Comparative Analysis of YOLO Model Performance for Freshwater Fish Detection

Vrica Gede Penggalih*[1], Esi Putri Silmina[2]

[1]Department of Information Technology, Faculty of Science and Technology, Aisyiyah University Yogyakarta, Yogyakarta, Indonesia

[2]Department of Information Technology, Faculty of Science and Technology, Aisyiyah University Yogyakarta, Yogyakarta, Indonesia

E-mail: vricagedepenggalih@gmail.com[1], esiputrisilmina@unisayogya.ac.id [2]

Abstract. Freshwater fish are an affordable source of animal protein that plays an important role in improving nutritional intake, particularly in developing countries such as Indonesia. However, accurate identification of freshwater fish species remains a challenge for the general public and small-scale farmers. Recent advances in deep learning–based object detection, especially the You Only Look Once (YOLO) family, provide a promising solution to this problem. This study presents a comparative performance analysis of lightweight YOLOv11n and YOLOv12n, for detecting five commonly consumed freshwater fish species in Indonesia: bandeng, gurame, lele, nila, and patin. A partial factorial experimental design was employed to evaluate the effect of selected hyperparameters, including epochs, image size, and optimizer, in order to determine the optimal configuration for each model. The models were trained using the same dataset and subsequently evaluated on test set using precision, recall, mAP@50, and F1-score metrics. Results show that YOLOv12n outperforms YOLOv11n in terms of overall detection accuracy, achieving a higher mAP@50 (0.808) and F1-score (0.779), while YOLOv11n demonstrates slightly higher recall (0.739) and faster inference speed. Further analysis using Precision–Recall curves and normalized confusion matrices confirms that YOLOv12n provides more stable predictions and lower background misclassification across most classes. These findings indicate that YOLOv12n is more suitable for applications prioritizing detection accuracy, whereas YOLOv11n offers a better trade-off for real-time systems with limited computational resources. This study contributes empirical insights into the selection of lightweight YOLO models for freshwater fish detection and supports their potential application in nutrition education and fisheries-related decision support systems.

Keywords: yolov11, yolov12, freshwater fish detection, partial factorial experimental, nutrition literacy.

1. Introduction

Indonesia still faces major challenges in terms of child nutrition [1]. Data from the 2024 Indonesian Nutrition Status Survey (SSGI) shows that the national prevalence of stunting reached around 19.8% in 2024 [2]. Although this represents a downward trend compared to the previous year, the figure is still far from the national target of 14.2% set in the 2024-2029 National Medium-Term Development Plan (RPJMN) [3]. One of the causes of stunting is low nutrition literacy [4]. This condition underscores that nutrition interventions are not only important in the short term, but also crucial for the long-term development of the nation.

Meanwhile, freshwater fish such as Lele, Patin, Nila, Bandeng, and Gurame are relatively affordable local commodities that are easily accessible to the Indonesian people [5], [6]. The use of freshwater fish as a source of animal protein can be an important strategy in the context of improving nutritional literacy and empowering local communities [7], [8]. However, in practice, there are obstacles in quickly and accurately identifying fish species by consumers or farmers, including in selecting fish that are highly nutritious and adequate for children's nutritional intake [9].

Object detection or automatic object detection based on digital images offers great potential to support nutrition literacy and public education [10]. In particular, the "You Only Look Once" framework of the YOLOv11 and YOLOv12 model families has shown significant improvements in terms of object detection accuracy and speed [11]. For example, a review study of the YOLO model series states that YOLOv12 is capable of providing a higher mAP with lower latency compared to previous versions [12]. This opens up the possibility that object detection technology can be applied to detect freshwater fish species from digital images, then integrated into educational applications or systems that help the public choose fish correctly.

Research on fish detection using YOLO-based models has been conducted extensively in recent years. The YOLO model was chosen because it provides better detection accuracy than other detection models, such as RF-DETR and Faster R-CNN [13], [14]. Justam et al. [14] reported that the YOLO model outperformed Faster R-CNN on the fish dataset (18 classes) with higher precision, recall, and accuracy metrics and a significantly lower weight estimation error (MAPE), demonstrating YOLO's superiority in terms of both detection accuracy and consistency of size/weight estimation.

A local comparative study, Putra et al. [15] found YOLOv11 to be more efficient in terms of memory usage and inference time, while YOLOv12 showed superior accuracy in several scenarios, results consistent with architectural benchmarks reporting an increase in YOLOv12 mAP while maintaining competitive speed. Zhang et al. [16] developed the YOLOv8 model, which is optimized for underwater fish detection, and the results showed a significant improvement in performance compared to YOLOv5 in low-light conditions and complex backgrounds. These studies confirm that the YOLO model is effective in the field of fisheries, but most of them still focus on marine fish.

