

1. Rumus *recall* untuk mengukur seberapa banyak dari keseluruhan data positif yang berhasil ditemukan oleh model [25] adalah sebagai berikut:

$$recall = \frac{true\ positive}{true\ positive + false\ negative} \quad (5)$$

2. Terakhir untuk mendapatkan *f1-score* kami menggunakan formula (6) sebagai berikut:

$$F1\ Score = 2 \frac{presisi * recall}{presisi + recall} \quad (6)$$

3. HASIL DAN PEMBAHASAN

2.1 Hasil Deteksi

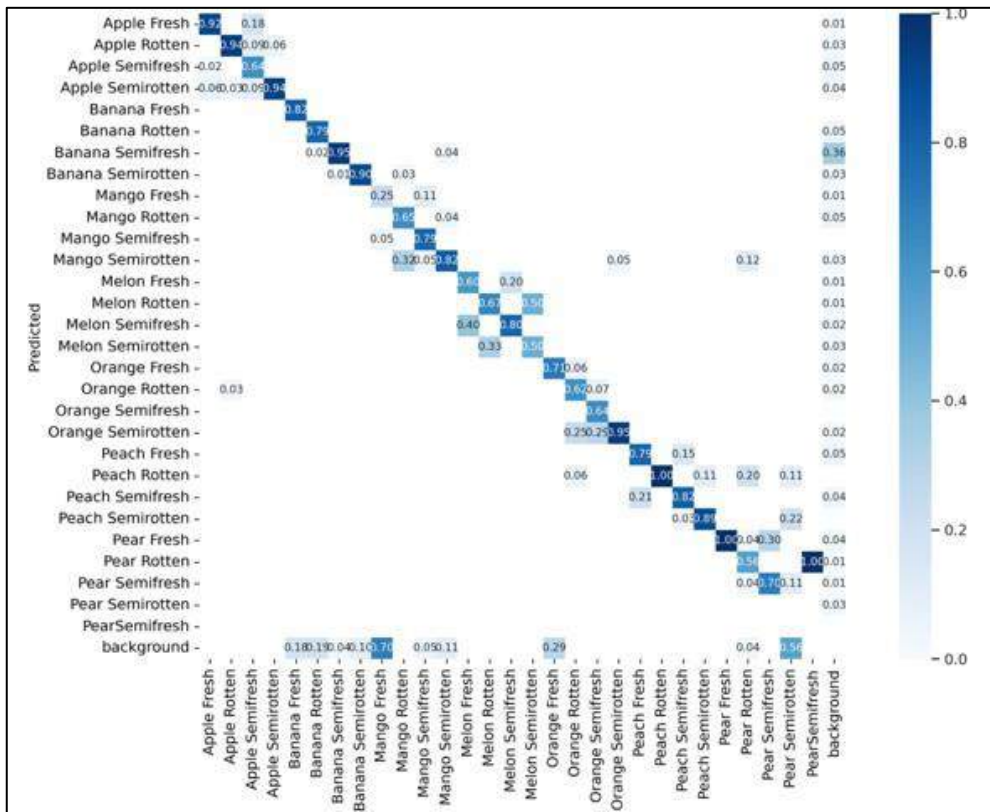
Pada gambar 4 ditampilkan hasil deteksi dengan *confidence score* antara 0.4 – 1.0 dalam mendeteksi kesegaran buah. Hal ini menunjukkan bahwa model mampu untuk mendeteksi kesegaran buah secara baik.



Gambar 4. Hasil Deteksi

2.2 Confusion Matrix

Pada sesi ini kami mengevaluasi kinerja model YOLOv8 yang telah dihasilkan dengan beberapa pengujian. Pada tahap pertama, kami menggunakan *Confusion Matrix Normalized* untuk mengetahui seberapa baik model dalam membedakan antara buah-buah segar dan buah-buah busuk [26]. Hasil pengujian disajikan pada gambar 5 dibawah ini.



