

Smart Aquaculture Vision: Deteksi dan Klasifikasi Ikan Otomatis Menggunakan YOLOv8

Riyo Irawan¹, Tresna Dewi¹, Pola Risma¹, Yurni Oktarina¹

¹Jurusan Teknik Elektro, Politeknik Negeri Sriwijaya
Palembang, Indonesia

Penulis yang berkorelas: tresna_dewi@polsri.ac.id

Abstrak— Akuakultur modern menuntut sistem pemantauan ikan yang efisien dan akurat guna meningkatkan produktivitas dan keberlanjutan. Penelitian ini mengusulkan pemanfaatan algoritma deteksi objek berbasis deep learning, yaitu YOLOv8, untuk mendeteksi dan mengklasifikasikan tiga jenis ikan secara otomatis: *Black Spotted Barb*, *Gourami*, dan *Mosquito Fish*. Dataset yang digunakan berasal dari Kaggle dan terdiri atas 730 gambar yang telah dilabeli ulang menggunakan Roboflow. Proses pelatihan dilakukan di Google Colab dengan konfigurasi GPU, batch size 32, selama 100 epoch. Model dievaluasi menggunakan metrik presisi, recall, dan mAP. Hasil evaluasi menunjukkan performa yang sangat baik dengan nilai precision sebesar 0.978, recall sebesar 0.928, mAP50 sebesar 0.973, dan mAP50-90 sebesar 0.616. Temuan ini membuktikan bahwa YOLOv8 mampu memberikan deteksi objek yang akurat dan efisien, serta berpotensi untuk diterapkan dalam sistem pemantauan akuakultur berbasis visi komputer secara real-time.

Kata kunci—Deep Learning, Artificial Intelligence, Klasifikasi, Deteksi Objek, YOLO (You Only Look Once)

Abstract— Modern aquaculture demands efficient and accurate fish monitoring systems to improve productivity and sustainability. This research proposes utilizing a deep learning-based object detection algorithm, YOLOv8, to automatically detect and classify three types of fish: *Black Spotted Barb*, *Gourami*, and *Mosquito Fish*. The dataset used comes from Kaggle and consists of 730 images that have been relabeled using Roboflow. The training process was performed on Google Colab with GPU configuration, batch size 32, for 100 epochs. The model was evaluated using precision, recall, and mAP metrics. The evaluation results showed excellent performance with a precision of 0.978, recall of 0.928, mAP50 of 0.973, and mAP50-90 of 0.616. These findings prove that YOLOv8 is capable of providing accurate and efficient object detection, and has the potential to be applied in real-time computer vision-based aquaculture monitoring systems.

Keyword—Deep Learning, Artificial Intelligence, Classification, Object Detection, YOLO (You Only Look Once)