To date, there have been few studies comparing the performance of the YOLOv11 and YOLOv12 models in the context of freshwater fish, especially Indonesian local species such as Lele, Patin, Nila, Bandeng, and Gurame. This is an important gap, considering that freshwater fish are an affordable source of protein that plays a strategic role in efforts to reduce stunting rates in Indonesia. Therefore, this study offers something new by conducting a comparative evaluation of the performance of the YOLOv11n and YOLOv12n models in detecting five types of local freshwater fish: Lele, Patin, Nila, Bandeng, and Gurame. Variation *n* was selected because, for each experimental trial (each hyperparameter configuration), both training and validation execute substantially faster on model *n* [17]. This efficiency enables a greater number of experimental runs within the available time and GPU resource constraints (Google Colab). Research conducted on the use of this technology is expected to be meaningful in reducing stunting rates, especially in Indonesia.

2. Research Method

This study uses a quantitative experimental approach with the aim of analyzing and comparing the performance of two YOLO-based object detection models, namely YOLOv11n and YOLOv12n, in detecting five types of local Indonesian freshwater fish. This approach was chosen because it allows for measurable evaluation of the accuracy, computational efficiency, and generalization ability of the models on a dataset of fish images that are homogeneous in domain but diverse in visual appearance.

2.1. Research Design

The research design consists of six main stages, namely: (1) dataset collection, (2) dataset annotation, (3) dataset splitting, (4) pre-processing & augmentation (5) training of the YOLOv11n and YOLOv12n models, (6) analysis of results and interpretation. The overall flowchart of the research can be seen in Figure 1.

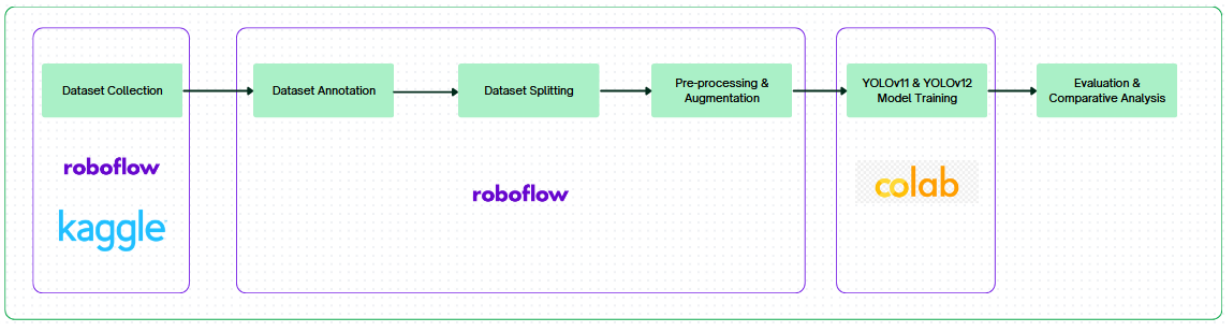

Figure 1. Research Stages

2.2. Dataset Collection

The dataset used in this study consisted of 1000 freshwater fish images, covering five popular fish species in Indonesia, namely: Lele (*Clarias batrachus*), Patin (*Pangasius hypophthalmus*), Nila (*Oreochromis niloticus*), Bandeng (*Chanos chanos*), and Gurame (*Osphronemus goramy*) [18]. Each fish class was taken from various public sources (Roboflow, Kaggle) with consideration given to variations in lighting, shooting angle, background, and object position. This diversity was intended to create a dataset that was representative of real-world conditions in the field, such as in fish markets or aquaculture environments. All image data were converted to .jpg format to maintain model input consistency. Table 1 shows the distribution of the research dataset.

Table 1. Research Datasets

Class	Count
Lele	200
Patin	200
Gurame	200
Nila	200
Bandeng	200

2.3. Dataset Annotation

The entire object annotation process was carried out using the Roboflow platform, which provides a web-based interface for data labeling. Each image was manually given a bounding box according to the area of the fish's body and labeled according to its species. Manual annotation was chosen to ensure high label accuracy, which is an important factor in the quality of object detection model training. Annotation errors can cause the model to learn incorrect features and reduce detection performance. Figure 2 shows the dataset annotation process in Roboflow.

Figure 2. Annotation Process in Roboflow

2.4. Dataset Splitting

After the annotation process is complete, the dataset is divided into three subsets with a ratio of 70% for training data (training set), 20% for validation data (validation set), and 10% for testing data (testing set). This proportion allows the model to have enough data to learn visual patterns during the training stage, while the validation data is used to assess the model's ability to generalize to new data during the training process. The testing data is used only at the final stage to measure the actual performance of the model after training is complete.