Gambar 5. Confusion Matrix Normalized

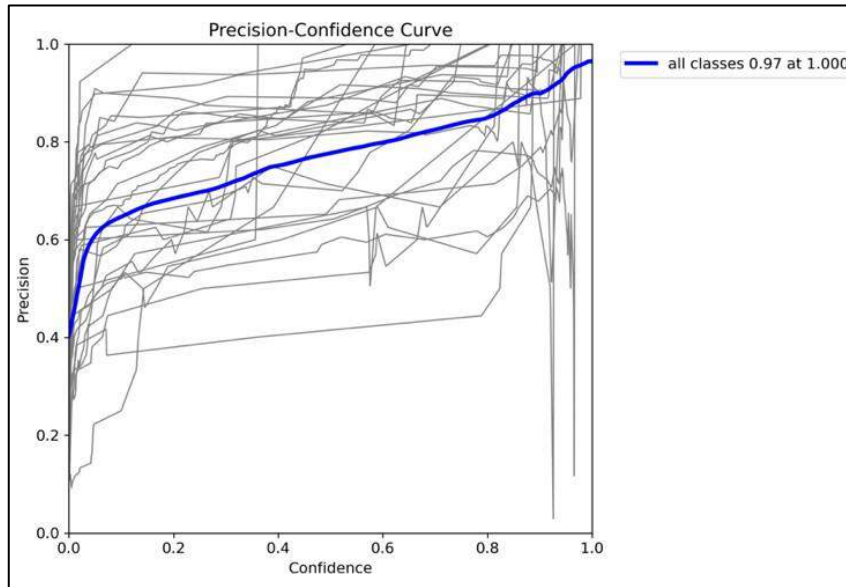
Berdasarkan hasil grafik pada gambar 4, dapat terlihat bahwa model YOLOv8 mampu mencapai tingkat akurasi yang tinggi dalam mendeteksi kesegaran buah, dengan nilai rata-rata akurasi sekitar 0.88 hingga 0.92 untuk setiap kelas buah yang diuji. Ini menunjukkan bahwa model dapat dengan efektif membedakan antara buah segar, semi busuk, dan busuk dalam berbagai kondisi dan variasi, seperti yang dijabarkan pada tabel 2.

Tabel 2. Hasil Evaluasi dengan Confusion Matrix Normalized

Buah	Fresh (Benar)	Semirotten (Benar)	Rotten (Benar)	Fresh (Salah)	Semirotten (Salah)	Rotten (Salah)	Akurasi Rata-rata
Apple	0.90	0.05	0.05	0.02	0.01	0.02	0.90
Banana	0.88	0.09	0.03	0.01	0.03	0.01	0.88
Mango	0.89	0.07	0.04	0.03	0.02	0.03	0.89
Melon	0.85	0.10	0.05	0.02	0.02	0.03	0.85
Orange	0.92	0.04	0.04	0.01	0.01	0.02	0.92
Pear	0.87	0.08	0.05	0.02	0.02	0.03	0.87
Peach	0.88	0.09	0.03	0.01	0.02	0.01	0.88

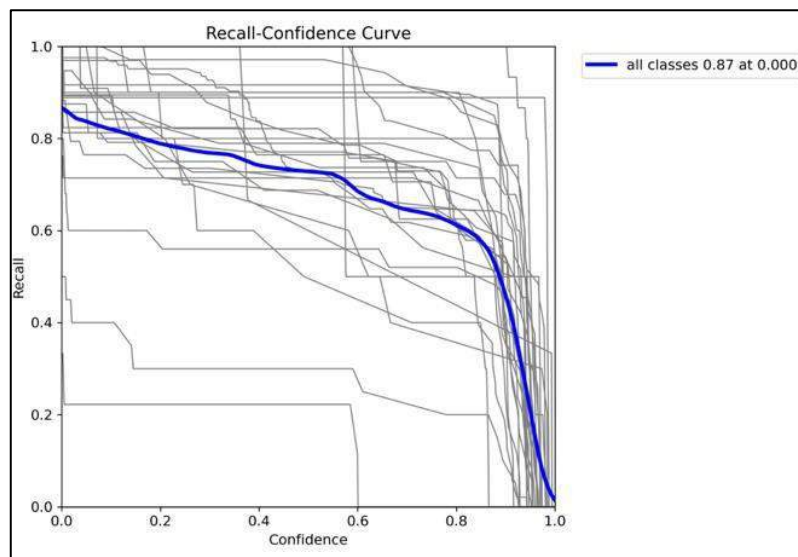
2.3 Presisi, Recall dan F1-Score Berdasarkan Tingkat Kepercayaan

Pada evaluasi pertama model YOLOv8 yang telah dilatih memiliki presisi sebesar 97% saat mencapai tingkat kepercayaan 100%. Hal ini menunjukkan pada tingkat kepercayaan yang tinggi, model memiliki tingkat akurasi yang optimal dalam mengidentifikasi dan mengklasifikasikan objek dalam semua kelas.



