Fish Detection and Classification using YOLOv8 for Automated Sorting Systems

[1]Ari Kuswantori, [2*]Dwi Joko Suroso
[1]Departement of Electronics Engineering, Politeknik Gajah Tunggal, Indonesia
[2]Departement of Nuclear Engineering and Engineering Physics, Universitas Gadjah Mada, Indonesia
Email: [1]ari@poltek-gt.ac.id, [2]dwi.jokosuroso@ugm.ac.id

Article Info

Article history:
Received May 24th, 2024
Revised Jun 19th, 2024
Accepted Jul 19th, 2024

Keyword:
Automatic Fish Sorting
Deep Learning
Fish Classification
Fish Harvest Technology
YOLOv8

ABSTRACT

Automation plays a crucial role in scaling up freshwater fish cultivation to address the future threat of food scarcity and meet growing nutrition needs. The fish industry, in particular, develops automation in the sorting and selection processes. However, research on this technology's development is still very limited. In this work, we propose an approach for detecting and classifying fish running on conveyors. We use YOLOv8, which is the most popular and newest deep learning model for object detection and classification. We conducted our test using the KMITLFish dataset, a moving conveyor belt recording that encompasses common cultivated freshwater fish in Thailand along with some endemic species. As a result, our proposed method was able to accurately detect and classify eight types of fish at a conveyor speed of 505.08 m/h. Moreover, we developed this work using a ready-to-use AI platform, intending to directly contribute to the advancement of automatic fish sorting system technology in the fish industry.

Corresponding Author:
Dwi Joko Suroso**,**
Departement of Nuclear Engineering and Engineering Physics,
Universitas Gadjah Mada,
Jl. Grafika 2, Yogyakarta 55281, Indonesia.
Email: dwi.jokosuroso@ugm.ac.id

DOI: http://dx.doi.org/10.24014/ijaidm.v7i2.30967

1. INTRODUCTION

Food source scarcity and insecurity have currently become a global issue that is attracting a lot of attention [1-3] due to the continued shrinking of agricultural land [4], the fact that the world's population continues to increase [5], and the impact of climate change [6, 7]. These factors raise concerns about the potential future scarcity of food supplies. Therefore, many preventive measures are being undertaken to increase food production, especially to boost agricultural productivity [8]. This includes the cultivation of freshwater fish in the agricultural industry, because fish is an important agricultural commodity. People all over the world consume fish because it is a source of protein and other beneficial nutrients for the body [9].

One way to increase agricultural productivity can be achieved through automation [10-12]. Apart from increasing productivity, implementing automation can also maintain the quality of agricultural products [13]. The fish farming industry implements automation for tasks like fish monitoring, water quality systems, feeding operations, and so on, up to the automatic fish sorting system [9]. For the automatic fish sorting system itself, machine or computer vision technology based on deep learning is considered the most effective. This technology relies on the development of a reliable deep-learning model to detect and classify fish [14-16].

In the last five years, many deep learning-based approaches have been developed to automatically detect and classify farmed fish, which can be used to develop automatic fish sorting technology. The author summarizes this in Table 1.

Table 1. Survey on fish classification work based on deep learning focused on farming fish in the last 5 years

Work	Dataset	Method	Accuracy	Note
Rauf H.T., et al. (2019) [17]	The Fish-Pak [18] dataset, consists of six fish species.	Applied a CNN with 32 layers, built from the VGGNet plus four convolutional layers.	96.94%	Fish classification on static images
Abinaya N.S., et al. (2021) [19]	Fish-Pak	Employ background removal using BLOB, auto-rotation using MSEE, and image segmentation as image preprocessing, then create 3 AlexNet networks for each body, scale, and head of fish, and finally summarize them with NBF.	98.64%	Fish classification on static images
Xu, et al. (2021) [20]	Fish-Pak	They proposed SE-ResNet152, transfer learning from ImageNet, and then applied it to the Fish-Pak dataset, optimized with a class-balanced focal loss function.	98.80%	Fish classification on static images
Shammi S.A., et al. (2021) [21]	Fish-Pak	Applied a classic CNN.	88.96%	Fish classification on static images
Banerjee A., et al. (2022) [22]	Their own dataset consists of 1,500 images of three Indian local carps.	Proposed Deep Convolutional Autoencoder (DCA).	97.33%	Fish classification on static images
Previous work by authors (2022) [23]	Fish-Pak	Employing YOLOv4	77.42%	Fish classification on static images
Md. Asif Ahmed, et al. (2023) [24]	BD Fish (constructed by themselves) consists of eight classes of fish.	CNN combined with a convolutional LSTM.	97%	Fish classification on static images
Bo Gong, et al. (2023) [25]	Fish-Pak	Proposed a transfer learning system combined with vision transformers.	98.34%	Fish classification on static images
Previous work by authors (2023) [26]	KMITLFish, consists of eight fish species.	Using 564 training images and YOLOv4, optimized with data augmentation, input image and training methods, and labeling technique	98.15%	Fish detection and classification run on a conveyor (video image)

Rauf H.T. et al. (2019) proposed a Convolutional Neural Network (CNN) with 32 layers, which was built from the VGGNet, and then added four more convolutional layers [17]. They applied the approach to Fish-Pak [18], a public dataset consisting of six freshwater-cultured fish species. This approach can classify accurately up to 96.94%. In their 2021 study, Abinaya N.S. et al. used a Binary Large Object (BLOB) to remove the background, a Multi-Stage Exhaustive Enumerative (MSEE) for auto-rotation, and image segmentation as image preprocessing. They then made three AlexNet networks, one for each fish's body, scale, and head, and finally summed them up with a Naive Bayesian Fusion (NBF). They also applied this approach to the Fish-Pak dataset, achieving an accuracy of up to 98.64% [19]. Xu et al. (2021) also utilized the Fish-Pak dataset. In their work, they proposed a SE-ResNet152 transfer learning from ImageNet and then applied it to the dataset, optimized with a class-balanced focal loss function [20]. With the approach they proposed, they were able to achieve very good final classification accuracy results, namely 98.80%.