2.5. Pre-processing and Augmentation Dataset

The pre-processing stage is carried out in Roboflow and consists of two main steps: image standardization and normalization. All images are resized to 640×640 pixels, in addition, auto-orientation is performed to ensure uniform image orientation, avoiding differences in rotation angles that can affect the feature extraction process.

To expand data variation and improve model generalization, data augmentation is performed using several transformation techniques. Table 2 shows a list of augmentation techniques applied to the dataset and their purposes.

Table 2. Augmentation Technique Applied to Dataset

Augmentation Technique	**Purpose**
Horizontal flip	Creating variations in fish orientation
Rotasi -15° to +15°	Simulating variations in the angle of view
Saturation-15% to+15% dan brightness -15% to +15%	Anticipating variations in natural light intensity
Exposure -10% to +10%	Enriching lighting level variation
Blur up to 2 pixel	Training the model still enables it to detect objects in images with low sharpness

The augmentation techniques in Table 2 were selected based on common practices in object detection research to improve model robustness without causing excessive distortion to the shape of the fish [19] and to avoid model overfitting [20].

2.6. Model Training

The training process was conducted using Google Colab with a GPU runtime configuration (NVIDIA Tesla T4) to ensure efficient computation. The augmented dataset was imported into the Colab environment to be trained using both models, YOLOv11n and YOLOv12n, separately but with identical parameters. The training parameters set can be seen in Table 3. The main objective of this stage was to obtain the best performance from each model by exploring several hyperparameter configurations that were considered to have the most significant impact on detection quality.

Table 3. Training Hyperparameter

HYPERPARAMETER	VALUE
Epoch	100 vs 150
Batch Size	16
Image Size (imgsz)	640 vs 768
Optimizer	AdamW vs SGD

This study uses a partial factorial experimental design method, in which each hyperparameter is tested separately while holding the other hyperparameters constant (One-Variable-at-a-Time). We first determine the baseline configuration, which is: img=640; epoch=100; optimizer=AdamW; batch=16. We will then test the hyperparameters one by one, with the other hyperparameters remaining the same as the baseline. Hyperparameters that show improved performance will be combined, and the resulting model will be evaluated and analyzed. This approach was chosen to maintain scientific validity, reduce computational costs, and facilitate the analysis of the influence of each hyperparameter.

2.7. Evaluation and Comparative Analysis

The evaluation stage was conducted to assess and compare the performance of the YOLOv11n and YOLOv12n models in detecting freshwater fish based on digital images. The evaluation was carried out on a testing set subset (10% of the total data) that was never used during training or validation, so that the test results objectively represented the model's generalization ability to new data.

Performance was assessed using four main evaluation metrics commonly used in object detection research, namely Precision (P), Recall (R), mAP@0.5 (mAP@0.5), and F1-Score. These four metrics provide a comprehensive overview of how well the model is able to detect objects accurately and consistently.

a. Precision (P)

Precision measures the accuracy of a model in detecting objects, i.e., how many predictions are correct (true positives) compared to all detections made by the model, including false positives. The Precision formula is shown in equation (1).

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

with:

TP (True Positive) = number of correct detections, i.e., fish objects that are correctly detected according to class and position.

FP (False Positive) = number of false detections, i.e., predictions that do not match the actual object (for example, the model detects “lele” when the object is actually “gurame”).

High precision indicates that the model rarely makes false detections or classification errors. In the context of this study, a high Precision value means that the model is able to identify fish with a low classification error rate.

b. Recall (R)

Recall measures the extent to which the model is able to find all objects that actually exist in the image. In other words, Recall measures the model's ability to find all fish that appear in the image. The Recall formula is expressed in equation (2).

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

with:
FN (False Negative) = the number of objects that should have been detected but were not detected by the model.
A high Recall indicates that the model has good sensitivity to objects, meaning that almost all fish in the image are successfully detected. In the context of this research, high Recall is important to ensure that no fish are "missed" by the detection system, which is particularly relevant for applications such as monitoring catches or classifying fish in the market.

c. Mean Average Precision (mAP@0.5)
Mean Average Precision (mAP) is the main metric in object detection tasks because it combines Precision and Recall into a single overall performance measure. mAP is calculated based on the Average Precision (AP) value of each object class, then averaged.
The AP calculation is based on the area under the Precision–Recall (PR) curve. The value of mAP@0.5 indicates that the evaluation was performed with an Intersection over Union (IoU) threshold of 0.5, meaning that predictions are considered correct if the overlap between the predicted bounding box and the ground truth is greater than or equal to 50%. The general formula for mAP is shown in equation (3).

$$mAP = \frac{1}{N}\sum_{i=1}^{N} APi \tag{3}$$

with:
N = number of classes (in this study N = 5 types of fish),
AP_i = Average Precision value for class i.
A high mAP@0.5 value indicates that the model has a good balance between Precision and Recall across all fish classes. This metric is often used as a key indicator of the success of an object detection model because it considers both the accuracy and completeness of detection.

d. F1-Score
F1-Score is a measure of the balance between Precision and Recall. This value is used to evaluate the overall performance of a model in a single number. The F1-Score formula is shown in equation (4).