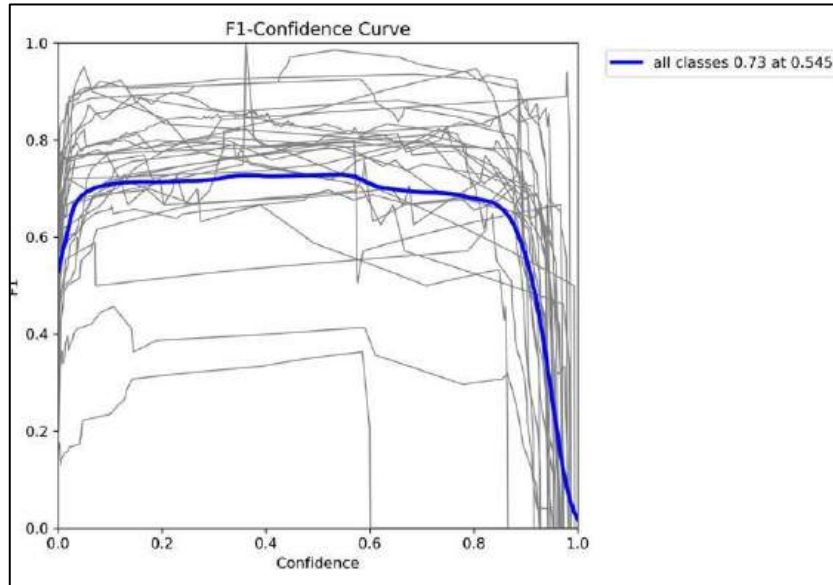
Gambar 6. Precision-Confidence Curve

Kemudian dalam tingkat kepercayaan yang paling rendah atau di 0%, model YOLOv8 memiliki *recall* atau sensitifitas deteksi sebesar 87%. Hal ini menunjukkan bahwa meskipun model memiliki tingkat keyakinan yang sangat rendah terhadap prediksinya, ia masih mampu mengidentifikasi dan mengenali kembali objek dalam semua kelas dengan akurat.



Gambar 7. Recall-Confidence Curve

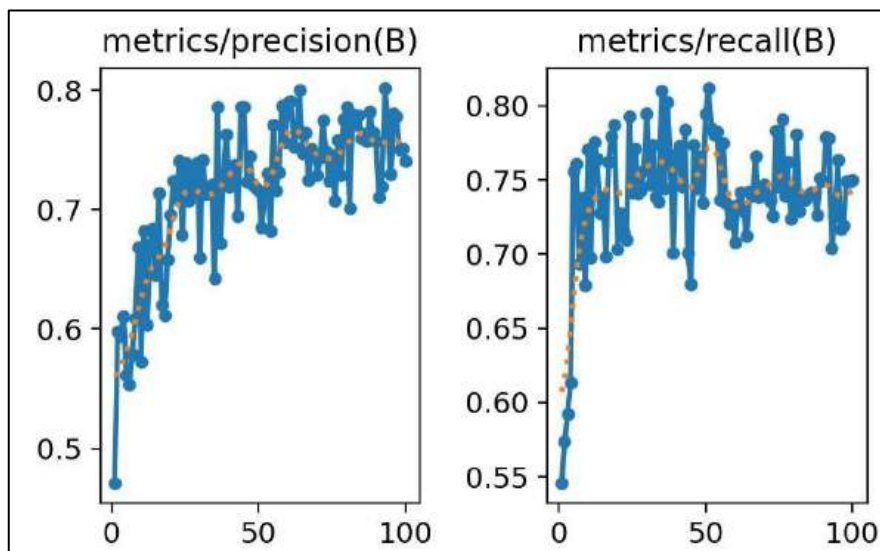
Grafik dari hasil evaluasi F1-Score pada grafik di gambar 5 menunjukkan bahwa presisi model adalah sebesar 0.73 atau 73% dengan ambang batas *threshold* diatas 50%. Hasil ini menunjukkan bahwa model memiliki keseimbangan yang cukup baik antara presisi dan *recall* pada tingkat kepercayaan diatas 50%.



Gambar 8. Hasil F1 Score

2.4 Metric Recall and Precision

Evaluasi terakir adalah dengan menghitung *metrics precision* untuk mengukur tingkat keakuratan dari prediksi positif yang dibuat oleh model deteksi objek. Semakin tinggi nilai presisi, semakin sedikit objek negatif yang salah diprediksi sebagai positif [27]. Sedangkan *metrics recall* untuk mengukur proporsi dari semua objek positif yang benar-benar diprediksi sebagai positif oleh model. Semakin tinggi nilai *recall*, semakin sedikit objek positif yang terlewatkan atau tidak terdeteksi oleh model [28]. Hasil dari evaluasi terhadap *metric recall* dan *precision* disajikan pada grafik di gambar 8.



Gambar 9. Grafik metrik presisi dan recall

Dari grafik pada gambar 9 dapat disimpulkan bahwa model YOLOv8 untuk mendeteksi kesegeran buah memiliki presisi sebesar 74% dan memiliki *recall* sebesar 75% pada parameter *epoch* ke 100 dengan 8 *batch*.