IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 3, September 2024, pp. 3314~3321
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i3.pp3314-3321

You only look once v8 for fish species identification

Nurfadjri Akbar Rizqi Basuki, Hustinawaty
Department of Information Systems Management, Gunadarma University, Jakarta, Indonesia

Article Info

Article history:

Received Sep 30, 2023
Revised Nov 6, 2023
Accepted Dec 2, 2023

Keywords:

Artificial intelligence
Fish species
Image processing
Object identification
YOLOv8

ABSTRACT

This research aims to test the performance of you only look once (YOLOv8) in identifying fish species in Indonesian waters. Fish image data is obtained from various sources to conduct testing. The results show that YOLOv8 is able to identify fish species with a mAP accuracy rate of 97%. These results reveal the great potential of deep learning technology in supporting the preservation of marine biodiversity as well as the development of various applications, such as fisheries monitoring, conservation, and marine-based tourism development in Indonesia. With its efficient object detection and classification capabilities, YOLOv8 can simplify and accelerate the process of identifying fish species, even on a large scale. Thus, this technology is a highly effective solution to overcome the challenges of manual fish species identification, which requires a lot of time and effort. The results of this study provide valuable insights into the potential utilization of Indonesia's natural resources in the context of scientific development, the tourism industry, and the fisheries sector, which is vital to the country's economy.

Corresponding Author:

Nurfadjri Akbar Rizqi Basuki
Department of Information Systems Management, Gunadarma University
Ciluar Asri Housing, Tanah Baru Village, South Bogor District, Bogor City, West Java, Indonesia
Email: nurfadjriarb@gmail.com

1. INTRODUCTION

Indonesia is one of the countries with very high marine biodiversity [1]. According to data released by the Ministry of Marine Affairs and Fisheries, it is estimated that there are more than 2,500 species of fish that can be found in Indonesian waters [2]. This marine biodiversity is not only an extraordinary natural wealth, but also one of the great potentials that can be utilized for various purposes, such as scientific research, the tourism industry, and the fisheries sector which is very vital for the country's economy [3]. Fish species identification is one of the important first steps in utilizing this marine biodiversity [4]. In general, fish species identification can be done manually or automatically. Manual identification involves a taxonomist using reference books to recognize and classify different types of fish. However, this method requires significant time and effort, making it less efficient when applied on a large scale [5]–[8].

To overcome these challenges, automatic identification using computer technology has become an increasingly popular solution. One method that has been used successfully is deep learning [9]. Deep learning methods utilize artificial neural networks that are capable of learning to recognize objects from image or video data with high accuracy [10], [11]. One of the most well-known deep learning models in fish species identification is you only look once (YOLO) [12]–[15]. YOLOv8 is an updated version of the object detection model that has the ability to efficiently detect and classify objects in a single frame of an image or video. With this technology, fish species identification can be done quickly and accurately, even on a large scale.

This research journal aims to test the performance of (YOLOv8) in the context of fish species identification in Indonesian waters. The journal will use fish image datasets obtained from various sources to conduct the test. The results of this research are expected to provide valuable insights into the effectiveness of YOLOv8 in supporting the preservation of marine biodiversity as well as the development of various applications such as fisheries monitoring, conservation, and marine-based tourism development.

2. RELATED WORKS

YOLO is an algorithm that already exists but has been developed by several researchers in the field of object detection and object identification. This research is inseparable from previous research as a reference. Several studies have focused on developing you only look once in various versions with fish objects. Liu *et al.* [16] developed a fish recognition system to improve the accuracy of fish image segmentation using YOLOv5. Experimental results show that object detection precision can reach 95.4% and semantic segmentation accuracy can reach 98.5% with the proposed algorithm structure. Kuswantori *et al.* [17] developed a mobile fish detection and classification system for the aquaculture industry. The proposed approach is based on the YOLOv4 detection algorithm, which is optimized with a unique labeling technique. The proposed method was tested by using video to detect fish moving on the conveyor, which are placed randomly in positions and sequences at a speed of 505.08 m/h and can achieve an accuracy of 98.15%. Malik *et al.* [18] developed an automatic fish species identification and categorization system proposing a new fish detection network (FD_Net) to detect nine different types of fish species using images captured by cameras based on the improved YOLOv7 algorithm by swapping Darknet53 with MobileNetv3 and depth-separated convolutions for size 3×3 filters in the bottleneck attention module of extended feature extraction network. The accuracy result reached 95.30%.

