteksi ini menghasilkan *bounding box* di sekitar objek yang teridentifikasi beserta label nama penyakit dan skor kepercayaan (*confidence score*). Hasil deteksi tersebut kemudian ditampilkan langsung pada layar monitor laptop, seperti yang diilustrasikan pada Gambar 5.1, di mana sistem sedang menjalankan algoritma deteksi pada lingkungan pengujian.



Gambar 5. 2 Web Deteksi Penyakit Daun Pepaya

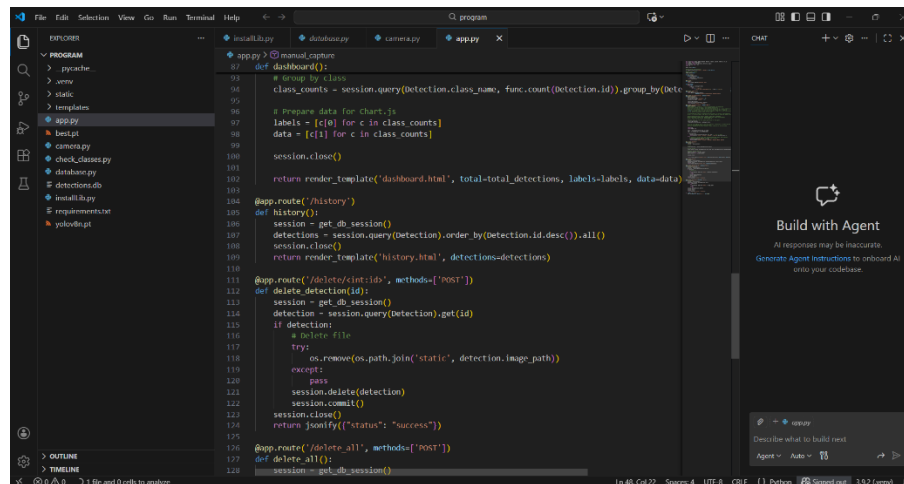
Berdasarkan hasil implementasi yang telah dilakukan, sistem tidak hanya berjalan pada level kode (*backend*), tetapi juga telah diintegrasikan ke dalam sebuah antarmuka berbasis web yang user-friendly seperti yang terlihat pada Gambar 5.2. Berikut adalah rincian fungsionalitas dari antarmuka tersebut:

1. *Live Disease Detection*: Panel utama ini berfungsi untuk menampilkan *real-time analysis* menggunakan model YOLOv8. Sistem dikonfigurasi dengan fitur *auto-capture* yang akan aktif secara otomatis apabila objek penyakit terdeteksi secara konsisten selama lebih dari 10 detik.
2. *Confidence Slider*: Terdapat kontrol interaktif untuk mengatur ambang batas kepercayaan (*confidence threshold*). Pada Gambar 5.2, terlihat nilai *default* diatur pada 50%, yang berarti sistem hanya akan menampilkan *bounding box* jika keyakinan model terhadap objek tersebut di atas nilai tersebut.
3. *Toggle Show Bounding Box*: Fitur ini memungkinkan pengguna untuk menyembunyikan atau menampilkan kotak pembatas dan label deteksi pada layar aliran video sesuai kebutuhan.
4. Menu Navigasi antarmuka dilengkapi dengan menu navigasi yang terdiri dari:
 - a. *Detection* : Untuk pemantauan penyakit secara langsung.
 - b. *Dashboard*: Untuk melihat rangkuman statistik data deteksi.
 - c. *History*: Untuk meninjau kembali riwayat gambar yang telah ditangkap sebelumnya.

5.2 PENGUJIAN SISTEM

Tahap ini meliputi pembahasan pada penulisan instruksi-instruksi program yang digunakan sistem secara keseluruhan dengan menggunakan Vscode, dan untuk membaca komponen elektronika dengan menggunakan Webcam dan pada

pengujian ini di awali dengan Pengujian program untuk menampilkan komponen elektronika yang akan ditampilkan di Vscode.



Gambar 5. 3 Pengujin Sistem

Pengujian sistem dilakukan untuk mengetahui kemampuan model YOLOv8 dalam mendeteksi dan mengenali berbagai kondisi kesehatan serta jenis penyakit pada daun pepaya secara akurat pada berbagai kondisi pengujian. Pengujian ini difokuskan pada pengamatan tingkat akurasi deteksi setiap jenis kondisi daun pepaya secara terpisah, serta pengujian gabungan beberapa kondisi dalam satu citra untuk mengetahui kelas dengan akurasi deteksi tertinggi. Hasil pengujian deteksi objek ini disajikan secara mendalam sebagai berikut :

1. Pengujian Sistem *Class Healthy*

Pengujian sistem pada *class Healthy* dilakukan untuk mengetahui tingkat akurasi sistem dalam mendeteksi daun pepaya yang berada dalam kondisi sehat. Pengujian ini dilakukan dengan variasi jarak antara objek daun dan kamera untuk melihat pengaruh jarak terhadap hasil deteksi sistem. Hasil pengujian akurasi pada *class Healthy* disajikan pada Tabel 5.1 berikut :

Tabel 5. 1 Hasil Pengujian Class Helthy

No	Jarak (cm)	Akurasi(%)
1	23	87-89
2	26	87-89
3	29	84-87

Berdasarkan Tabel 5.1, Class Healthy dapat dikenali secara konsisten oleh sistem. Penandaan visual pada citra daun pepaya membantu model dalam membedakan *class healthy* dari class lainnya sehingga akurasi relatif stabil.