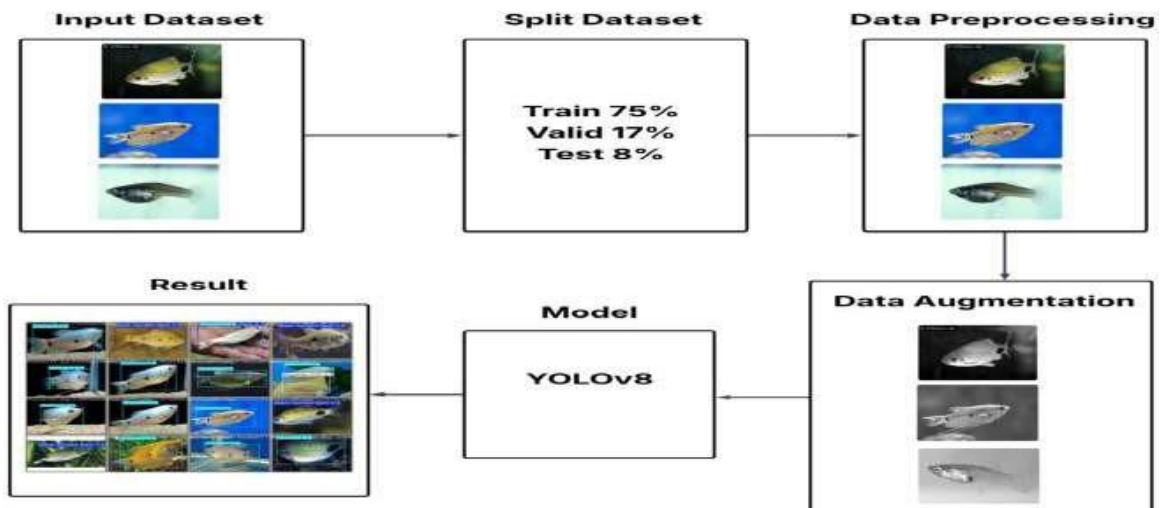
I. PENDAHULUAN

Akuakultur telah menjadi salah satu sektor pertanian dengan pertumbuhan tercepat di dunia dan memainkan peran penting dalam ketahanan pangan global. Dengan meningkatnya permintaan terhadap produk perikanan, muncul kebutuhan untuk meningkatkan efisiensi produksi melalui otomatisasi dan digitalisasi sistem budidaya [1], [2]. Salah satu tantangan utama dalam akuakultur adalah pemantauan populasi ikan, yang umumnya masih dilakukan secara manual. Proses ini tidak hanya memakan waktu dan tenaga, tetapi juga rentan terhadap kesalahan dan kurang konsisten. Oleh karena itu, integrasi teknologi visi komputer dan kecerdasan buatan (AI) menjadi solusi potensial untuk meningkatkan efisiensi dan akurasi dalam sistem monitoring akuakultur. Deteksi objek merupakan bagian penting dari visi komputer, dan dalam beberapa tahun terakhir, algoritma berbasis deep learning seperti YOLO (You Only Look Once) telah menunjukkan performa unggul dalam mendeteksi objek secara real-time. YOLOv8, sebagai iterasi terbaru, menghadirkan peningkatan dari segi arsitektur dan akurasi deteksi dibanding versi sebelumnya [3], [4].

Penelitian ini mengusulkan pemanfaatan YOLOv8 untuk mendeteksi dan mengklasifikasikan ikan secara otomatis dari citra yang diambil di lingkungan perairan budidaya. Dataset yang digunakan mencakup tiga kelas ikan, dan pelatihan dilakukan menggunakan sumber daya cloud melalui Google Colab. Diharapkan, pendekatan ini mampu menjadi solusi efektif untuk sistem pemantauan otomatis dalam akuakultur cerdas [1].

II. METODOLOGI

Penelitian ini menggunakan metode deep learning untuk klasifikasi dan mendeteksi ikan secara otomatis dalam citra digital menggunakan model YOLOv8 dengan mengikuti beberapa tahapan utama, yaitu pengumpulan data, pra pemrosesan data, data split, training model, dan model testing evaluasi [5].



Gambar 1. Pembuatan Model

A. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini diperoleh dari situs Kaggle, dengan judul *Fish Dataset* oleh Mark Daniel Lampa (Kaggle, 2022) [<https://www.kaggle.com/datasets/markdaniellampa/fish-dataset>]. Dataset ini terdiri dari 730 gambar ikan dengan beragam kondisi latar belakang, pencahayaan, serta posisi objek. Terdapat tiga kelas ikan yang digunakan dalam penelitian ini, yaitu: Black Spotted Barb, Gourami, dan Mosquito Fish [6]. Citra-citra tersebut kemudian dilabeli ulang menggunakan platform Roboflow untuk menyesuaikan dengan format dan struktur data yang dibutuhkan oleh YOLOv8. Proses pelabelan dilakukan secara manual dengan menandai bounding box pada setiap objek ikan yang terdeteksi dalam citra [7].