Shammi S.A. et al. (2021) tested a classic CNN on the Fish-Pak dataset. However, the classification accuracy results were only 88.96% [21]. Banerjee A., et al. (2022) proposed a Deep Convolutional Autoencoder (DCA) to classify carp. They compiled a dataset consisting of 1,500 images of three Indian local carps. The resulting accuracy is 97.33% [22]. In the previous work (2022), we applied YOLOv4 to the Fish-Pak dataset and were only able to produce a classification accuracy of 77.42% for the complete class [23]. Md. Asif Ahmed et al. (2023) created their own dataset, and they called it the BD Fish dataset. The dataset consists of eight classes of fish. In their work, they proposed a CNN combined with a Convolutional LSTM (Long Short-Therm Memory Unit) [24]. And this approach is able to classify with an accuracy of 97%. Bo Gong et al. (2023) discovered the combination of transfer learning and vision transformers, resulting in a classification accuracy of 98.34% on the Fish-Pak dataset [25].

The review above highlights the limited availability of public datasets containing images of freshwater fish. Most of the work reviewed uses the FishPak dataset, which contains only static fish images. Our previous work considered dynamic data from Thai endemic fish videos moving on a conveyor in 2023 [26]. The fish in the video move randomly on a conveyor, making the dataset ideal for developing deep learning-based auto-sorting technology for cultivated fish. We utilized 564 images and various techniques to optimize YOLOv4, achieving the highest accuracy of 98.15% in that work. Meanwhile, the training images are considered large, and currently YOLOv8 has been developed, which, of course, was created with various improvements from the previous versions.

From the works that we have reviewed above, we can summarize the research gap: almost all published work on the detection or classification of freshwater-cultured fish does not use moving datasets, so it is less applicable for the purpose of developing fish auto-sorting technology. The work that uses moving images is our previous work, which we limited to using YOLOv4, but the number of training images used is still considered large. This research gap led us to question whether we could propose a method using a different model, aiming to reduce the number of training images while still achieving satisfactory recognition results.

So in this work, we propose different models with different training data treatments, as a development from previous studies. The model we use in this work is YOLOv8, which is the latest version of YOLO. We apply it to the same dataset (KMITLFish) to compare its performance with the same goal: developing auto-sorting technology for farmed fish. Ultimately, this work aims to develop an effective and ready-to-use auto-sorting system for farmed fish, by utilizing a suitable dataset, the latest version of the YOLO deep learning model, and a simple method.

2. RESEARCH MATERIAL AND METHOD

Figure 1 depicts the research flow or method in this work, which can be broadly divided into two parts: the training process for preparing the model and testing the model. The next subsection will provide a more detailed explanation of each part of the flow chart.

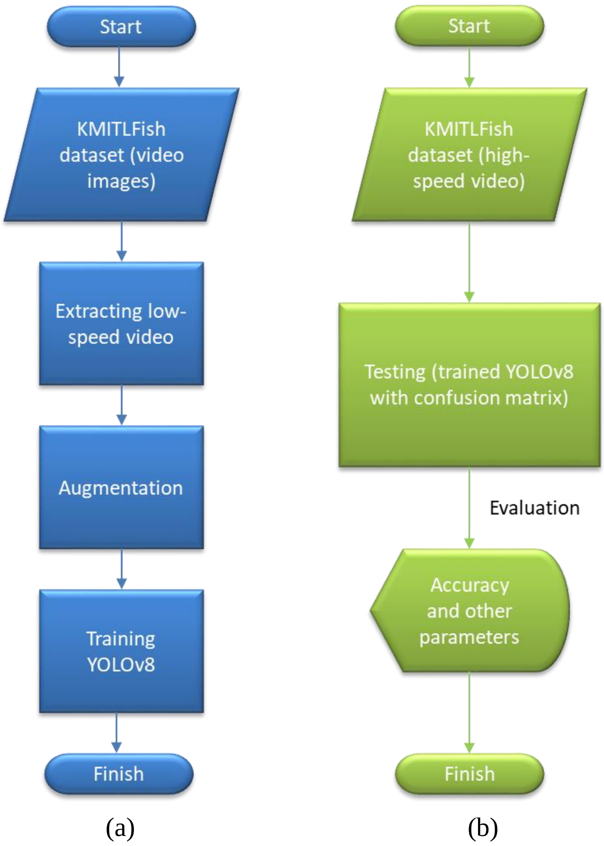


Figure 1. The work flow chart, where (a) describes the training process and (b) for the testing

2.1. Dataset and Image Augmentation

This work utilizes the KMITLfish dataset, which was introduced and utilized for the first time by the author in the previous work [26]. This dataset was created in the laboratory of the Food Engineering Department, School of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand, in collaboration with the Instrumentation and Control Engineering Department at the same university and a small freshwater fish farming company in Thailand. The dataset consists of 8 types of farmed fish species that are commonly cultivated, sold, and consumed in Southeast Asian countries, especially in Thailand and surrounding countries such as Cambodia, Laos, Myanmar, etc. Moreover, 3 of the 8 fish species are endemic to the Chao Phraya River, the main river in Thailand. The eight species of fish are listed in Table 2, and examples of fish for each species in the KMITLfish dataset can be seen in Figure 2.