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

The F1-Score is high if both Precision and Recall are high, and low if either one is low. In this study, the F1-Score is used to provide an overview of the balance between the model's ability to detect fish correctly (without classification errors) and ensuring that all fish are successfully detected.

3. Results and Discussion

This section presents the experimental results and discussion of YOLO-based object detection models for freshwater fish detection. Two recent YOLO variants, YOLOv11n and YOLOv12n, were evaluated to identify the model configuration that achieves the best detection performance under controlled experimental settings.

The experiments were conducted using a partial factorial experimental design, where one hyperparameter was varied at a time while the remaining parameters were fixed at a predefined baseline configuration. This approach enables a controlled analysis of the individual impact of each hyperparameter while maintaining computational efficiency. The evaluated hyperparameters include the number of training epochs, input image size, and optimizer type.

Model performance was assessed using four evaluation metrics: precision, recall, mean Average Precision at IoU 0.5 (mAP@50), and F1-score. These metrics were selected to provide a comprehensive evaluation of detection accuracy, detection completeness, and the balance between precision and recall. In the context of freshwater fish detection, recall and F1-score are particularly important, as missed detections may limit the practical applicability of the system.

The discussion is structured as follows. First, the performance of YOLOv11n is analyzed to examine its sensitivity to hyperparameter variations and to determine its optimal configuration. Next, a similar analysis is conducted for YOLOv12n to evaluate its robustness and stability. Finally, a comparative analysis between YOLOv11n and YOLOv12n is presented to highlight their respective strengths and limitations in freshwater fish detection tasks. Table 4 shows the results of the partial factorial experimental design on the both YOLOv11n dan YOLOv12n.

Table 4. Partial Factorial Experimental Results

	Baseline	**Epoch 150**	**Image Size 768**	**SGD**
Yolov11n	P: 0.85 R: 0.743 mAP50: 0.819 F1-score: 0.792	P: 0.816 R: 0.774 mAP50: 0.816 F1-score: 0.794	P: 0.851 R: 0.769 mAP50: 0.831 F1-score: 0.807	P: 0.865 R: 0.734 mAP50: 0.816 F1-score: 0.790
Yolov12n	P: 0.836 R: 0.783 mAP50: 0.832 F1-score: 0.808	P: 0.835 R: 0.756 mAP50: 0.812 F1-score: 0.793	P: 0.811 R: 0.787 mAP50: 0.817 F1-score: 0.798	P: 0.776 R: 0.769 mAP50: 0.819 F1-score: 0.772

3.1. YOLOv11n Results Discussion

The performance of YOLOv11n was evaluated using a partial factorial experimental design by varying one hyperparameter at a time while keeping the others at baseline values. The baseline configuration (epoch = 100, image size = 640, optimizer = AdamW) achieved a precision of 0.85, recall of 0.743, mAP@50 of 0.819, and F1-score of 0.792, serving as a reference for subsequent comparisons.

Increasing the number of training epochs to 150 resulted in improved recall (0.774) and a slight increase in F1-score (0.794), indicating better detection completeness. However, this improvement was accompanied by reductions in precision (0.816) and mAP@50 (0.816), suggesting a trade-off between detecting more objects and maintaining localization accuracy. Increasing the input image size to 768 yielded the best overall performance among all configurations. This setting improved recall to 0.769 and increased mAP@50 to 0.831, while maintaining high precision (0.851). Consequently, the F1-score increased to 0.807, indicating a more balanced detection performance. These results suggest that higher-resolution inputs enhance feature representation and object localization for YOLOv11n. Replacing AdamW with the

SGD optimizer increased precision to 0.865 but reduced recall to 0.734 and slightly lowered mAP@50 and F1-score. This behavior indicates that while SGD improves prediction confidence, it negatively impacts detection completeness.

Based on these results, the configuration with image size 768 and AdamW optimizer was selected as the final YOLOv11n model, as it provides the best trade-off between precision, recall, and overall detection accuracy.

3.2. YOLOv12n Results Discussion

YOLOv12n was evaluated using the same experimental design and baseline hyperparameter configuration as YOLOv11n to ensure a fair comparison. The baseline model achieved a precision of 0.836, recall of 0.783, mAP@50 of 0.832, and F1-score of 0.808, indicating strong initial performance, particularly in detection completeness.