Gambar 2. Informasi Dataset

B. Pra Pemrosesan Data

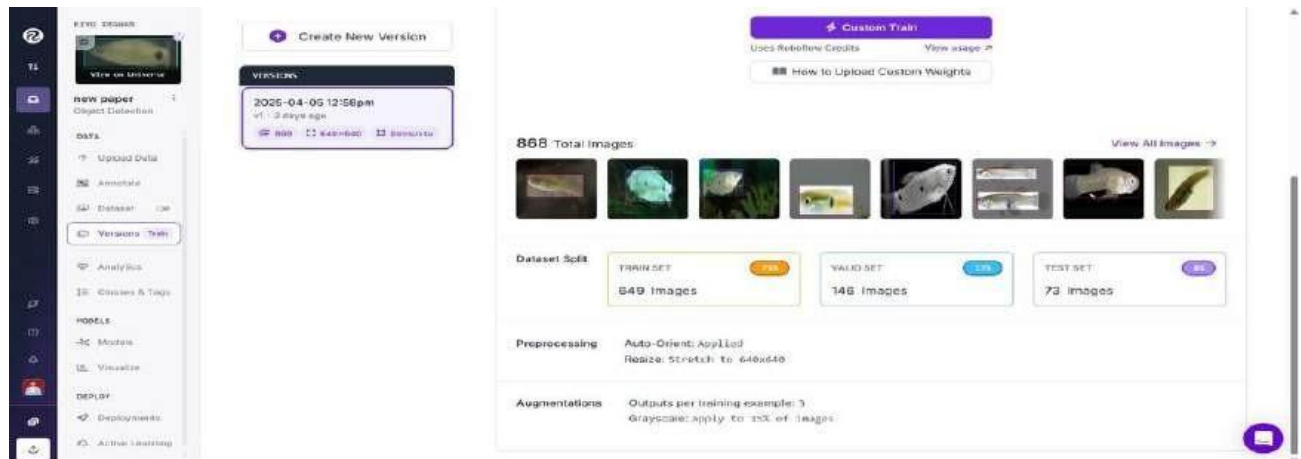
Setelah mengumpulkan dan menyortir data gambar, tahap pertama pengolahan dataset di Roboflow adalah preprocessing. Untuk memenuhi kebutuhan model, gambar 4000×3000 piksel diubah menjadi ukuran standar 640×640 piksel [8]. Labeling atau anotasi gambar dilakukan setelah preprocessing selesai. Di Roboflow, anotasi dilakukan dengan menambahkan ground truth, kotak pembatas yang menunjukkan posisi dan kategori objek dalam gambar. Sehingga model dapat memahami dan membedakan setiap kategori, setiap objek dalam dataset diberi label berdasarkan jenisnya [9], [10].

Tabel 1. Dataset Fish Informasi

Resolution	(640 px) × (640 px)	
Fish dataset (730 images)	Black Spotted Barb	212 images
	Gourami	313 images
	Mosquito Fish	280 images

C. Data Split

Setelah mendapatkan data, langkah berikutnya adalah membagi dataset ke dalam tiga kelompok utama: kelompok pelatihan, kelompok validasi, dan kelompok pengujian ini dilakukan untuk memastikan bahwa model dapat belajar dan diuji secara optimal. 75% dari total data digunakan untuk set pelatihan, yang membantu model mempelajari pola dan karakteristik objek. Selanjutnya, 17% data digunakan sebagai set validasi, yang mengevaluasi kinerja model selama proses pelatihan dan membantu menyesuaikan parameter untuk meningkatkan akurasi deteksi. 8% terakhir data diberikan sebagai set pengujian untuk menguji kinerja akhir model pada data yang belum pernah dilihat sebelumnya. Pembagian ini dilakukan untuk memastikan bahwa model tidak hanya dapat mengidentifikasi pola dari data latihan tetapi juga dapat menggeneralisasi dengan baik ke data baru yang belum pernah dipelajari [5].



Gambar 3. Detail Datasplit

D. Training Model

Pelatihan model YOLOv8 dilakukan menggunakan 868 gambar latih dengan konfigurasi batch size sebesar 32 dan jumlah epoch sebanyak 100 [8]. Waktu rata-rata pelatihan per epoch tercatat sekitar 45,25 menit. Proses pelatihan menggunakan GPU untuk mempercepat komputasi dan algoritma optimasi Adam [11], [12]. Evaluasi dilakukan secara berkala pada data validasi dengan metrik utama yaitu precision, recall, mAP50, dan mAP50-90. Model terbaik dipilih berdasarkan performa validasi dan disimpan dalam format siap implementasi [13].