While there has been progress in the field of deep learning-based methods for fish detection, there are still some obstacles that need to be overcome. One of the main obstacles is the urgent need for more comprehensive and varied data sets to train and evaluate these methodologies. Also, it should be noted that the use of substandard cameras or under poor lighting conditions may affect the accuracy of the fish detection algorithms. YOLOv8 is a state-of-the-art model that builds on the achievements of previous iterations of YOLO by introducing new features and enhancements to improve performance and adaptability. As a single-stage object detection approach, YOLOv8 retains the same benefits as YOLOv1 to YOLOv7. In this regard, YOLOv8 requires only one stage of feature extraction to achieve object detection, making it faster than other two-stage algorithms [19]–[21].

The most noticeable changes in YOLOv8 are the use of the C2f module as the backbone instead of the C3 module, and the number of blocks per stage changing from [3, 6, 9, 3] to [3, 6, 6, 3]. In addition, there is the replacement of anchor base with anchor free, which is inspired by the concept found in TOOD and ppyoloe's YOLOv6. YOLOv8 is adaptable and remains compatible with all previous versions of the YOLO framework, making it easy for users to switch between different versions and evaluate their performance. This makes YOLOv8 an excellent choice for individuals who want to utilize the latest YOLO technology while still having the option of using existing YOLO models. The architecture of YOLOv8 is depicted in Figure 1.

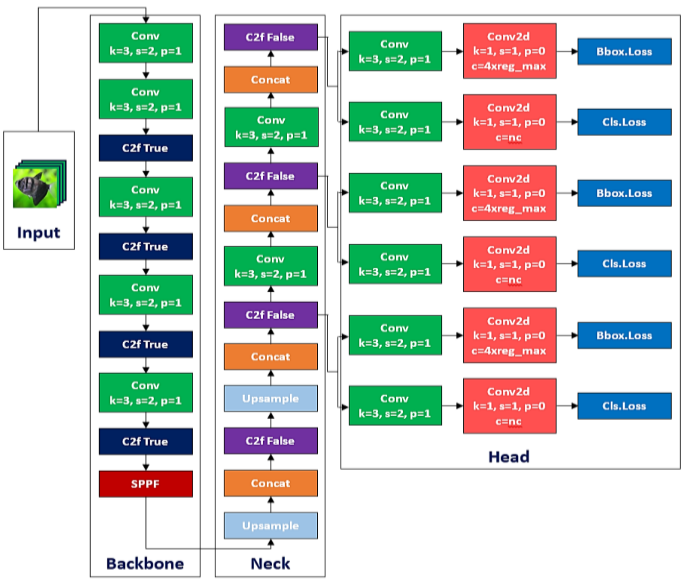

Figure 1. YOLOv8 architecture

3. METHOD

Figure 2 shows the five steps of the proposed fish species identification method: data collection, data preprocessing, YOLOv8 model training, YOLOv8 model evaluation and YOLOv8 model testing. Data collection is the collection of images of fish species from the internet (kaggel and roboflow). Preprocessing consists of normalizing image size, labeling and bounding boxes on data, and data augmentation. YOLOv8 model training is training data with YOLOv8 model. Evaluation of the YOLOv8 model is evaluating the YOLOv8 model of training results, and test the YOLOv8 model to test the capabilities of the model.

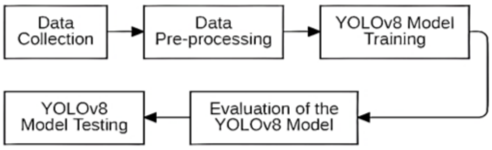

Figure 2. Flowchart for making fish species identification models

3.1. Data collection

The data collected consisted of 982 images with 21 different types of fish including black tetra, botia, koi asagi, koi bekko, koi benigoi, koi chagoi, koi goromo, koi hariwake, koi kohaku, koi kujaku, koi platinum, koi sanke, koi showa, koi tancho, koi utsuri, chef black ranchu, chef oranda, chef ranchu, lemon, manfish, and rainbow. This data will be used to study the patterns and characteristics of the fish. The following is an example of an image of the results of data collection that has been collected, which can be seen in Figure 3.

Figure 3. Image of fish species

3.2. Data pre-processing

In Figure 4 is a data preprocessing flowchart. This step is carried out to adjust the data to the YOLOv8 model format before the model training process is carried out. The purpose of data pre-processing is to improve data quality and YOLOv8 model performance.