Increasing the number of training epochs to 150 did not improve performance. Although precision remained stable (0.835), recall decreased to 0.756, leading to reductions in mAP@50 (0.812) and F1-score (0.793). This suggests that extended training may cause overfitting for YOLOv12n, especially when applied to the given dataset. Increasing the input image size to 768 resulted in a slight increase in recall (0.787), but this was accompanied by a decrease in precision (0.811) and mAP@50 (0.817). Consequently, the F1-score declined to 0.798, indicating that higher-resolution inputs did not consistently benefit YOLOv12n and may introduce additional noise or instability. Replacing AdamW with the SGD optimizer led to the lowest overall performance among the tested configurations. Precision decreased to 0.776, recall dropped to 0.769, and the F1-score fell to 0.772, demonstrating that SGD is less suitable for training YOLOv12n on this dataset.

Overall, the baseline configuration provided the best balance across all evaluation metrics for YOLOv12n. Unlike YOLOv11n, YOLOv12n showed limited sensitivity to hyperparameter modifications, suggesting that its default configuration is already well-optimized for lightweight object detection tasks.

3.3. Evaluation and Comparative Analysis

The final hyperparameter that was chosen is (epoch: 100, batch size: 16, imgsz: 768, optimizer AdamW) for YOLOv11 and (epoch: 100, batch size: 16, imgsz: 640, optimizer AdamW) for YOLOv12. Table 5 shows the final results of both models with test set.

Table 5. Final Model Performance Results with Test Dataset

	Precision	**Recall**	**mAP50**	**F1-score**	**Inference Speed**
YOLOv11n	0.789	0.739	0.781	0.763	7.5ms
YOLOv12n	0.844	0.725	0.808	0.779	10.8ms

The final test results on the test dataset show that both models are capable of detecting five types of freshwater fish with competitive performance, but with different characteristics. Overall, YOLOv12n shows superior detection performance compared to YOLOv11n, particularly in terms of precision, mAP@50, and F1-score metrics. YOLOv12n achieved a precision of 0.844, mAP@50 of 0.808, and an F1-score of 0.779, higher than YOLOv11n, which achieved 0.789, 0.781, and 0.763, respectively. Conversely, YOLOv11n has a slightly higher recall value (0.739) than YOLOv12n (0.725), indicating a tendency for YOLOv11n to detect more objects with a greater risk of false positives. In terms of computational efficiency, YOLOv11n is also faster with an average inference time of 7.5 ms per image, while YOLOv12n requires 10.8 ms per image, indicating a trade-off between accuracy and inference speed.