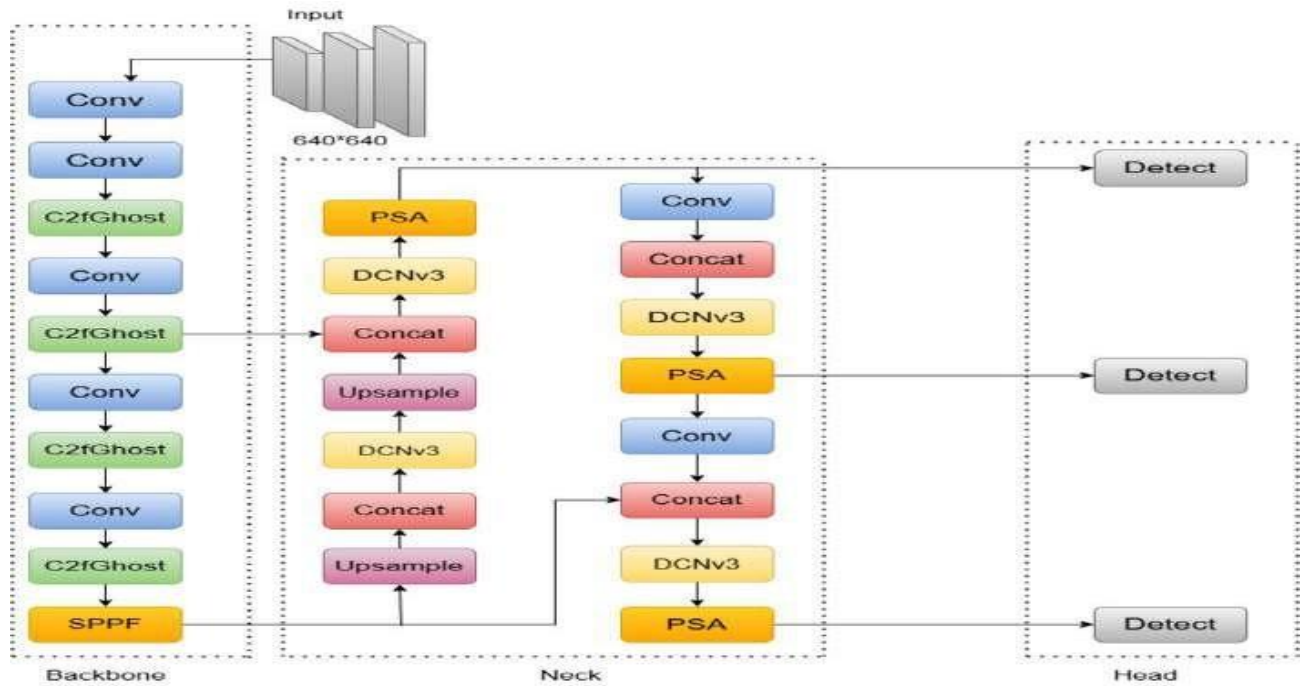
Tabel 2. Detail Training Model

Model (pt)	Image Training	Batchsize	Estimation Training(m)	Epoch
Yolov8	868	32	45.25	100

E. Architecture YOLOv8

Arsitektur yang digunakan dalam penelitian ini merupakan pengembangan dari model YOLOv8 yang terdiri atas tiga bagian utama, yaitu Backbone, Neck, dan Head [8]. Pada bagian Backbone, digunakan modul C2fGhost dan SPPF (Spatial Pyramid Pooling - Fast) untuk mengekstraksi fitur secara efisien sekaligus mempertahankan konteks spasial dari berbagai skala. C2fGhost merupakan kombinasi dari mekanisme Cross Stage Partial dan GhostNet, yang berfungsi mengurangi kompleksitas komputasi tanpa mengorbankan akurasi [7].

Pada bagian Neck, dilakukan penggabungan fitur dari berbagai skala melalui proses upsampling dan concatenation. Penambahan modul Parallel Split Attention (PSA) bertujuan untuk meningkatkan fokus model terhadap fitur penting, sementara Deformable Convolution v3 (DCNv3) memberikan fleksibilitas pada konvolusi agar lebih adaptif terhadap bentuk objek yang tidak beraturan seperti ikan. Di bagian Head, fitur-fitur ini diproses untuk menghasilkan output akhir berupa bounding box, confidence score, dan klasifikasi objek. Kombinasi komponen ini memungkinkan model bekerja secara real-time dengan akurasi tinggi pada berbagai ukuran dan orientasi objek ikan [13].