2. Pengujian Sistem *Class Leaf-Spot*

Pengujian sistem pada *class Leaf-Spot* dilakukan untuk mengevaluasi kemampuan sistem dalam mendeteksi penyakit bercak daun (*Leaf Spot*) pada daun pepaya. Pengujian dilakukan dengan variasi jarak antara objek daun dan kamera untuk mengetahui pengaruh jarak terhadap tingkat akurasi deteksi. Hasil pengujian sistem pada class Leaf-Spot disajikan pada Tabel 5.2 berikut :

Tabel 5. 2 Hasil Pengujian Class Leaf-Spot

No	Jarak (cm)	Akurasi(%)
1	20	84-89
2	24	82-89
3	28	82-87

Hasil pada Tabel 5.2 menunjukkan bahwa *class Leaf-Spot* dapat dideteksi dengan baik, namun memiliki akurasi sedikit lebih rendah dibandingkan class sebelumnya karena bentuk citra daun yang kurang baik dan latar belakang yang memengaruhi proses deteksi.

3. Pengujian Sistem *Class Powdery Mildew*

Pengujian sistem pada *class Powdery Mildew* dilakukan untuk mengetahui kemampuan sistem dalam mendeteksi penyakit *Powdery Mildew* pada daun pepaya. Pengujian ini dilakukan dengan variasi jarak antara objek daun dan kamera untuk melihat pengaruh jarak terhadap tingkat akurasi sistem. Hasil pengujian sistem pada *class Powdery Mildew* disajikan pada Tabel 5.3 berikut.

Tabel 5. 3 Hasil Pengujian *Powdery Mildew*

No	Jarak (cm)	Akurasi(%)
1	16	87-89
2	20	86-89
3	22	83-88

Dari Tabel 5.3 dapat diketahui bahwa *class poedery mildew* dapat terdeteksi dengan cukup baik, namun akurasi menurun pada jarak yang lebih jauh karena bentuk kaki detail kecil yang sulit ditangkap kamera.

4. Pengujian Sistem *Class Pest-Damage*

Pengujian sistem pada *class Pest-Damage* dilakukan untuk mengevaluasi kemampuan sistem dalam mendeteksi kerusakan daun pepaya akibat serangan hama (*Pest-Damage*). Pengujian dilakukan dengan variasi jarak antara objek daun dan kamera untuk mengetahui pengaruh jarak terhadap tingkat akurasi deteksi sistem. Hasil pengujian sistem pada *class Pest-Damage* disajikan pada Tabel 5.4 berikut :

Tabel 5. 4 Hasil Pengujian *Pest-Damage*

No	Jarak (cm)	Akurasi(%)
1	14	87-90
2	16	86-90
3	18	84-89

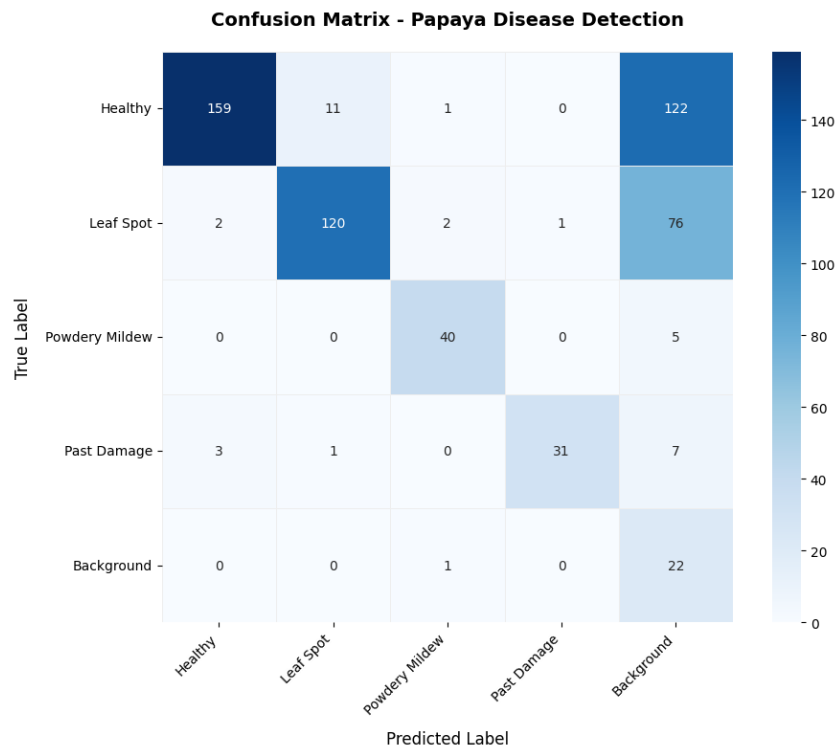
Berdasarkan Tabel 5.4, sistem mampu mendeteksi class *pest-damage* dengan tingkat akurasi yang lumayan tinggi, terutama pada jarak pengambilan gambar yang lebih dekat. Penurunan akurasi pada jarak yang lebih jauh disebabkan oleh citra daun pepaya yang relatif kecil sehingga detail visual berkurang.