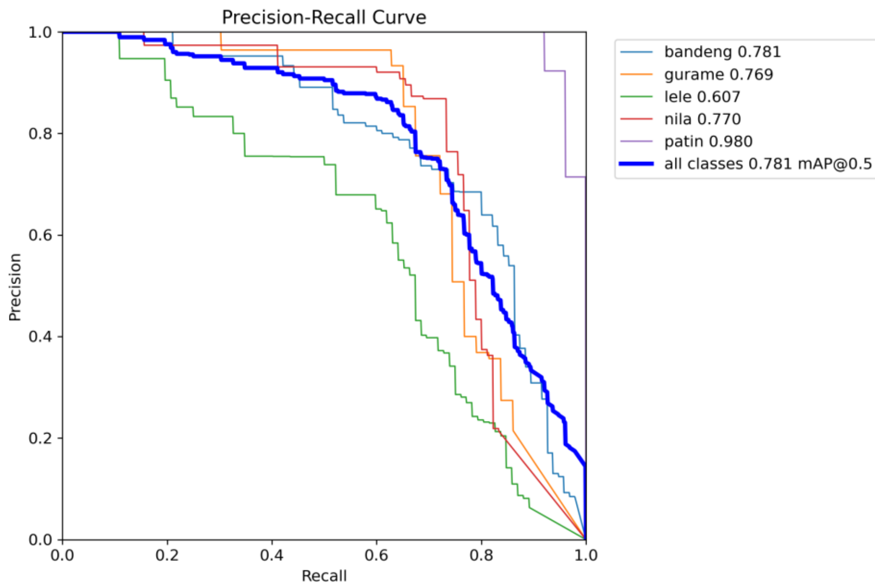

Figure 3. YOLOv11n PR Curve

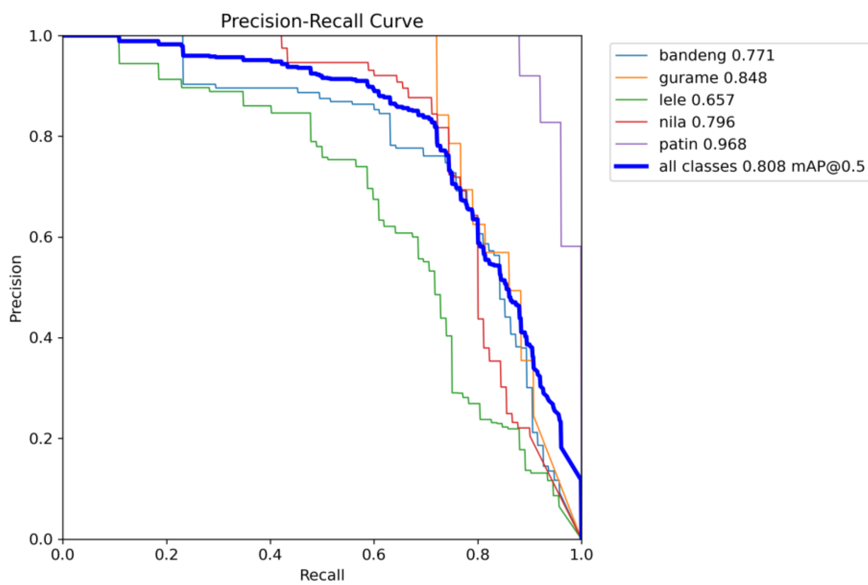

Figure 4. YOLOv12n PR Curve

Based on the Precision–Recall (PR) curve, YOLOv12n (Figure 4) has a larger area under the curve (AUC), which is reflected in an overall mAP@50 value of 0.808, higher than the 0.781 value for YOLOv11n (Figure 3). This shows that YOLOv12n is able to maintain higher precision at various recall levels. Class-by-class analysis shows that the patin class performs best on both models, with very high AP values (YOLOv11n: 0.980, YOLOv12n: 0.968), indicating that the visual characteristics of this class are relatively easy to distinguish. In contrast, the lele class consistently performed the worst on both models (YOLOv11n: 0.607, YOLOv12n: 0.657), which may be due to its visual similarity to the background and the variation in object shape and orientation. In general, the PR curve of YOLOv12n is more stable and above YOLOv11n in most recall ranges, confirming the superiority of YOLOv12n in maintaining a balance between precision and recall.

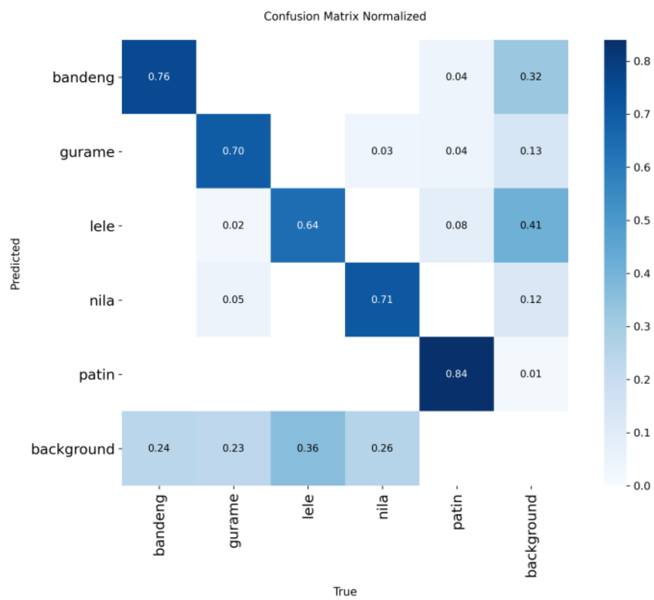

Figure 5. YOLOv11n Normalized Confusion Matrix

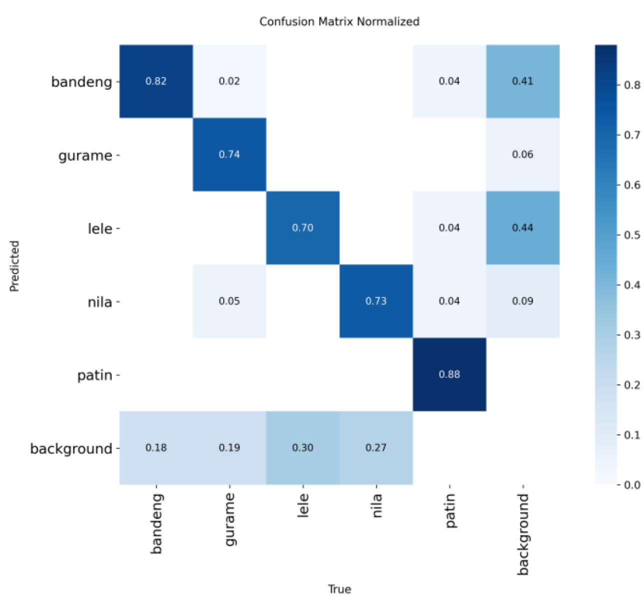

Figure 6. YOLOv12n Normalized Confusion Matrix

The normalized confusion matrix results show that YOLOv12n (Figure 6) has a higher true positive rate in almost all classes compared to YOLOv11n (Figure 5). For example, the diagonal value for the bandeng class increased from 0.76 (YOLOv11n) to 0.82 (YOLOv12n), the lele class from 0.64 to 0.70, and the patin class from 0.84 to 0.88. In addition, YOLOv12n shows a decrease in the proportion of classification errors to the background, indicating better object localization capabilities. However, both models still show a relatively high level of confusion between fish objects and the background, especially in the lele and nila classes, indicating visual challenges due to lighting, water texture, and the similarity of object colors to the surrounding environment.