Gambar 4. Detail Arsitektur Yolov8

F. Model Testing dan Evaluasi

Beberapa metrik evaluasi utama digunakan untuk menilai kinerja model ini termasuk Intersection over Union (IoU), Precision, Recall, Mean Average Precision (mAP), dan Inference Time. IoU mengukur tingkat tumpang tindih antara bounding box prediksi dan bounding box sebenarnya (ground truth). IoU dapat dihitung dengan menggunakan persamaan berikut:

$$IoU = \frac{|R \cap R'|}{|R \cup R'|}$$

Dimana R adalah area *bounding box* prediksi dan R' adalah area *bounding box* sebenarnya. Jika $IoU \geq 0.5$, prediksi dianggap benar (*True Positive*), sedangkan jika $IoU < 0.5$, prediksi dianggap sebagai *False Positive*. Jika objek yang ada dalam *ground truth* tidak terdeteksi oleh model, maka dianggap sebagai *False Negative* [13].

Precision dan *Recall* digunakan untuk mengukur keakuratan dan kelengkapan deteksi model. *Precision* dihitung menggunakan rumus:

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

Sementara itu, *Recall* dihitung menggunakan rumus:

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

di mana TP adalah jumlah prediksi yang benar, FP adalah jumlah prediksi salah, dan FN adalah jumlah objek yang tidak terdeteksi oleh model.

Selain itu, mAP (*Mean Average Precision*) digunakan untuk mengukur rata-rata *Average Precision* (AP) dari semua kategori ikan yang terdeteksi. Perhitungan mAP dilakukan dengan rumus:

$$mAP = \frac{1}{C} \sum AP \times 100\%$$

di mana C adalah jumlah kategori ikan yang dideteksi. Semakin tinggi nilai mAP, semakin baik akurasi model dalam melakukan deteksi objek.

III. HASIL DAN PEMBAHASAN

Hasil evaluasi model YOLOv8 terhadap dataset ikan menunjukkan performa yang sangat baik dalam mendeteksi dan mengklasifikasikan objek. Berdasarkan hasil yang ditampilkan pada Tabel tersebut, model berhasil mencapai nilai Precision sebesar 0.978, yang berarti bahwa 97.8% dari semua prediksi positif merupakan deteksi yang benar (true positive). Hal ini mengindikasikan bahwa model sangat jarang memberikan deteksi palsu (false positive). Selain itu, nilai Recall sebesar 0.928 menunjukkan bahwa model mampu menemukan 92.8% dari seluruh objek ikan yang ada di dalam gambar, sehingga sangat efektif dalam menghindari kesalahan tidak terdeteksi (false negative).

Metrik mean Average Precision (mAP) juga menunjukkan performa yang unggul. Nilai mAP50 sebesar 0.973 mengindikasikan bahwa rata-rata presisi pada ambang batas IoU (Intersection over Union) sebesar 0.5 sangat tinggi, mencerminkan keberhasilan model dalam melakukan prediksi lokasi bounding box secara akurat. Sedangkan nilai mAP50-90 sebesar 0.616 memberikan gambaran bahwa performa model tetap cukup stabil meskipun ambang batas IoU diperketat, yang menunjukkan ketahanan model terhadap variasi posisi dan bentuk objek ikan. Secara keseluruhan, hasil evaluasi ini membuktikan bahwa YOLOv8 sangat layak digunakan untuk sistem deteksi dan klasifikasi ikan otomatis dalam lingkungan akuakultur pintar.

Tabel 3. Hasil Model YOLOv8

Model	Precision	Recall	mAP50	mAP50-90
Yolov8	0.978	0.928	0.973	0.616

Gambar 5 menunjukkan hasil deteksi objek ikan menggunakan model YOLOv8 yang telah dilatih. Setiap kotak biru menunjukkan deteksi objek dengan label nama spesies ikan dan nilai confidence (tingkat keyakinan model terhadap prediksi tersebut). Dua jenis ikan yang terdeteksi dalam gambar adalah Black Spotted Barb dan Gourami, dengan nilai confidence berkisar antara 0.5 hingga 0.99, menandakan tingkat akurasi yang tinggi dalam identifikasi objek.