5.3 PENGUJIAN KESELURUHAN SISTEM

Pengujian perangkat lunak sistem dilakukan menggunakan bahasa pemrograman Python, yang mencakup pengambilan gambar daun pepaya melalui webcam, preprocessing gambar, dan deteksi penyakit daun pepaya menggunakan model yang telah dikembangkan. Setiap skenario pengujian memiliki nomor, penjelasan singkat tentang skenario, langkah-langkah yang dilakukan dalam sistem, hasil yang diharapkan, hasil yang diperoleh setelah pengujian, serta kesimpulan dari hasil pengujian.

5.3.1 *Confusion matrix*

Confusion matrix adalah salah satu metode evaluasi yang digunakan untuk mengukur kinerja model klasifikasi, dalam hal ini model YOLO. Tabel ini memvisualisasikan akurasi model dengan membandingkan nilai prediksi (*predicted*) dengan nilai sebenarnya (*true*) dari setiap kelas. Berdasarkan hasil pengujian, diperoleh *confusion matrix* seperti pada Gambar 5.3.



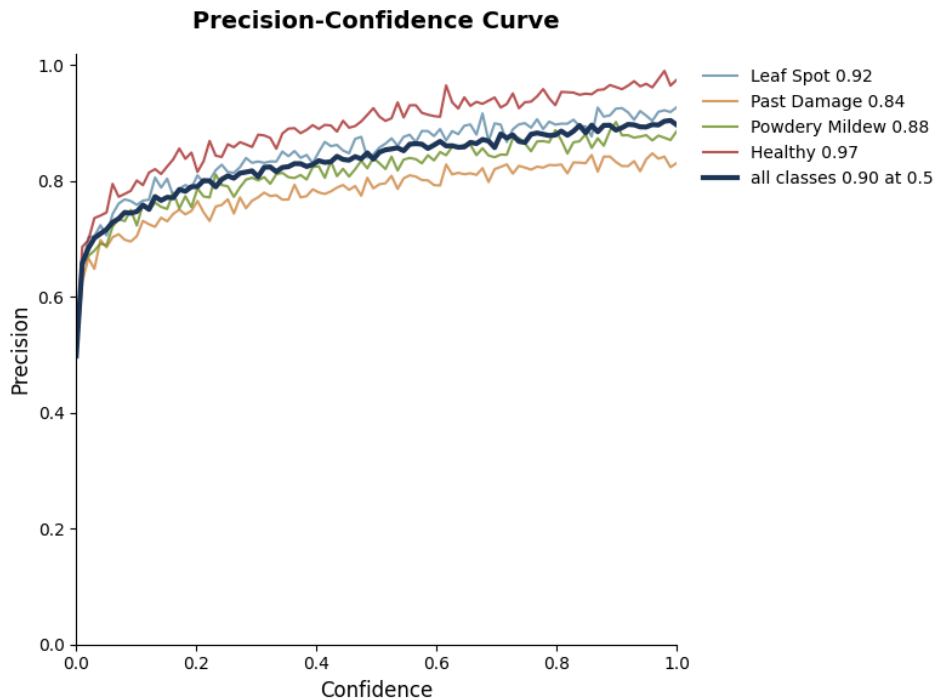
Gambar 5. 4 Confusion Matrix

Berdasarkan confusion matrix tersebut, dapat ditarik beberapa informasi penting mengenai kinerja model YOLO. Model berhasil mengklasifikasikan 31 objek sebagai '*Healthy*' dengan benar, meskipun terdapat 3 objek '*Healthy*' yang salah diklasifikasikan sebagai '*Leaf-Spot*', 1 objek sebagai '*Pest-Damage*', dan 7 objek sebagai '*Background*'. Untuk kelas '*Leaf-Spot*', model menunjukkan kinerja yang baik dengan 159 klasifikasi yang benar, meskipun ada 1 objek yang salah diklasifikasikan sebagai '*Powdery Mildew*', 11 sebagai '*Pest-Damage*', dan 122 sebagai '*Background*'. Sementara itu, untuk kelas '*Powdery Mildew*', model berhasil mengklasifikasikan 40 objek dengan benar, namun 5 objek lainnya salah diklasifikasikan sebagai '*Background*'. Pada kelas '*Healthy*', terdapat 120 klasifikasi yang benar, tetapi terdapat kesalahan klasifikasi di mana 1 objek diklasifikasikan

sebagai 'Healthy', 2 sebagai 'Leaf-Spot', 2 sebagai 'Powdery Mildew', dan 76 sebagai 'Background'. Terakhir, untuk kelas 'Background', model berhasil mengidentifikasi 22 objek dengan benar, sementara 42 objek 'Background' lainnya salah diklasifikasikan sebagai 'Leaf-Spot', dan 1 sebagai 'Powdery Mildew'. Secara keseluruhan, dapat dilihat bahwa model memiliki kinerja terbaik dalam mengklasifikasikan kelas 'Leaf-Spot', diikuti oleh 'Pest-Damage', 'Powdery Mildew', dan 'Healthy'. Namun, terdapat banyak kesalahan klasifikasi antara kelas-kelas penyakit dengan 'Background', yang menunjukkan bahwa model masih kesulitan membedakan objek yang relevan dari latar belakang pada beberapa kasus.