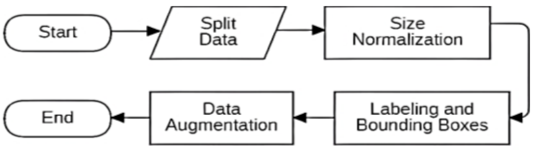

Figure 4. Flowchart pre-processing data

In data pre-processing, the first step is to divide the data into 80% for training, 10% for testing, and 10% for validation. The goal is to prevent overfitting and ensure that the model can generalize well to new data. Then, the image size was normalized to 640×640 pixels to match the model's input format. Next, the data is annotated including labeling and bounding boxes. Labeling is the process of labeling objects in the image, while bounding box is the process of drawing a box around those objects. The last step is to perform data augmentation, which is the process of creating copies of the image with different variations, such as brightness, contrast, and rotation. This is done to increase the size and diversity of the dataset, so that the model can learn better.

3.2.1. Data augmentation

In Figure 5 is a data augmentation flowchart. This step is carried out to multiply data and prevent overfitting before the YOLOv8 model training process is carried out so that the YOLOv8 model is optimal. The purpose of data augmentation is to train the YOLOv8 model to be more robust to changes in data appearance, so that it can produce more accurate object detection.

At the data augmentation stage, the steps taken are flip: horizontal, vertical, which means rotating the image horizontally or vertically. 90° rotate: clockwise, counter-clockwise, upside down means clockwise rotation, anti-clockwise rotation and the image is rotated 180°. Saturation: between -50% and +50%, means to adjust the color intensity in the image. Brightness: between 0% and +20% means to adjust the brightness level of the image. Exposure: between -15% and +15% means to adjust how much light enters the camera sensor. Blur: up to 1.5 px means to make the image blurry. After the data has been augmented to 2946, the next stage will be YOLOv8 model training.

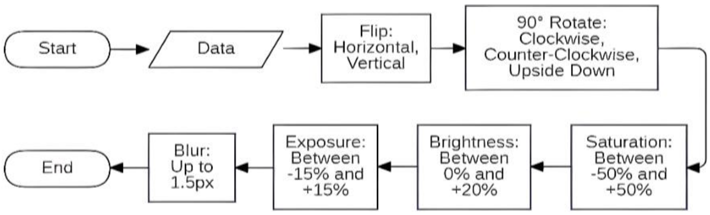

Figure 5. Data augmentation flowchart

3.3. YOLOv8 model training

In Figure 6 is the YOLOv8 model training flowchart. This stage is carried out to train the data with the YOLOv8 model to produce a model that can identify objects, in this case identifying fish species. This model will be used to identify fish species in the image. The YOLOv8 model training step starts with importing the data that has gone through the preprocessing stage. Next, the data is trained with the YOLOv8 model for 200 epochs. Finally, the YOLOv8 model training results are saved in the best.pt file.

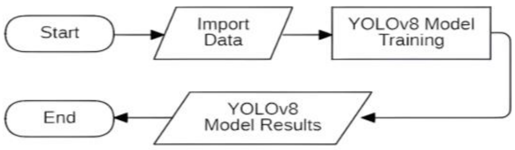

Figure 6. YOLOv8 model training flowchart

3.4. YOLOv8 model evaluation

In Figure 7 is the YOLOv8 model evaluation flowchart. This stage is carried out to evaluate the model that has been made. This evaluation is carried out to measure the performance of the model in identifying objects, in this case identifying fish species. In the model evaluation process, the first thing to do is to visualize the results of the YOLOv8 model training in graphical form against the box_loss value to measure how well the bounding box produced by the model matches the actual bounding box around the object in the image, cls_loss to measure how well the model can classify detected objects into the right class, dfl_loss to measure class imbalance problems in the object detection task, precision to measure the extent to which the positive predictions made by the model are correct. Recall to measure the extent to which the model is able to identify all true positive instances. The F1-score is the harmonic average of precision and recall, providing a more balanced view of the performance of the YOLOv8 model when precision and recall differ significantly. Mean average precision @ intersection over union (IoU) 50% (mAP50) to measure mAP performance at an IoU threshold of 50%. mAP50-mean average precision @ IoU 90% (mAP90) measures mAP performance at an IoU threshold of 50% to 90% and. The confusion matrix is a table used to describe the performance of a classification model by comparing the model's predictions with the actual value of the data. Finally, the YOLOv8 model was evaluated.