Table 2. Fish species in the dataset

No.	Fish class (species)
1	Yeesok *(Labeo rohita)*
2	Nuanchan *(Cirrhinus microlepsis)*
3	Tapian *(Barbonymus gonionotus)*
4	Nai *(Cyprinus carpio)*
5	Jeen Ban *(Hypophthalmichthys molitrix)*
6	Jeen To *(Hypophthalmichthys nobilis)*
7	Nin *(Oreochromis niloticus)*
8	Sawai *(Pangasianodon hypophthalmus)*

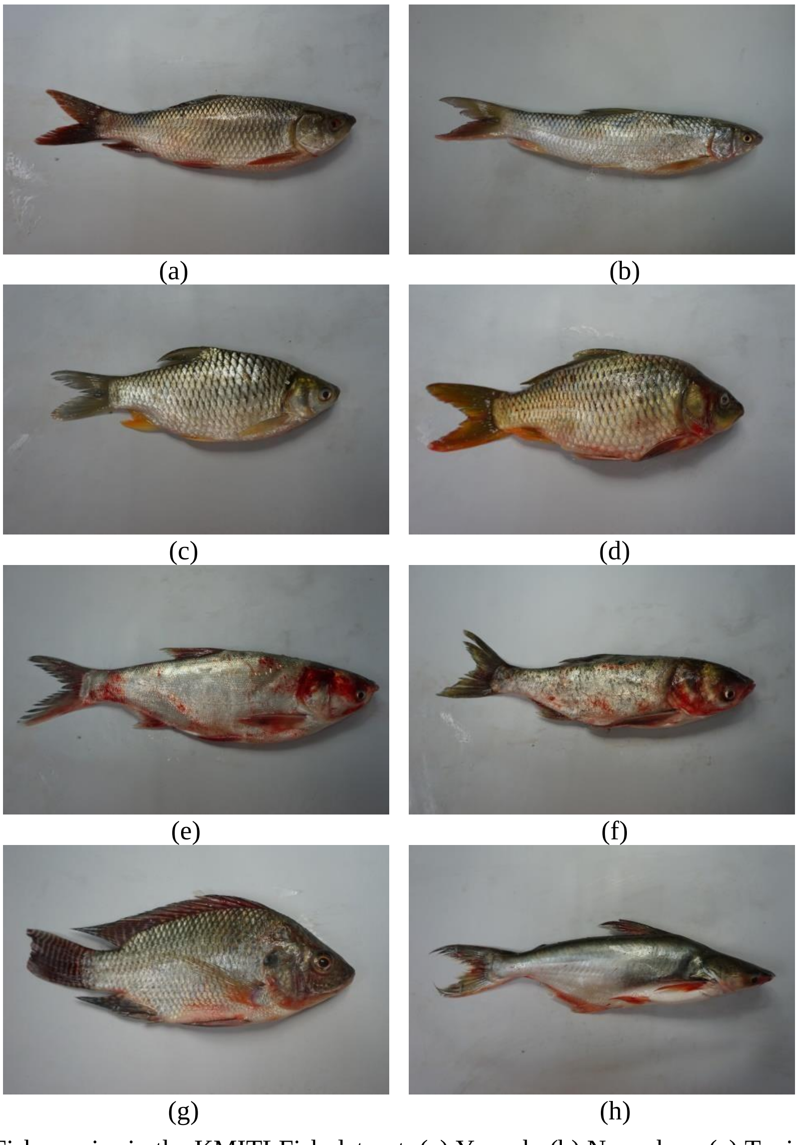


Figure 2. Fish species in the KMITLFish dataset: (a) Yeesok, (b) Nuanchan, (c) Tapian, (d) Nai, (e) Jeen Ban, (f) Jeen To, (g) Nin, and (h) Sawai [26]

In Thailand, farmers cultivate these fish simultaneously in a single pond. Once harvested, the workers manually sort the fish at the fish hub before shipping and selling each type of fish [26]. As a result, this dataset is ideal for developing automatic fish detection and classification technology because it will have a direct impact on the industry. This dataset consists of static images and videos. The video image is a recording of fish running on a conveyor in random positions and sequences with two different speed settings, namely low (116.65 m/h) and high (505.08 m/h). The video has a frame size of 1920x1080 (width x height) and a frame rate of 29.97 frames per second. We use this video because this work aims to detect and classify fish to develop automatic sorting system technology. Figure 3 illustrates the setup for capturing video images and provides an overview of the results.