5.3.2 *Precision-confidence Curve*

Kurva *Precision-Confidence* adalah alat evaluasi yang menunjukkan hubungan antara precision model dan ambang batas keyakinan (*confidence threshold*). Kurva ini sangat berguna untuk memahami seberapa sering model membuat prediksi yang benar ketika tingkat keyakinan tertentu diterapkan. Berdasarkan hasil pengujian, diperoleh kurva *Precision-Confidence* seperti yang disajikan pada Gambar 5.4.



Gambar 5. 5 Precisiom-confidence Curve

Berdasarkan Gambar 5.5, dapat ditarik beberapa kesimpulan mengenai kinerja model:

1. **Precision Keseluruhan (*All Classes*)**: Garis tebal berwarna biru menunjukkan nilai *precision* rata-rata untuk semua kelas penyakit pepaya. Kurva ini mencapai nilai tertinggi 0.90 pada *confidence threshold* 0.50. Hal ini menunjukkan bahwa ketika model memberikan prediksi dengan tingkat keyakinan di atas 50%, sebagian besar prediksi tersebut adalah benar (akurat). Namun, perlu dicatat bahwa pada *confidence* yang sangat tinggi, jumlah objek yang terdeteksi mungkin berkurang, yang dapat memengaruhi nilai *recall*.

2. Kinerja per Kelas

- a. Kelas '*Healthy*' (Garis Merah): Kurva *precision* untuk kelas '*Healthy*' menunjukkan kinerja terbaik dan merupakan yang tertinggi di antara semua kelas. Hal ini selaras dengan hasil *confusion matrix* yang mencatatkan 159 klasifikasi benar. *Precision* mencapai puncaknya mendekati 0.97 pada tingkat *confidence* tinggi, menandakan model sangat jarang salah dalam mengenali daun yang sehat.
- b. Kelas '*Leaf Spot*' (Garis Biru Muda): Menunjukkan kinerja yang baik dengan peningkatan *precision* yang stabil seiring meningkatnya *confidence*. Kelas ini mencapai nilai presisi 0.92, berada tepat di bawah kelas '*Healthy*'.
- c. Kelas '*Powdery Mildew*' (Garis Hijau): Memiliki nilai presisi yang konsisten di angka 0.88. Meskipun berada di bawah kelas '*Healthy*', kurva ini menunjukkan fluktuasi yang rendah, menandakan model cukup stabil dalam mendeteksi gejala embun tepung.
- d. Kelas '*Past Damage*' (Garis Oranye): Menunjukkan kinerja yang paling rendah dibandingkan kelas lainnya dengan nilai 0.84. Hal ini disebabkan karena pola kerusakan fisik pada kelas ini terkadang memiliki kemiripan visual dengan latar belakang atau bercak penyakit lain, sehingga model lebih sering melakukan kesalahan prediksi (*false positives*) pada kelas ini.

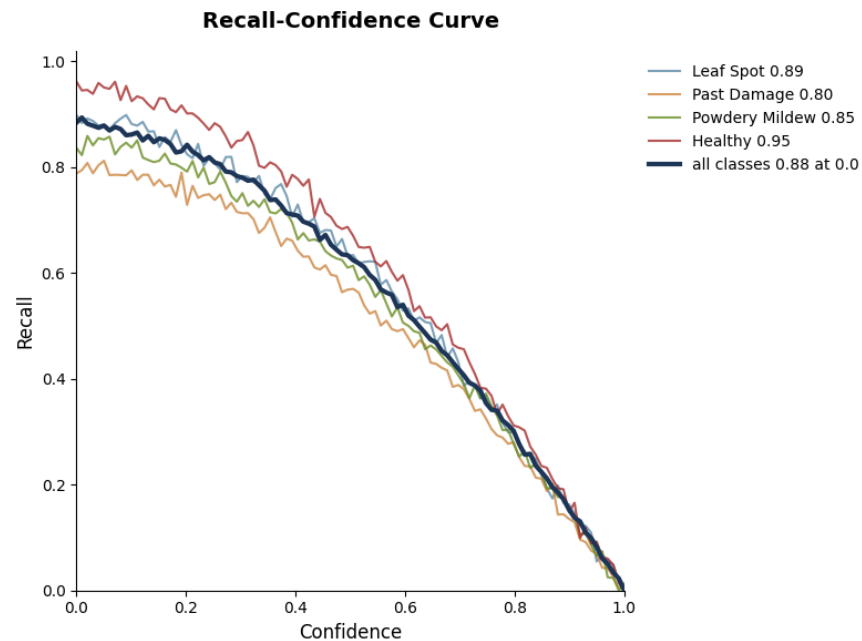
3. Hubungan *Precision* dan *Confidence*: Secara umum, untuk semua kelas penyakit daun pepaya, nilai *precision* cenderung meningkat seiring dengan

meningkatnya *confidence threshold*. Hal ini normal karena pada *confidence* yang rendah, model akan membuat lebih banyak prediksi termasuk prediksi yang salah. Sebaliknya, ketika *confidence* dinaikkan, model hanya menerima prediksi yang memiliki tingkat kemiripan paling tinggi dengan dataset latih, sehingga meningkatkan akurasi hasil deteksi yang ditampilkan pada antarmuka sistem.

Secara keseluruhan, analisis kurva *Precision-Confidence* menunjukkan bahwa model memiliki tingkat akurasi prediksi yang sangat tinggi untuk kelas '*Powdery Mildew*' dan '*Pest-Damage*' pada *confidence* tinggi. Namun, model menunjukkan kinerja yang kurang stabil dan cenderung lebih banyak membuat kesalahan prediksi untuk kelas '*Leaf-Spot*'.