3.4. Implications

Based on all evaluation results, it can be concluded that YOLOv12n excels in terms of prediction accuracy and stability, while YOLOv11n is more efficient in terms of inference speed. Thus, the choice of model depends heavily on the application requirements. YOLOv12n is more suitable for systems that prioritize accuracy, such as educational applications and decision support systems related to the selection of nutritious fish, while YOLOv11n is more suitable for real-time applications with limited computing resources.

These results confirm that the architectural improvements in YOLOv12 enhance detection performance without significant changes to the training pipeline, and reinforce the finding that newer versions of YOLO not only offer improved accuracy but also better prediction stability in freshwater fish object detection scenarios.

4. Conclusion

This study presents a comparative performance analysis of YOLOv11n and YOLOv12n models for freshwater fish detection, focusing on five commonly consumed Indonesian species: bandeng, gurame, lele, nila, and patin. A partial factorial experimental design was employed to identify the optimal hyperparameter configuration for each model, followed by a final evaluation on an unseen test dataset to ensure objective performance assessment.

Experimental results demonstrate that YOLOv12n achieves superior overall detection performance, as indicated by higher precision, mAP@50, and F1-score compared to YOLOv11n. This indicates that architectural improvements in YOLOv12 enhance the model's ability to produce more accurate and stable predictions across multiple object classes. Meanwhile, YOLOv11n exhibits slightly higher recall and faster inference speed, highlighting its suitability for real-time applications where computational efficiency is a primary concern.

Analysis of the Precision–Recall curves and normalized confusion matrices further reveals that YOLOv12n consistently produces higher true positive rates and lower background misclassification across most fish classes. Both models perform particularly well in detecting patin, while lele remains the most challenging class due to visual similarity with the background and greater intra-class variation. These findings suggest that dataset characteristics and object appearance play a critical role in influencing detection performance.

Overall, this research confirms that YOLOv12n is the more suitable model when detection accuracy is prioritized, whereas YOLOv11n offers a favorable trade-off between accuracy and inference speed. The results contribute to the growing body of research on YOLO-based object detection in the fisheries domain and provide a practical reference for selecting lightweight detection models for freshwater fish identification. Future work may explore advanced data augmentation strategies, class imbalance handling, and deployment on edge devices to further improve robustness and real-world applicability.

5. References

[1] F. Ernawati, A. Syauqy, A. Y. Arifin, M. Y. E. Soekatri, and S. Sandjaja, "Micronutrient Deficiencies And Stunting Were Associated With Socioeconomic Status In Indonesian Children Aged 6–59 months," *Nutrients*, vol. 13, no. 6, 2021, doi: 10.3390/nu13061802.

[2] Kemenkes, *Survei Status Gizi Indonesia (SSGI) 2024*. 2025. Accessed: Oct. 25, 2025. [Online]. Available: https://www.badankebijakan.kemkes.go.id/survei-status-gizi-indonesia-ssgi-2024/

[3] P. P. Indonesia, *Peraturan Presiden (Perpres) Nomor 12 Tahun 2025 tentang Rencana Pembangunan Jangka Menengah Nasional Tahun 2025 - 2029* . 2025. Accessed: Oct. 25, 2025. [Online]. Available: https://peraturan.bpk.go.id/Details/314638/perpres-no-12-tahun-2025

[4] F. C. Wahyuni, U. Karomah, R. W. Basrowi, N. L. Sitorus, and L. A. Lestari, "The Relationship between Nutrition Literacy and Nutrition Knowledge with the Incidence of Stunting: A Scoping Review," 2023, *Airlangga University Faculty of Public Health*. doi: 10.20473/amnt.v7i3SP.2023.71-85.

[5] G. H. Martono, A. Azhari, and K. Mustofa, "An Extended Approach Of Weight Collective Influence Graph For Detection Influence Actor," *International Journal of Advances in Intelligent Informatics*, vol. 8, no. 1, pp. 1–11, Mar. 2022, doi: 10.26555/ijain.v8i1.800.

[6] M. Suparta, M. D. Pradana, and K. S. Jit Singh, "Financial Analysis of Catfish Cultivation Businesses in Sedenganmijen Village, Krian District, Sidoarjo Regency," *International Journal of Economics, Business and Management Research*, vol. 07, no. 12, pp. 54–66, 2023, doi: 10.51505/ijebmr.2023.71204.