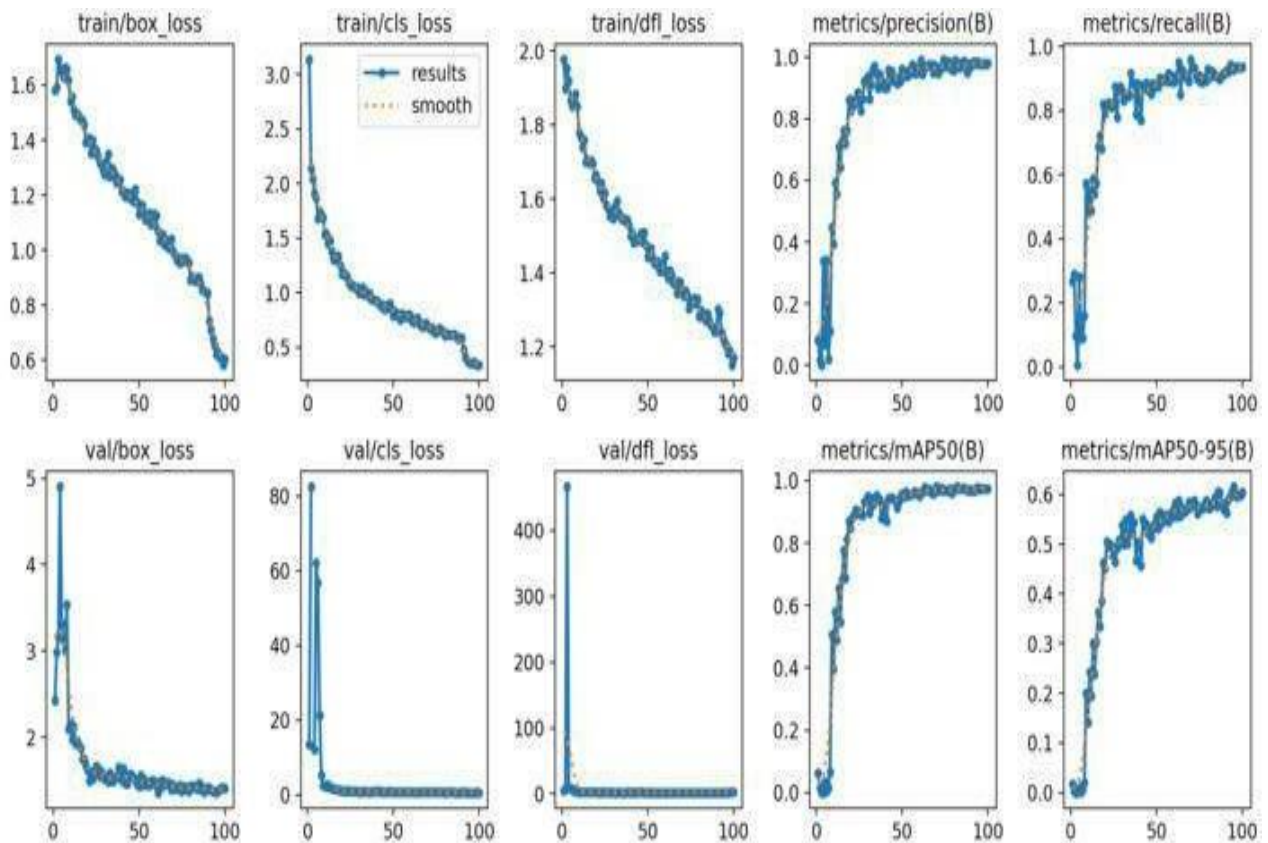
Hasil visualisasi ini membuktikan bahwa model mampu mengenali ikan dengan berbagai posisi, ukuran, dan latar belakang yang berbeda. Deteksi yang konsisten pada spesies yang sama menunjukkan bahwa model telah berhasil mempelajari fitur visual khas dari masing-masing jenis ikan. Hal ini mendukung efektivitas model dalam tugas klasifikasi objek secara real-time pada lingkungan yang bervariasi.



Gambar 5. Hasil Model YOLOv8

Gambar 6 memperlihatkan hasil visualisasi pelatihan dan validasi model YOLOv8 selama 100 epoch. Tiga metrik utama pada proses pelatihan—box loss, classification loss, dan distribution focal loss—mengalami penurunan secara konsisten, menandakan bahwa model berhasil meminimalkan kesalahan deteksi objek. Selain itu, nilai precision dan recall meningkat seiring waktu dan menunjukkan konvergensi yang stabil, mengindikasikan kemampuan model dalam mendeteksi objek dengan akurasi yang baik.