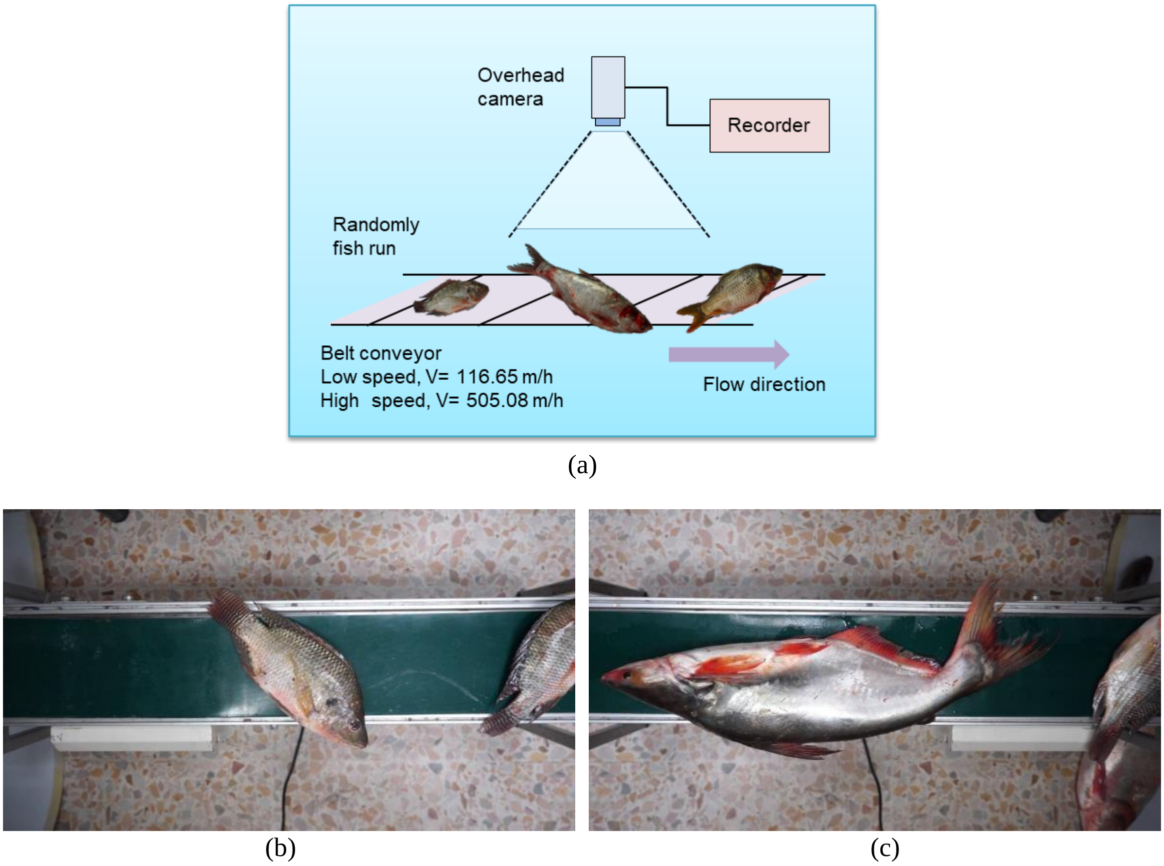


Figure 3. (a) Set-up to take the video, and (b,c) are examples of capturing video results [26]

When developing fish auto-detection and classification technology using these video images, we encountered several challenges, particularly those based on deep learning:

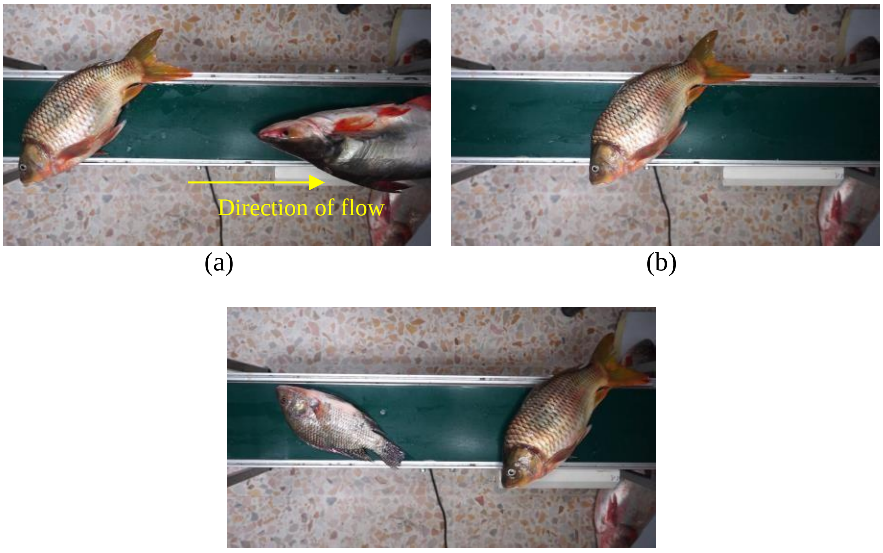
1. First, the position of the fish running on the conveyor is random. This requires that the deep learning model created have high translation invariant capabilities so that the model can still classify fish correctly even though the fish is passing in various possible positions, such as turning, turning upside down, and so on. Despite the conditions of this dataset, the model must maintain its accuracy in classifying fish, even when parts of the fish, like the bent tail, are not visible.
2. Second, the conveyor used is small and does not dominate the frame. This results in a varied background for the fish object, rather than a monotonous conveyor system. This is a challenge and can affect detection and classification accuracy [26, 27].
3. Thirdly, a significant challenge lies in the fact that the fish in the dataset exhibit high similarity. For instance, the Jeen Ban-Jeen To, Yeesok-Nuanchan, and Nai-tapian fish exhibit striking similarities. Jeen Ban and Jeen To ish have almost the same appearance. The only difference is that Jeen To's head is larger, and his scales have black spots. Yeesok and Nuanchan fish also have almost the same shape. The most striking difference between them is that Yeesok has darker-colored scales, whereas the Nuanchan fish has brighter (shinier) scales and orange fins. Meanwhile, Nai and Tapian fish also have almost the exact same shape. A significant characteristic that differentiates the two is the color of their scales, where the Nai fish has yellow scales while the Tapian fish has bright (shiny) skin. This similarity between one class and another presents a high-level challenge. The techniques used in an approach to classify these fish automatically must be able to maintain these distinguishing features, and the deep learning model employed must be able to extract them effectively so that the model can classify each fish correctly.