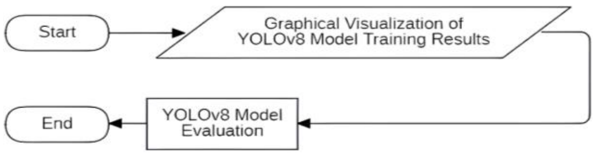

Figure 7. YOLOv8 model evaluation flowchart

4. RESULTS AND CONCLUSION

4.1. YOLOv8 model evaluation results

The YOLOv8 model that has been made trained with 200 epochs produces precision, recall, F1-score, mAP50, and mAP50-90 values which can be seen in Table 1. Precision is the fraction of detections that are true positives, recall is the fraction of ground truth objects detected, F1-score (F1) is the harmonic mean of precision and recall, mAP50 is the average precision of all classes at the IoU threshold of 50% and mAP50-90 is the average precision of all classes at the IoU threshold between 50% and 90%. These values indicate the performance of the model in identifying fish species. The higher the value, the better the performance of the model [22], [23]. Precision, recall, F1-score, and IoU are calculated as (1)-(4):

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{1}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{2}$$

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{3}$$

$$IOU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \tag{4}$$

In (1) to (3) where true positives are the number of true positives, false positives are the number of false positives, false negatives are the number of false negatives [24]. In (4) area of intersection is the overlapping area between the prediction bounding box and the ground truth bounding box, and area of union is the combined area between the prediction bounding box and the ground truth bounding box [25].

Table 1. YOLOv8 model training results

Epoch	Precision	Recall	F1-Score	mAP50	mAP50-90
200	94%	93%	93.4%	97%	94%

The YOLOv8 model training results from 200 epochs have very high values of precision, recall, F1-score, mAP50, and mAP50-90. Next, a visualization of the results of the YOLOv8 model training with a total of 200 epochs was carried out as seen from the values of box_loss, cls_loss, dfl_loss, precision and recall, mAP50 and mAP50-90, and confusion matrix. The YOLOv8 model training results graph can be seen in Figure 8. The box_loss graph shows that the box_loss value decreases, meaning that the model is getting better at producing detection boxes that match the actual object. It can also be seen from the graph that the box_loss train and the box_loss val overlap a little. This means that the model is slightly overfitting. This indicates that the model is able to generalize well to new data, because the performance is not too different between training data and validation data. The cls_loss graph shows that the lower the cls_loss value, the better the model is at classifying objects. It can also be seen from the graph that train cls_loss and val cls_loss coincide slightly. This means that the model is slightly overfitting. This indicates that the model is able to generalize well to new data, because its performance is not too different between training data and validation data. The dfl_loss graph shows that the value of dfl_loss is decreasing, meaning that the model is getting better at solving class imbalance problems in the object detection task. It can also be seen from the graph that the train dfl_loss and val dfl_loss overlap a little. This means that the model is slightly overfitting. This indicates that the model is able to generalize well to new data, because its performance is not too different between training data and validation data. The precision and recall graphs show that the precision and recall values increase, meaning that the model performance increases in classifying or detecting. It can also be seen from the graph that precision and recall coincide a little. This means that the model has a good balance between the ability to produce accurate predictions (precision) and the ability to identify as many objects that actually exist (recall). The mAP50 and mAP50-90 graphs show that the mAP50 and mAP50-90 values are increasing, meaning an overall increase in the performance of the object detection model. It can be seen from the graph that the mAP50 and mAP50-90 overlap slightly. This means that the model has a fairly consistent and uniform performance. at various levels of overlap (IoU thresholds).

Figure 9 shows the confusion matrix which indicates that the model has a good performance in classifying or detecting objects. This is evidenced by the large number of confusion matrix values that predict correctly and the small number of confusion matrix values that predict incorrectly. The fewer confusion matrix values that predict incorrectly, the better the model performance [26].