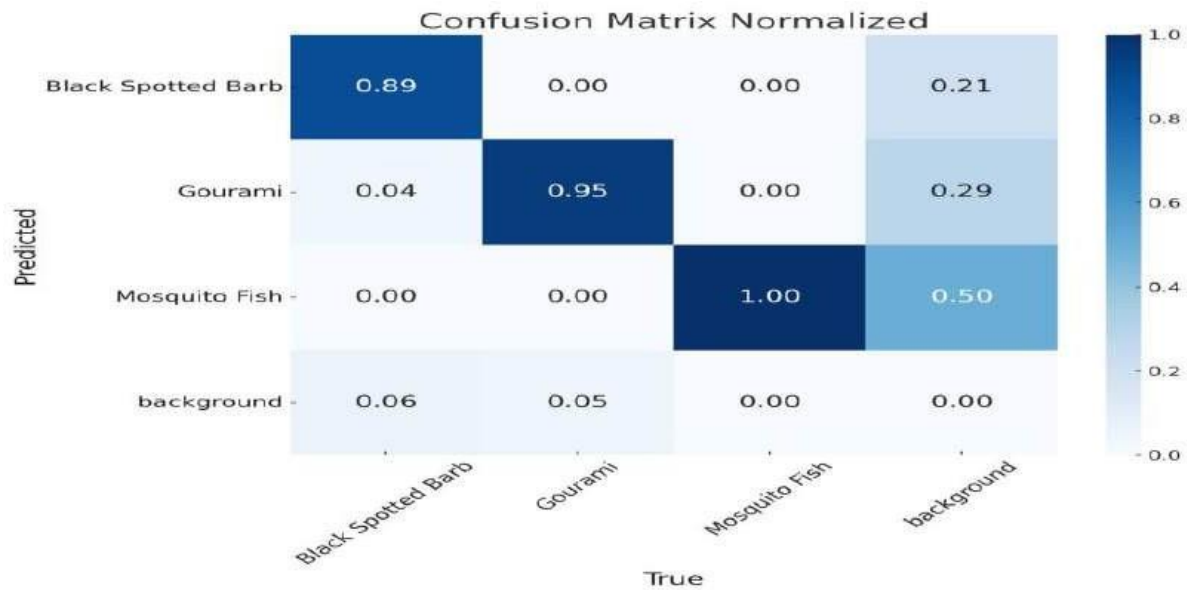
Pada proses validasi, nilai loss menunjukkan tren penurunan meskipun terdapat fluktuasi awal, yang mencerminkan proses adaptasi terhadap data uji. Sementara itu, nilai mAP50 meningkat tajam hingga mendekati 1.0 dan mAP50-95 stabil di sekitar 0.6, memperlihatkan bahwa model mampu mengenali objek dengan presisi tinggi pada berbagai tingkat threshold. Secara keseluruhan, tren grafik menunjukkan bahwa pelatihan berjalan optimal tanpa indikasi overfitting yang signifikan.



Gambar 6. Grafik Hasil Training YOLOv8

Gambar 7 menunjukkan performa model YOLOv8 dalam mengklasifikasikan tiga jenis ikan: Black Spotted Barb, Gourami, dan Mosquito Fish. Matriks ini telah dinormalisasi untuk mempermudah interpretasi proporsi klasifikasi. Hasilnya menunjukkan bahwa model memiliki akurasi sangat tinggi, terutama pada kelas Mosquito Fish yang berhasil diklasifikasikan dengan benar sebesar 100%. Gourami dan Black Spotted Barb juga menunjukkan performa tinggi dengan akurasi masing-masing 0.95 dan 0.89.

Meskipun akurasi keseluruhan tinggi, masih ditemukan sejumlah kecil kesalahan klasifikasi, terutama pada deteksi latar belakang (background), di mana sebagian objek Black Spotted Barb dan Gourami salah dikenali. Hal ini kemungkinan disebabkan oleh tumpang tindih objek atau kemiripan visual dengan latar belakang. Secara keseluruhan, confusion matrix ini mengindikasikan bahwa model telah bekerja dengan sangat baik, dengan potensi peningkatan di aspek deteksi yang lebih presisi terhadap latar dan objek serupa.