In this work, we use three videos from the dataset. We extracted the video "low speed 1 no audio.mp4", which has a duration of 17 minutes and 13 seconds and used it as training data. Meanwhile, the video "high speed no audio.mp4" (hereinafter referred to as "video high speed-1") is 8 minutes and 24 seconds long, and the video "washed-high speed no audio.mp4" (hereinafter referred to as "video high speed-2"), with a duration of 17 minutes and 13 seconds, is used as test data. Table 3 summarizes the details of this video's use.

Table 3. Video images utilized in the KMITLFish dataset

No.	The file's name in the dataset	Names in this work	Conveyor speed (m/h)	Frame size (W x H)	Frame rate (frame/s)	Duration	Note
1	low speed 1 no audio.mp4	Low-speed video	116.65			17 min, 13 s	Extracted for training data
2	high speed no audio.mp4	High-speed video 1	505.08	1920x 1080	29.97	8 min, 24 s	For testing data
3	washed- high speed no audio.mp4	High-speed video 2	505.08			17 min, 13 s	For testing data

We obtained 140 static images from low-speed video extraction, which will serve as training data. We obtained this image by taking a screenshot of each fish as it passed through the conveyor. We took a screenshot of each fish when it passes through three positions: when it appears fully visible, when it is in the exact middle position, and before it leaves the frame. Figure 4 exemplifies the extraction of the Nai fish.



(a) (b)

(c)

Figure 4. Example of extracting images for one fish from the low-speed video: (a) when the fish appears in its entirety, (b) at the exact center position, and (c) before the fish leaves the frame [26]

We then augmented this image using horizontal and vertical flips techniques. These augmentation techniques are suitable for recognizing fish on conveyors [19]. After the augmentation process, the image data grew to 443 images. When compared with our previous work, the number of these training images is lower, whereas in our previous study, we used 564 images. In this work, we eliminated images that contained fish objects that were not fully visible.

2.2. YOLOv8 and The Training Process

You Only Look Once (YOLO) is the most popular and widely used deep learning algorithm or model for object detection and classification. YOLO has been widely developed and applied in various fields such as industry, agrotechnology, autonomous vehicles, etc., even in the development of military technology [29, 30]. Since its first release, YOLO has continued to be developed until the current version 8, which was introduced in January 2023 by Ultralytics, the company that developed YOLOv5. In version 8, YOLO is introduced in several variants, including YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x. Each variant has the same size but has a different speed, number of parameters, and FLOPs [31]. Figure 5 from the YOLOv8 website presents a simple and easier-to-understand YOLOv8 architecture.

YOLOv8 reportedly supports multiple vision tasks such as object detection, pose estimation, tracking, segmentation, and classification [32]. The major development of YOLOv8 compared to the previous version is that YOLOv8 now uses mosaic data augmentation, anchor-free detection, a C2F module, a decoupled head, and loss for final classification [32]. Architecturally, YOLOv8 uses a modified CSPDarknet53 backbone. The CSPLayer used in previous versions was replaced by the C2f module. Through feature pooling into a fixed-size map, a spatial pyramid pooling fast (SPPF) layer expedites computation. SiLU activation and batch normalization are present in every convolution. To handle objectivity, classification, and regression tasks separately, the head is detached. These architectural details of YOLOv8 are nicely described by [33].

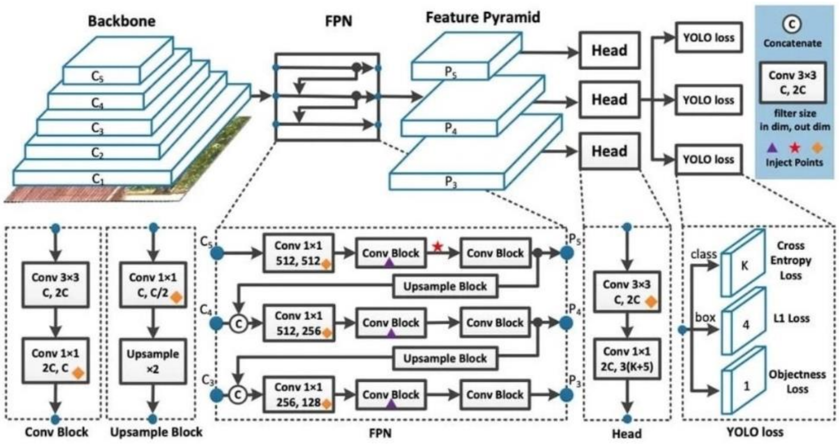


Figure 5. Architecture of YOLOv8 [28]