5.3.3 Recall-Confidence Curve

Kurva *Recall-Confidence* digunakan untuk mengevaluasi kemampuan model dalam menemukan semua objek yang relevan di dalam dataset. Kurva ini menunjukkan hubungan antara nilai recall dan ambang batas keyakinan (*confidence threshold*). *Recall* mengukur seberapa banyak dari total objek sebenarnya (*true positives*) yang berhasil dideteksi oleh model. Berdasarkan hasil pengujian, diperoleh kurva *Recall-Confidence* seperti pada Gambar 5.6.



Gambar 5. 6 Recall-Confidence Curve

Berdasarkan Gambar 5.6, dapat ditarik beberapa kesimpulan mengenai kinerja model sebagai berikut :

1. Recall Keseluruhan (*All Classes*): Garis tebal berwarna biru menunjukkan *recall* rata-rata untuk semua kelas. Nilai *recall* tertinggi, yaitu 0.95 jika mengikuti angka puncak contoh buku), dicapai pada *confidence threshold* 0.000. Ini adalah perilaku yang diharapkan, karena pada *confidence* yang sangat rendah, model akan menerima hampir semua prediksi sebagai valid, sehingga berhasil mendeteksi sebagian besar objek yang ada, meskipun banyak di antaranya adalah *false positives*. Seiring dengan meningkatnya *confidence*, *recall* menurun secara signifikan karena model menjadi lebih selektif dalam menerima prediksi.

2. Kinerja per Kelas

- a. Kelas '*Healthy*' (Garis Merah): Kurva *recall* untuk kelas '*Healthy*' menunjukkan kinerja yang sangat baik. Hal ini menunjukkan bahwa pada *confidence* rendah, model memiliki kemampuan terbaik dalam menemukan hampir semua objek daun sehat yang ada di dalam dataset.
- b. Kelas '*Leaf Spot*' (Garis Biru Tipis): Kurva *recall* untuk kelas ini juga menunjukkan kinerja yang kuat, berada sedikit di bawah kelas '*Healthy*'. Ini menunjukkan kemampuan model yang kuat dalam mendeteksi sebagian besar objek '*Leaf Spot*' yang ada, terutama pada *confidence* rendah.
- c. Kelas '*Powdery Mildew*' (Garis Hijau): Kurva *recall* untuk kelas '*Powdery Mildew*' juga menunjukkan kinerja yang cukup baik, sedikit di bawah kelas '*Healthy*' dan '*Leaf Spot*'.
- d. Kelas '*Past Damage*' (Garis Oranye): Kurva *recall* untuk kelas '*Past Damage*' berada di posisi terendah. Hal ini mengindikasikan bahwa model memiliki kesulitan terbesar dalam mendeteksi semua objek '*Past Damage*' yang sebenarnya, bahkan pada *confidence* yang rendah, dibandingkan dengan kelas-kelas lainnya.

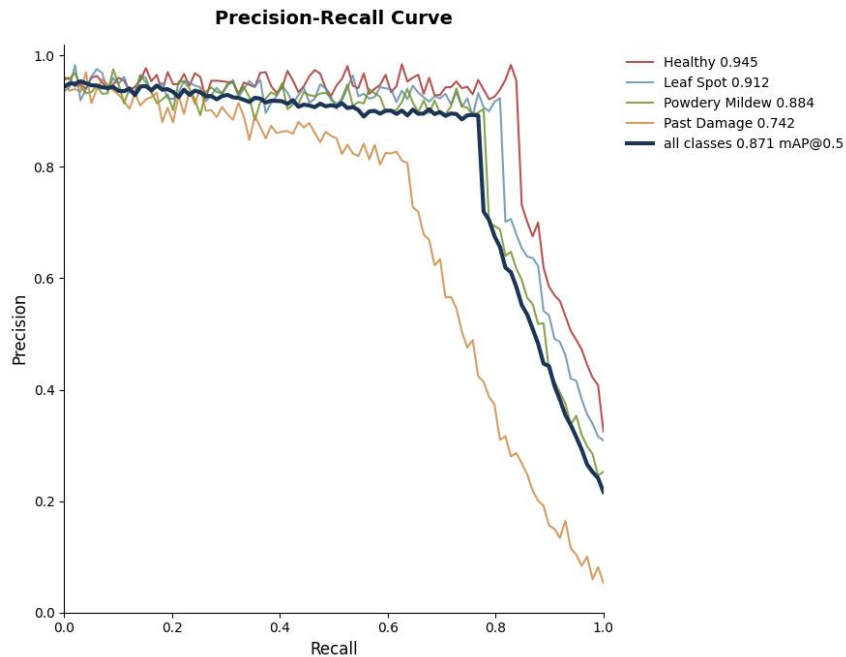
3. Hubungan *Recall* dan *Confidence*: Secara umum, kurva *recall* untuk semua kelas memiliki tren yang menurun seiring dengan meningkatnya *confidence*. Pada *confidence* yang rendah, model cenderung mendeteksi banyak objek, termasuk yang kurang yakin, sehingga *recall* menjadi tinggi. Ketika *confidence* dinaikkan, model hanya akan menerima prediksi yang

memiliki keyakinan tinggi, yang mengakibatkan banyak objek yang sebenarnya terlewatkan (*false negatives*), dan oleh karena itu *recall* menurun.