Gambar 7. Confusion Matrix

IV. KESIMPULAN

Berdasarkan hasil eksperimen dan evaluasi performa, dapat disimpulkan bahwa algoritma YOLOv8 menunjukkan kapabilitas yang sangat baik dalam tugas deteksi dan klasifikasi ikan secara otomatis. Model berhasil mendeteksi tiga kelas ikan dengan tingkat akurasi yang tinggi, dibuktikan oleh nilai precision sebesar 0.978, recall sebesar 0.928, serta mAP50 sebesar 0.973 dan mAP50-90 sebesar 0.616. Selain itu, grafik pelatihan menunjukkan konvergensi model yang stabil dan tidak mengalami overfitting secara signifikan. Dengan demikian, YOLOv8 sangat potensial untuk diimplementasikan dalam sistem akuakultur cerdas guna meningkatkan efisiensi pemantauan populasi ikan secara real-time.

REFERENSI

- [1] Jocher, G., Chaurasia, A., Qiu, J., & Stoken, A. (2023). YOLO by Ultralytics. *GitHub repository*. <https://github.com/ultralytics/ultralytics>
- [2] Han, K., Wang, Y., Tian, Q., Guo, J., Xu, C., & Xu, C. (2020). GhostNet: More Features from Cheap Operations. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 1580–1589. <https://doi.org/10.1109/CVPR42600.2020.00165>
- [3] Zhu, X., Hu, H., Lin, S., & Dai, J. (2022). Deformable ConvNets v3: Detecting Objects with Deeper ConvNets and Deformable Modules. *arXiv preprint arXiv:2304.07154*. <https://arxiv.org/abs/2304.07154>
- [4] Zhang, C., Li, X., & Yang, J. (2021). PSA: A Novel Pyramid Split Attention Module for Effective Feature Extraction. *Pattern Recognition Letters*, 146, 67–73. <https://doi.org/10.1016/j.patrec.2021.03.008>
- [5] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint arXiv:2004.10934*. <https://arxiv.org/abs/2004.10934>
- [6] Lampa, M. D. (2022). *Fish Dataset*. Kaggle. <https://www.kaggle.com/datasets/markdaniellampa/fish-dataset>
- [7] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint arXiv:2004.10934*.
- [8] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal Loss for Dense Object Detection. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2980–2988. <https://doi.org/10.1109/ICCV.2017.324>
- [9] Lin, T. Y., et al. (2017). Feature Pyramid Networks for Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2117–2125. <https://doi.org/10.1109/CVPR.2017.106>
- [10] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Advances in Neural Information Processing Systems (NIPS)*, 91–99.
- [11] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- [12] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556*.

- [13] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *You Only Look Once: Unified, Real-Time Object Detection*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779–788. <https://doi.org/10.1109/CVPR.2016.91>
- [14] Jocher, G., et al. (2023). *YOLOv8: Ultralytics Documentation*. <https://docs.ultralytics.com/>
- [15] Mark Daniel Lampa. (2022). *Fish Dataset*. Kaggle. <https://www.kaggle.com/datasets/markdaniellampa/fish-dataset>
- [16] Nasir, M., et al. (2020). *Fish detection and tracking using computer vision techniques: A review*. Aquacultural Engineering, 90, 102088. <https://doi.org/10.1016/j.aquaeng.2020.102088>
- [17] Zhang, C., Li, X., & Yang, J. (2021). *PSA: A Novel Pyramid Split Attention Module for Effective Feature Extraction*. Pattern Recognition Letters, 146, 67–73. <https://doi.org/10.1016/j.patrec.2021.03.008>
- [18] D. C. Love, J. P. Fry, X. M. Li, E. S. Hill, L. Genello, K. Semmens, and R. E. Thompson, "Commercial aquaponics production and profitability: Findings from an international survey," *Aquaculture*, vol. 435, pp. 67–74, 2015, doi: [10.1016/j.aquaculture.2014.09.023](https://doi.org/10.1016/j.aquaculture.2014.09.023).
- [19] P. Domenici, "Context-dependent variability in the components of fish escape response: Integrating locomotor performance and behavior," *J. Exp. Zool. Part A: Ecol. Integr. Physiol.*, vol. 313, no. 2, pp. 59–79, 2010, doi: [10.1002/jez.580](https://doi.org/10.1002/jez.580).
- [20] J. Y. Xu, Y. Liu, S. R. Cui, and X. W. Miao, "Behavioral responses of tilapia (*Oreochromis niloticus*) to acute fluctuations in dissolved oxygen levels as monitored by computer vision," *Aquacult. Eng.*, vol. 35, no. 3, pp. 207–217, 2006, doi: [10.1016/j.aquaeng.2006.02.004](https://doi.org/10.1016/j.aquaeng.2006.02.004).