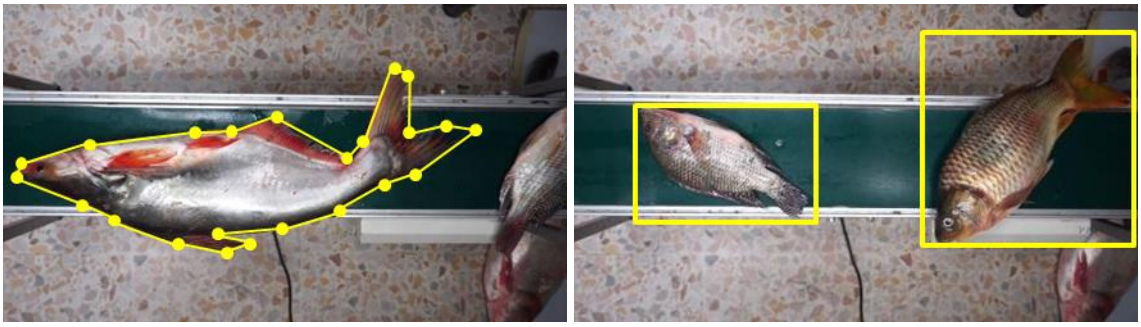
In the training process, we used 443 images, as explained previously. Among the 443 training data images, many contain more than one different type of fish in a single image, allowing the ground truth data to exceed the total number of training images (443). Nin fish appear most frequently in this training data because they are the most numerous species in the dataset. Meanwhile, for the other fish, the numbers are relatively even, between 60 and 66. Figure 6 illustrates the ground truth labeling used in this training process. In one training image, we can see that there may be more than one fish object. Table 4 lists the number of ground truths for each fish class.

Table 4. Ground truth for every class in the training process

No.	Fish Class	Ground truth from the training images	Used for training
1	Yeesok	63	63
2	Nuanchan	66	66
3	Tapian	63	63
4	Nai	63	63
5	Jeen Ban	60	60
6	Jeen To	60	60
7	Nin	333	333
8	Sawai	60	60

During the training process, we employ a combination of labeling techniques, specifically the landmark technique for large fish (Sawai) and the conventional square labeling technique for other small fish. The landmarking technique removes various objects from large Sawai fish, potentially disrupting the learning process in the deep learning model. For small fish, the surrounding background is not very diverse. The conveyor still dominates the background and incorporating it into the training process is beneficial as it aids in extracting fish body shape features that are crucial for fish differentiation. For other smaller fish, we employ conventional labeling techniques. The work [26, 27] provides a more detailed explanation of the combination labeling technique, which has been reported to increase accuracy when using YOLO. Figure 6 also provides an overview of this labeling technique.

This work uses the YOLOv8m 640, which is 169 MB in size. The training and work process uses CiRA-CORE, an AI deep learning platform from AMI (Advanced Manufacturing Innovation)-KMITL, Bangkok, Thailand. The main advantage of this platform is its ease of use and connection to other devices, making it a ready-to-use technology [34]. The training process uses a batch size of 16, subdivision 8, and only rotation for data enrichment with a setting of -180 to 180 degrees in 90 steps. We can successfully complete the training process in just 27 minutes and 52 seconds, with a loss of 0.0570. We carried out the training in 2197 iterations. The PC used is a desktop PC with an Intel® 1151 CoreTM i7-9700 3.0 GHz processor, 32 GB of DDR4/3200 RAM (Random Access Memory), and an NVIDIA GeForce RTX 3070 8 GB GDDR6 GPU (Graphical Processing Unit). As a practical recommendation, readers can use the same YOLOv8 type and training method to help achieve effective training, as demonstrated in this work, in the same or similar cases.



(a) (b)

Figure 6. Labeling technique in the training process: (a) with the landmarking technique for Sawai; (b) using the conventional (square) technique for other classes

2.3. Confusion Matrix

A confusion matrix evaluates the detection and classification results of the proposed approach. This matrix consists of 4 parts, namely the truth base group, namely True Positive (TP) and True Negative (TN), and the error base group, namely False Positive (FP) and False Negative (FN). TP is defined as when the model can detect and classify fish correctly. TN occurs when the model fails to identify fish due to their absence. We do not use this parameter in our work. FP is determined when the model detects incorrectly or can detect but misclassify, including double classification or more. When the model fails or cannot detect fish, it yields FN. We measure these parameters when the fish object is precisely in the middle of the frame.

Equation (1) to (4) in this work denote the values of accuracy, precision, recall or sensitivity, and F1 score using this confusion matrix.

.

$$Accuracy = \frac{TP}{TP + FP + FN} \times 100\% \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \times 100\,\%. \tag{2}$$

$$Recall\,/\,Sensitivity = \frac{TP}{TP + FN} \times 100\,\%. \tag{3}$$

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall} \times 100\,\%. \tag{4}$$

3. EXPERIMENTAL RESULTS AND DISCUSSION

After using training images from low-speed video extraction for the YOLOv8 training process, we test the trained model using high-speed videos 1 and 2. We evaluate the test results using a confusion matrix to derive accuracy values and other parameters. These parameters include precision, recall or sensitivity, and the F1 score. Figure 7 depicts the entire flow of this work.