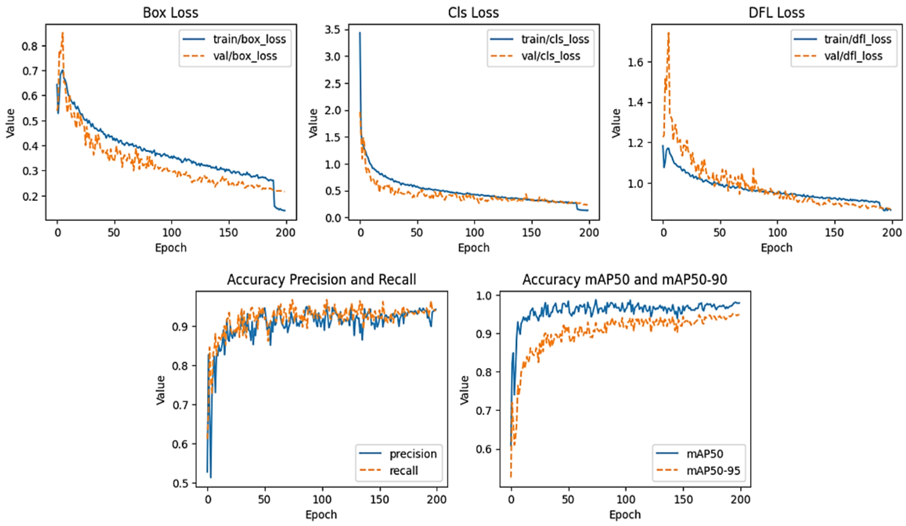

Figure 8. Graph of YOLOv8 model training results

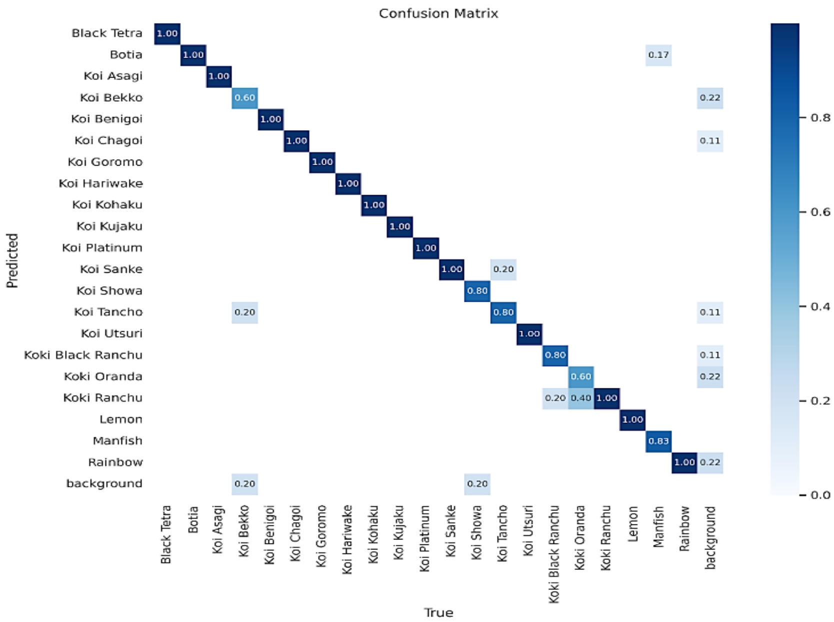

Figure 9. Confusion matrix

4.2. YOLOv8 model trial results

This stage tests the ability of the YOLOv8 model to predict objects in the context of object identification. The purpose of this test is to evaluate whether the YOLOv8 model is able to recognize and classify objects in the image or not. By conducting this trial, we can measure the extent to which the YOLOv8 model can work accurately in identifying objects that appear in various visual situations. The following results of trials that have been carried out can be seen in Figure 10.

Figure 10. YOLOv8 model trial results

The results of testing the YOLOv8 model on 6 fish images showed a very high accuracy of 95.67% on average. The model was able to correctly identify 6 fish species with an accuracy of 98% for black tetra, 95% for botia, 95% for koi asagi, 97% for koi bekko, 98% for koi benigoi, and 91% for koi chagoi. These results indicate that the YOLOv8 model has good potential for use in identifying fish species.

5. CONCLUSION

Identification of fish species with the YOLOv8 model has been successfully developed to identify objects in the form of fish species images. In this study, 982 data were collected with 21 types of fish that had been collected, then pre-processed image size normalization, labeling, and bounding boxes for fish images, data augmentation, the data increased to 2946. Then the YOLOv8 model training process was carried out on 200 epoch data. produces a precision value of 94%, recall of 93%, F1-score of 93.4%, mAP50 of 97% and mAP50-90 of 94%. Evaluation of the model is done by looking at the graphs of the box_loss, cls_loss, dfl_loss, and confusion matrix values that the loss graph shows a decrease in each epoch meaning that the model works well, because at each epoch the loss value decreases and the results of the confusion matrix show that many predict objects correctly and predict little wrong and to make sure the YOLOv8 model actually performs well. The results of tests conducted on 6 fish images showed that the YOLOv8 model was able to correctly identify fish species and had an average accuracy of 95.67% from 6 fish images.