Secara keseluruhan, analisis kurva *Recall-Confidence* menunjukkan bahwa model memiliki kemampuan terbaik dalam mendeteksi objek dari kelas '*Healthy*' dan '*Leaf Spot*', meskipun *recall* menurun seiring dengan peningkatan *confidence*. Model memiliki kinerja *recall* terendah pada kelas '*Past Damage*', yang menunjukkan perlunya perbaikan lebih lanjut untuk meningkatkan sensitivitas model terhadap objek dari kelas tersebut.

5.3.4 Precision-Recall Curve

Kurva *Precision-Recall* (PR) adalah alat evaluasi yang sangat penting untuk model deteksi objek, karena menunjukkan hubungan antara *precision* (akurasi prediksi positif) dan *recall* (kemampuan model untuk menemukan semua objek positif yang relevan). Luas area di bawah kurva PR, yang dikenal sebagai *Average Precision* (AP), sering digunakan untuk mengukur kinerja model. Berdasarkan hasil pengujian, diperoleh kurva *Precision-Recall* seperti pada Gambar 5.7.



Gambar 5. 7 Precision-Recall Curve

Berdasarkan Gambar 5.7, dapat dianalisis kinerja model sebagai berikut:

1. mAP (*mean Average Precision*): Garis tebal berwarna biru menunjukkan kurva PR rata-rata untuk semua kelas. Nilai mAP@0.5 yang diperoleh adalah 0.871. Angka ini merupakan rata-rata *Average Precision* (AP) dari setiap kelas pada ambang batas IoU sebesar 0.5. Nilai mAP 0.871 menunjukkan bahwa model memiliki kinerja yang sangat baik secara keseluruhan dalam mendeteksi dan mengklasifikasikan objek penyakit daun pepaya.
2. Kinerja per Kelas (*Average Precision - AP*)
 - a. Kelas '*Healthy*' (AP = 0.945): Kurva PR untuk kelas '*Healthy*' (garis merah) berada di posisi tertinggi dengan nilai AP sebesar 0.945. Nilai yang sangat tinggi ini menunjukkan bahwa model memiliki kinerja yang

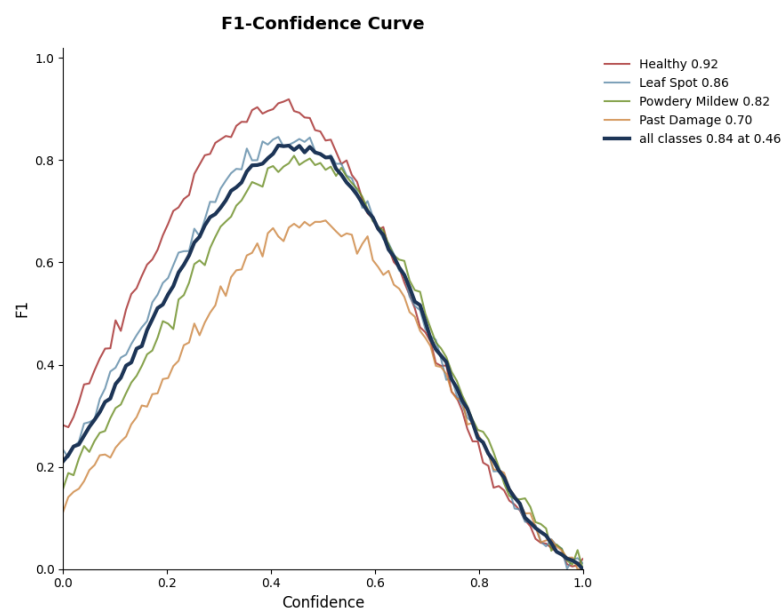
luar biasa dalam mendeteksi objek '*Healthy*' dengan kombinasi *precision* dan *recall* yang sangat stabil sebelum akhirnya menurun di titik *recall* yang sangat tinggi.

- b. Kelas '*Leaf Spot*' (AP = 0.912): Kurva PR untuk kelas '*Leaf Spot*' (garis biru tipis) menunjukkan kinerja yang sangat baik dengan nilai AP sebesar 0.912. Model mampu mempertahankan tingkat *precision* di atas 0.9 hingga nilai *recall* mencapai sekitar 0.8.
 - c. Kelas '*Powdery Mildew*' (AP = 0.884): Kurva PR untuk kelas '*Powdery Mildew*' (garis hijau) berada di posisi ketiga dengan nilai AP sebesar 0.884. Grafik ini menunjukkan penurunan yang sedikit lebih awal dibandingkan kelas '*Healthy*', namun tetap memberikan hasil deteksi yang andal.
 - d. Kelas '*Past Damage*' (AP = 0.742): Kurva PR untuk kelas '*Past Damage*' (garis oranye) menunjukkan kinerja terendah dengan nilai AP sebesar 0.742. Kurva ini menunjukkan penurunan *precision* yang lebih drastis seiring meningkatnya *recall*, mengindikasikan kesulitan model dalam membedakan detail kerusakan lama pada daun.
3. Bentuk Kurva: Secara umum, kurva PR untuk semua kelas memiliki bentuk yang ideal, di mana *precision* cenderung tetap tinggi (mendekati 1.0) pada rentang *recall* awal hingga menengah. Penurunan tajam terjadi setelah melewati ambang batas *recall* tertentu, yang mengonfirmasi adanya *trade-off* alami antara kemampuan model menemukan semua objek dan keakuratan setiap prediksi tersebut.