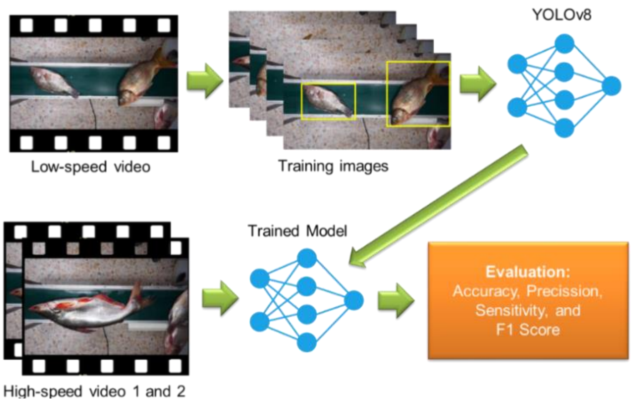


Figure 7. The Workflow of This Paper

Table 5 summarizes the results of these experiments, and Figure 8 and Figure **9** provide a more detailed depiction of the entire performance model for all classes and the confusion matrix of model performance, respectively.

Table 5. Experimental results

Fish classes	Ground truth	Correctly detected and classified (TP)	Wrong/ double classified or more (FP)	Miss detects (FN)	Accuracy (%)	Precission (%)	Sensitivity (%)	F1 Score (%)	Note
Yeesok	18	17	1	0	94.44	94.44	100	97.14	1 double classified also as Nuanchan
Nuanchan	21	20	0	1	95.24	100	95.24	97.56	1 miss detect because doubt beetween Nuanchan and Yeesok below 50% confidence score
Tapian	20	20	0	0	100	100	100	100	-
Nai	20	19	0	1	95.00	100	95.00	97.44	1 miss detect because doubt beetween Nai and Tapian below 50% confidence score
Jeen Ban	21	21	0	0	100	100	100	100	-
Jeen To	21	15	2	4	71.43	88.24	78.95	83.33	2 wrong classifieds as Jeen Ban, 4 miss detect because doubt beetween Jeen To and Jeen Ban below 50% confidence score
Nin	171	171	0	0	100	100	100	100	-
Sawai	21	21	0	0	100	100	100	100	-
Total	313	304	3	6	97.12	99.02	98.06	98.54	-

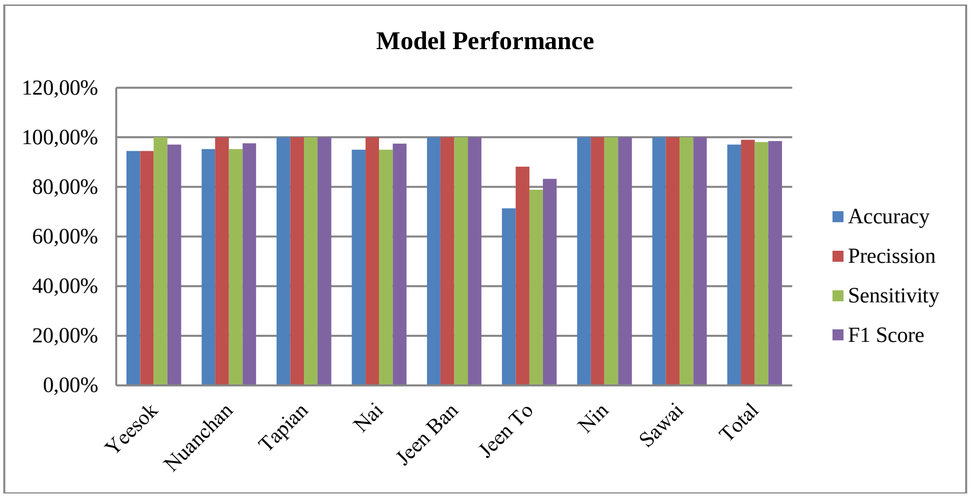


Figure 8. Model performance

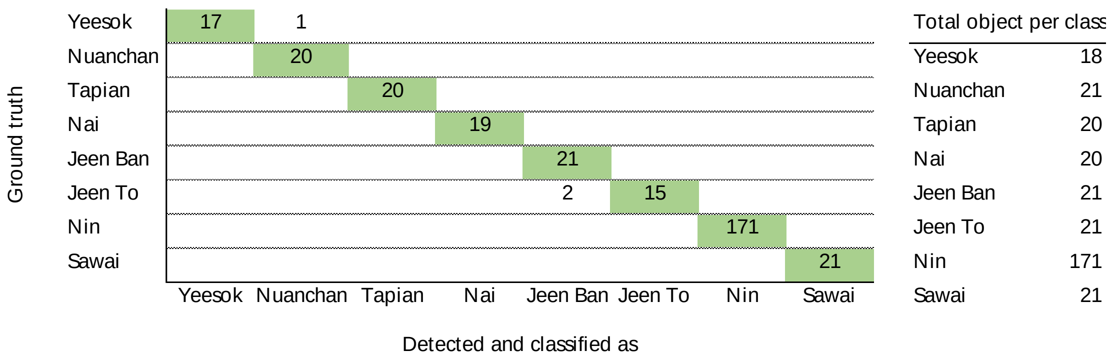
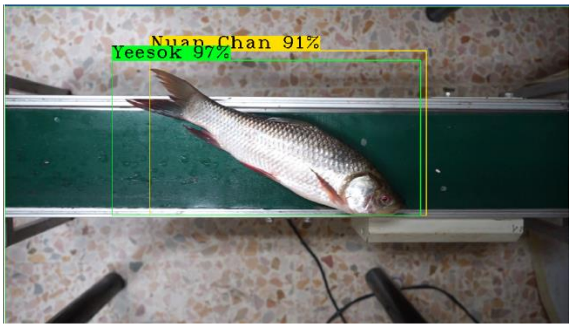