[7] Candra Wahyuni and Retno Palupi Yonni Siwi, "Increasing Public Awareness of the Importance of 'Gemari' (Love to Eat Fish) Through Mentoring Catfish Cultivation and Processing," *Journal of Community Engagement in Health*, vol. 7, no. 1, pp. 84–88, Mar. 2024, doi: 10.30994/jceh.v7i1.578.

[8] D. S. Sudaryanti, A. Helmi Munawar, N. Sahroni, and T. Badriatin, "Community Empowerment Through Diversification of Processed Fish Towards Food Security Based on Local Wisdom in the Pandemic COVID-19," vol. 5, no. 1, pp. 29–40, 2022, doi: 10.31764/jces.v3i1.5982.

[9] H. Mustafidah, R. Setiawan, and A. Karim, "Image-Based Classification of Freshwater Fish Species to Support Feed Recommendation Using Random Forest," 2025.

[10] R. Agil, Y. Arjun, and E. P. Silmina, "Deteksi Bahan Pangan Tinggi Protein Menggunakan Model You Only Look Once (YOLO)," *Technology and Science (BITS)*, vol. 6, no. 4, 2025, doi: 10.47065/bits.v6i4.6889.

[11] M. Chaman, A. El Maliki, H. El Yanboiy, H. Dahou, H. Laâmari, and A. Hadjoudja, "Comparative Analysis of Deep Neural Networks YOLOv11 and YOLOv12 for Real-Time Vehicle Detection in Autonomous Vehicles," *International Journal of Transport Development and Integration*, vol. 9, no. 1, pp. 39–48, Mar. 2025, doi: 10.18280/ijtdi.090104.

[12] R. Sapkota *et al.*, "YOLO Advances To Its Genesis: A Decadal And Comprehensive Review Of The You Only Look Once (YOLO) Series," *Artif Intell Rev*, vol. 58, no. 9, Sep. 2025, doi: 10.1007/s10462-025-11253-3.

[13] R. Sapkota, R. H. Cheppally, A. Sharda, and M. Karkee, "RF-DETR Object Detection vs YOLOv12 : A Study of Transformer-based and CNN-based Architectures for Single-Class and Multi-Class Greenfruit Detection in Complex Orchard Environments Under Label Ambiguity," Apr. 2025, [Online]. Available: http://arxiv.org/abs/2504.13099

[14] J. Justam, A. Malik, E. Erlita, D. Mangellak, and Y. Yuyun, “Perbandingan Kinerja YOLO vs Faster R-CNN untuk Deteksi & Estimasi Berat Ikan,” *Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI)*, vol. 7, no. 2, pp. 363–376, Oct. 2024, doi: 10.57093/jisti.v7i2.273.

[15] G. Albeth, A. Putra, and A. D. Wowor, “Analisis Perbandingan Performa YOLO v11 Dan v12 Menggunakan Model N Dan S,” *INNOVATIVE: Journal Of Social Science Research*, vol. 5, pp. 8593–8601, 2025.

[16] Z. Zhang *et al.*, “An Improved YOLOv8n Used for Fish Detection in Natural Water Environments,” *Animals*, vol. 14, no. 14, Jul. 2024, doi: 10.3390/ani14142022.

[17] R. Sapkota, Z. Meng, M. Churuvija, X. Du, Z. Ma, and M. Karkee, “Comprehensive Performance Evaluation of YOLOv12, YOLO11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments,” Feb. 2025, [Online]. Available: http://arxiv.org/abs/2407.12040

[18] Pusaran-KP Foundation, “Laporan Hasil Survey Perkembangan Usaha Budidaya Ikan Air Tawar Di Indonesia,” Oct. 2025. Accessed: Nov. 20, 2025. [Online]. Available: https://pusarankp.org/wp-content/uploads/2025/10/LAPORAN-HASIL-SURVEY-PERKEMBANGAN-BUDIDAYA-AIR-TAWAR-FINAL.pdf

[19] E. Putri Silmina, S. Sunardi, and A. Yudhana, “Comparative Analysis Of YOLO Deep Learning Model For Image-Based Beef Freshness Detection,” *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 11, no. 1, pp. 250–265, Aug. 2025, doi: 10.33480/jitk.v11i1.6784.

[20] A. Yudhana, E. P. Silmina, and Sunardi, “Deteksi Kesegaran Daging Sapi Menggunakan Augmentasi Data Mosaic pada Model YOLOv5sM,” *JRST (Jurnal Riset Sains dan Teknologi)*, pp. 63–71, Apr. 2025, doi: 10.30595/jrst.v9i1.24990.