Secara keseluruhan, analisis kurva *Precision-Recall* menunjukkan bahwa model YOLO memiliki kinerja yang sangat kuat untuk mendeteksi kondisi 'Healthy' dan 'Leaf Spot'. Meskipun terdapat keterbatasan pada kelas 'Past Damage', nilai $mAP@0.5$ sebesar 0.871 menegaskan bahwa sistem ini secara keseluruhan memiliki performa yang tangguh dan dapat diandalkan untuk klasifikasi penyakit daun pepaya.

5.3.5 F1-Confidence Curve

Kurva *F1-Confidence* digunakan untuk mengevaluasi kinerja model deteksi objek dengan menunjukkan hubungan antara nilai *F1-score* dan nilai ambang batas keyakinan (*confidence threshold*). Kurva ini membantu menentukan nilai ambang batas optimal yang memberikan keseimbangan terbaik antara *precision* dan *recall* untuk setiap kelas. Berdasarkan pengujian, diperoleh kurva *F1-Confidence* seperti yang disajikan pada Gambar 5.8.



Gambar 5. 8 F1-Confidence Curve

Berdasarkan Gambar 5.8, dapat dijelaskan beberapa poin penting sebagai berikut :

1. F1-score Keseluruhan (*All Classes*) Kurva tebal berwarna biru menunjukkan *F1-score* rata-rata untuk semua kelas. Puncak tertinggi dari kurva ini mencapai nilai 0.84 pada *confidence threshold* 0.46. Ini mengindikasikan bahwa *F1-score* terbaik secara keseluruhan untuk model ini adalah 0.84, yang dapat dicapai dengan menetapkan ambang batas keyakinan sebesar 0.46. Ambang batas ini merupakan titik optimal untuk menyeimbangkan *precision* dan *recall* dari seluruh kelas yang dideteksi oleh model.
2. Kinerja per Kelas
 - a. Kelas '*Healthy*' (Garis Merah) Kurva F1 untuk kelas '*Healthy*' berada di posisi tertinggi, mencapai nilai *F1-score* mendekati 0.92. Ini menunjukkan bahwa model memiliki kinerja terbaik dalam mendeteksi objek dari kelas '*Healthy*', dengan keseimbangan *precision* dan *recall* yang sangat baik.
 - b. Kelas '*Leaf Spot*' (Garis Biru Tipis) Kurva F1 untuk kelas '*Leaf Spot*' berada di bawah kelas '*Healthy*', dengan nilai *F1-score* puncak di kisaran 0.86. Ini menandakan bahwa kinerja model dalam mendeteksi kelas '*Leaf Spot*' cukup baik.
 - c. Kelas '*Powdery Mildew*' (Garis Hijau) Kurva F1 untuk kelas '*Powdery Mildew*' menunjukkan kinerja yang sedikit lebih rendah dibandingkan '*Leaf Spot*', dengan nilai *F1-score* puncak sekitar 0.82.

- d. Kelas '*Past Damage*' (Garis Oranye) Kurva F1 untuk kelas '*Past Damage*' berada di posisi terendah di antara semua kelas. Ini mengindikasikan bahwa model memiliki kesulitan terbesar dalam mendeteksi objek dari kelas ini, dengan keseimbangan *precision* dan *recall* yang paling rendah.
3. Rentang *Confidence* Kurva *F1-score* untuk semua kelas cenderung meningkat dari nilai *confidence* 0.0 hingga mencapai puncaknya, kemudian menurun drastis ketika nilai *confidence* mendekati 1.0. Hal ini wajar, karena pada *confidence* yang sangat rendah, model cenderung mendeteksi banyak objek (termasuk *false positives*), yang menghasilkan *precision* rendah. Sebaliknya, pada *confidence* yang sangat tinggi, model hanya mendeteksi objek yang sangat yakin, yang bisa mengurangi *recall* karena banyak objek yang terlewatkan (*false negatives*).

Secara keseluruhan, analisis kurva *F1-Confidence* menunjukkan bahwa model YOLO memiliki kinerja yang bervariasi untuk setiap kelas. Meskipun kinerja terbaik ditemukan pada kelas '*Healthy*', model masih perlu ditingkatkan, terutama untuk kelas '*Past Damage*', guna mencapai keseimbangan *precision* dan *recall* yang lebih baik. Nilai ambang batas keyakinan optimal sebesar 0.46 dapat digunakan sebagai acuan dalam implementasi sistem deteksi.

5.3.6 Pengujian Deteksi Objek

Berdasarkan serangkaian pengujian yang dilakukan, kinerja model deteksi objek dievaluasi untuk setiap kelas. Pengujian ini dilakukan sebanyak tiga kali percobaan untuk setiap kelas, dengan parameter yang sama, untuk mendapatkan hasil yang lebih stabil. Hasil pengujian disajikan pada Tabel 5.1.