REFERENCES

[1] N. Hanif, A. Murni, C. Tanaka, and J. Tanaka, "Marine natural products from Indonesian waters," *Marine Drugs*, vol. 17, no. 6, 2019, doi: 10.3390/md17060364.

[2] D. Hidayati, B. Oclandhi, N. Maulidinia, N. N. Sa'adah, and A. P. D. Nurhayati, "Short communication: the species and body size composition of pelagic fishes that caught by troll line in the fish landing of Sendang Biru, East Java, Indonesia," *Biodiversitas*, vol. 20, no. 6, pp. 1764–1769, 2019, doi: 10.13057/biodiv/d200638.

[3] G. Canonico *et al*., "Global observational needs and resources for marine biodiversity," *Frontiers in Marine Science*, vol. 6, 2019, doi: 10.3389/fmars.2019.00367.

[4] S. C. Mana and T. Sasipraba, "An intelligent deep learning enabled marine fish species detection and classification model," *International Journal on Artificial Intelligence Tools*, vol. 31, no. 1, 2022, doi: 10.1142/S0218213022500178.

[5] H. T. Rauf, M. I. U. Lali, S. Zahoor, S. Z. H. Shah, A. U. Rehman, and S. A. C. Bukhari, "Visual features based automated identification of fish species using deep convolutional neural networks," *Computers and Electronics in Agriculture*, vol. 167, 2019, doi: 10.1016/j.compag.2019.105075.

[6] E. M. Ditria, S. L. -Marcano, M. Sievers, E. L. Jinks, C. J. Brown, and R. M. Connolly, "Automating the analysis of fish abundance using object detection: optimizing animal ecology with deep learning," *Frontiers in Marine Science*, vol. 7, 2020, doi: 10.3389/fmars.2020.00429.

[7] A. Jalal, A. Salman, A. Mian, M. Shortis, and F. Shafait, "Fish detection and species classification in underwater environments

using deep learning with temporal information," *Ecological Informatics*, vol. 57, 2020, doi: 10.1016/j.ecoinf.2020.101088.
[8] A. Banan, A. Nasiri, and A. T. -Garavand, "Deep learning-based appearance features extraction for automated carp species identification," *Aquacultural Engineering*, vol. 89, 2020, doi: 10.1016/j.aquaeng.2020.102053.
[9] X. Yang, S. Zhang, J. Liu, Q. Gao, S. Dong, and C. Zhou, "Deep learning for smart fish farming: applications, opportunities and challenges," *Reviews in Aquaculture*, vol. 13, no. 1, pp. 66–90, 2021, doi: 10.1111/raq.12464.
[10] S. Cui, Y. Zhou, Y. Wang, and L. Zhai, "Fish detection using deep learning," *Applied Computational Intelligence and Soft Computing*, vol. 2020, pp. 1–13, 2020, doi: 10.1155/2020/3738108.
[11] K. M. Knausgård *et al.*, "Temperate fish detection and classification: a deep learning based approach," *Applied Intelligence*, vol. 52, no. 6, pp. 6988–7001, 2022, doi: 10.1007/s10489-020-02154-9.
[12] A. A. Muksit, F. Hasan, M. F. H. B. Emon, M. R. Haque, A. R. Anwary, and S. Shatabda, "YOLO-Fish: a robust fish detection model to detect fish in realistic underwater environment," *Ecological Informatics*, vol. 72, 2022, doi: 10.1016/j.ecoinf.2022.101847.
[13] D. Xu and Y. Wu, "Improved YOLO-V3 with densenet for multi-scale remote sensing target detection," *Sensors*, vol. 20, no. 15, pp. 1–24, 2020, doi: 10.3390/s20154276.
[14] E. Prasetyo, N. Suciati, and C. Fatichah, "Yolov4-tiny with wing convolution layer for detecting fish body part," *Computers and Electronics in Agriculture*, vol. 198, 2022, doi: 10.1016/j.compag.2022.107023.
[15] Z. Zhao, Y. Liu, X. Sun, J. Liu, X. Yang, and C. Zhou, "Composited FishNet: fish detection and species recognition from low-quality underwater videos," *IEEE Transactions on Image Processing*, vol. 30, pp. 4719–4734, 2021, doi: 10.1109/TIP.2021.3074738.
[16] Q. L. Liu *et al.*, "A multitask model for realtime fish detection and segmentation based on YOLOv5," *PeerJ Computer Science*, vol. 9, pp. 1–31, 2023, doi: 10.7717/PEERJ-CS.1262.
[17] A. Kuswantori, T. Suesut, W. Tangsrirat, G. Schleining, and N. Nunak, "Fish Detection and Classification for Automatic Sorting System with an Optimized YOLO Algorithm," *Applied Sciences*, vol. 13, no. 6, 2023, doi: 10.3390/app13063812.
[18] H. Malik, A. Naeem, S. Hassan, F. Ali, R. A. Naqvi, and D. K. Yon, "Multi-classification deep neural networks for identification of fish species using camera captured images," *PLoS ONE*, vol. 18, no. 4 April, Apr. 2023, doi: 10.1371/journal.pone.0284992.
[19] J. M. Liang, S. Mishra, and Y. L. Cheng, "Applying image recognition and tracking methods for fish physiology detection based on a visual sensor," *Sensors*, vol. 22, no. 15, 2022, doi: 10.3390/s22155545.
[20] L. Li, G. Shi, and T. Jiang, "Fish detection method based on improved YOLOv5," *Aquaculture International*, vol. 31, no. 5, pp. 2513–2530, 2023, doi: 10.1007/s10499-023-01095-7.
[21] E. Mahoro and M. Akhloufi, "Automated fish detection and classification on sonar images using detection transformer and YOLOv7," in *Sixteenth International Conference on Quality Control by Artificial Vision*, SPIE, 2023, doi: 10.1117/12.2688330.
[22] P. Sarkar, S. De, and S. Gurung, "Fish detection from underwater images using yolo and its challenges," in *Advances in Intelligent Systems and Computing*, Springer Nature Singapore, 2023, pp. 149–159. doi: 10.1007/978-981-99-1472-2_13.
[23] E. Prasetyo, N. Suciati, and C. Fatichah, "Yolov4-tiny and spatial pyramid pooling for detecting head and tail of fish," in *ICAICST 2021 - 2021 International Conference on Artificial Intelligence and Computer Science Technology*, IEEE, 2021, pp. 157–161. doi: 10.1109/ICAICST53116.2021.9497822.
[24] H. Li, L. Deng, C. Yang, J. Liu, and Z. Gu, "Enhanced YOLO v3 tiny network for real-time ship detection from visual image," *IEEE Access*, vol. 9, pp. 16692–16706, 2021, doi: 10.1109/ACCESS.2021.3053956.
[25] H. M. Lathifah, L. Novamizanti, and S. Rizal, "Fast and accurate fish classification from underwater video using you only look once," *IOP Conference Series: Materials Science and Engineering*, vol. 982, no. 1, 2020, doi: 10.1088/1757-899X/982/1/012003.
[26] N. Radha, R. Swathika, and P. S. Shreya, "Automatic fish detection in underwater videos using machine learning," in *6th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), I-SMAC 2022 - Proceedings*, IEEE, Nov. 2022, pp. 587–592. doi: 10.1109/I-SMAC55078.2022.9987363.

BIOGRAPHIES OF AUTHORS

Nurfadjri Akbar Rizqi Basuki holds a Bachelor of Computers (S.Kom.) in Information Systems, and is currently pursuing a Master of Information Systems Management (MMSI.) in Information Systems Software in addition to several certificates and professional skills. He is currently staff at Gunadarma University. His research areas of interest include artificial intelligent, image processing, and data science. He can be contacted at email: nurfadjriarb@staff.gunadarma.ac.id.

Hustinawaty holds a Doctor of Information Technology degree from Gunadarma University, Indonesia, awarded in 2012. She also obtained her MMSI degree from Gunadarma University, Indonesia, in 1994. Currently, she serves as the Head of the Master of Information Systems Management Study Program at Gunadarma University, Indonesia. Her research interests encompass information systems management and information systems technology. She can be contacted at email: hustina@staff.gunadarma.ac.id.