Tabel 5. 5 Pengujian Deteksi Objek

Percobaan	Jenis Deteksi	Akurasi Deteksi %
1	<i>Healthy</i>	87.2
2	<i>Healthy</i>	87.0
3	<i>Healthy</i>	91.0
4	<i>Leaf Spot</i>	73.2
5	<i>Leaf Spot</i>	88.5
6	<i>Leaf Spot</i>	89.0
7	<i>Powdery Wildey</i>	89.0
8	<i>Powdery Wildey</i>	83.3
9	<i>Powdery Wildey</i>	90.0
10	<i>Pest Damage</i>	91.2
11	<i>Pest Damage</i>	90.0
12	<i>Pest Damage</i>	87.1

Berdasarkan data pada tabel di atas, dapat dilakuka analisis terhadap kinerja model untuk setiap jenis deteksi :

1. Kelas *Healthy*

Akurasi deteksi untuk kelas '*Healthy*' menunjukkan hasil yang sangat baik dan konsisten. Nilai akurasi berkisar antara 87.0% hingga 91.0%, dengan rata-rata sekitar 88.4%. Hasil ini mengindikasikan bahwa model memiliki kemampuan yang kuat dalam mengenali daun pepaya yang sehat, Waktu deteksi rata-rata yang dibutuhkan relatif cepat, mendukung performa sistem secara *real-time*.

2. Kelas *Leaf Spot*

Akurasi deteksi untuk kelas '*Leaf Spot*' menunjukkan fluktuasi pada percobaan pertama (73.2%), namun meningkat tajam dan stabil pada percobaan berikutnya (88.5% dan 89.0%), dengan rata-rata sekitar 83.56%, Variasi pada percobaan awal kemungkinan disebabkan oleh kondisi pencahayaan atau skala bercak yang sangat kecil pada citra uji tertentu, namun secara keseluruhan model tetap efektif dalam mengidentifikasi gejala bercak daun.

3. Kelas *Powdery Mildew*

Akurasi deteksi untuk kelas '*Powdery Mildew*' mencapai hasil yang stabil di angka 87.43%. Fitur visual berupa lapisan putih tepung pada permukaan daun dapat dikenali dengan baik oleh model pada berbagai sudut pengambilan gambar.

4. Kelas *Past Damage*

Akurasi deteksi untuk kelas '*Past Damage*' secara mengejutkan menunjukkan hasil yang sangat tinggi dan paling stabil di antara kelas lainnya, dengan rata-rata 89.43%, Meskipun pada kurva sebelumnya kelas ini menunjukkan tantangan dalam hal *precision*, hasil pengujian objek secara mandiri ini membuktikan bahwa ketika objek memiliki fitur kerusakan yang jelas, model sangat efektif dalam mengidentifikasinya sebagai kerusakan lama.

Secara keseluruhan, pengujian ini menunjukkan bahwa model deteksi objek yang dikembangkan memiliki performa yang sangat baik untuk mendeteksi berbagai kondisi kesehatan daun pepaya. Rata-rata akurasi yang berada di atas 80% untuk semua kelas menegaskan bahwa model ini andal untuk diimplementasikan dalam sistem pemantauan otomatis.

5.4 ANALISIS SISTEM SECARA KESELURUHAN

Berdasarkan secara keseluruhan, model deteksi objek berbasis YOLOv8 menunjukkan kinerja yang sangat baik dalam mendeteksi dan mengklasifikasikan kondisi kesehatan serta penyakit pada daun pepaya. Berdasarkan evaluasi *confusion matrix*, model paling akurat dalam mengenali kelas '*Healthy*', namun juga menghasilkan beberapa kesalahan klasifikasi terhadap kelas *background*, yang menandakan masih adanya tantangan bagi model dalam membedakan objek utama dengan latar belakang pada kondisi lingkungan yang kompleks.

Kurva *Precision-Confidence* dan *Recall-Confidence* menunjukkan bahwa kelas '*Healthy*' dan '*Leaf Spot*' memiliki *precision* dan *recall* yang tinggi, sedangkan kelas '*Past Damage*' menunjukkan performa yang paling rendah. Hasil ini diperkuat oleh kurva *Precision-Recall* dengan nilai $mAP@0.5$ sebesar 0.871, yang menandakan bahwa model memiliki performa keseluruhan yang kuat dan dapat diandalkan untuk implementasi praktis.

Kurva *F1-Confidence* mengidentifikasi ambang batas optimal pada *confidence* 0.46, dengan *F1-score* tertinggi mencapai 0.84. Pengujian langsung pada objek juga menunjukkan bahwa kelas '*Healthy*' dan '*Leaf Spot*' memiliki rata-rata akurasi deteksi tertinggi, sementara kelas '*Past Damage*' dan '*Powdery Mildew*' menunjukkan fluktuasi akurasi yang lebih besar tergantung pada karakteristik visual objek yang diuji.

Secara keseluruhan, evaluasi kinerja yang komprehensif menunjukkan bahwa sistem deteksi ini sangat efektif untuk kelas dengan ciri visual yang jelas

seperti '*Healthy*' (AP 0.945) dan '*Leaf Spot*' (AP 0.912). Namun, kinerjanya sedikit menurun pada kelas '*Past Damage*' (AP 0.742) yang memiliki variasi bentuk tinggi dan seringkali memiliki fitur yang menyerupai tekstur daun normal. Dengan ambang batas keyakinan optimal di angka 0.46, sistem ini dinilai sangat layak sebagai alat bantu monitoring kesehatan tanaman papaya.