Figure 9. Confusion matrix for model performance

According to the experimental results, the proposed approach can achieve a total accuracy of 97.12%, with precision, sensitivity, and F1 score values of 99.02%, 98.06%, and 98.54%, respectively. The Tapian, Jeen Ban, Nin, and Sawai classes demonstrate perfect detection and classification results, achieving complete correct detection and classification. The Jeen To class displayed the worst results, with 6 errors, followed by the Yeesok, Nuanchan, and Nai classes, each with 1 error. If detailed, the six errors in the Jeen To class consisted of two misclassifications of Jeen Ban fish and four failed detections. The model made a double classification error in the Yeesok class, classifying the Yeesok fish as a Nuanchan fish. In the Nuanchan class, one error is a failed detection. In the Nai class, one error is also a detection failure.

If we observe, the problem lies in the similarities between fish in several classes. Jeen To fish bears a striking resemblance to Jeen Ban fish, Yeesok shares similarities with Nuanchan, and Nai fish resembles Tapian. This is what makes the model misclassify or double-classify fish with other similar fish. The model misclassified the Jeen To fish as Jeen Ban and double-classified the Yeesok fish as Nuanchan. In addition, the model finds it challenging to distinguish between the fish in the two classes due to their similarity, leading to uncertainty at a confidence score level of less than 50%. Consequently, the model's failure to detect results from setting the detection threshold at a minimum confidence score of 50%. This happens when the model fails to identify a Jeen To fish because it isn't sure whether it is a Jeen Ban or Jeen To fish (confidence score of less than 50%). It also happens when the model fails to identify a Yeesok or Tapian fish when it isn't sure whether it is a Nuanchan or Nai fish.This could possibly be caused by two causes. First, the model's ability to differentiate between similar types of fish is imperfect and can be improved. Second, the condition of the fish image is disturbed when the model performs classification.

Regarding the second cause, visual observations during the testing process reveal that certain conditions, such as glare from lights, can disrupt the appearance of the fish image object during the classification process. This glare distorts the pattern and color of fish scales, despite these being the primary distinguishing features of fish, particularly those of similar species. In the Jeen To class, as the classification progresses, the glare exposes the Jeen To fish's body, distorting the black spot pattern on its scales. This distortion makes the fish's body appear predominantly silver in color, a characteristic of the Jeen Ban fish, leading to the classification of the fish as a Jeen Ban fish. In other detections, this also makes the model hesitate to classify Jeen To and Jeen Ban fish with confidence scores below 50% for each, so that in the end it fails to detect the fish that should be Jeen To. This also happened in Yeesok and Nai's classes. Since the glares in the middle of fish' body, its scales take on a dominant silver color, leading to the classification of the Yeesok fish as a Nuanchan fish. However, the Nai fish eluded detection due to its challenging model, raising doubts about its Tapian classification. This is because Nuanchan and Tapian fish have predominantly silver-scale colors. Examples of misclassification due to this (glare interference) can be seen in Figure 10 and Figure **11**.



(a)

(b) (c)

Figure 10. (a) an example of a double detection of a Yeesok and Nuanchan, while (b) represents the ground truth for Yeesok, and (c) represents Nuanchan

(a)

(b) (c)

Figure 11. (a) illustrates a false detection of a Jeen To fish classified as Jeen Ban, while (b) represents the ground truth for Jeen To, and (c) represents Jeen Ban

Similar to the previously published work by the authors [26], the YOLOv4-based approach also cannot perfectly detect and classify fish due to this glare problem. The model's uncertainty led to the misclassification or non-detection of some fish, with a confidence score below 50%. Therefore, by reducing or eliminating glare, which can disrupt the fish's appearance, we can improve the performance of this work in the future. If the selection process is conducted indoors, we should avoid placing light or lighting persistently above the fish object. If the selection process takes place outdoors, we can incorporate a light net to prevent the dazzling light from disrupting the fish's recognition process.

4. CONCLUSION

This work aims to detect and classify fish automatically using YOLOv8, which is the latest version of YOLO, to develop automatic sorting system technology in the fish industry. To validate our proposal, we utilized the KMITLFish dataset. We intend to use this video-formatted dataset to develop an automatic sorting system. The dataset comprises 8 species of freshwater fish, collectively cultivated in a single pond, and manually sorted by humans at the fish hub after harvest, prior to their sale to customers per species. The experimental results evaluated the proposed approach using a confusion matrix, yielding a final accuracy of 97.12%, along with precision, sensitivity, and F1 score values of 99.02%, 98.06%, and 98.54%, respectively. Interestingly, in this case, using a different model, YOLOv8, the performance results are not much different from using YOLOv4, as in the author's previously published work. In fact, the number of training images in this work is 27.31% less. However, the performance of our proposal is relatively high, and because we developed it on a ready-to-use software platform, it could directly contribute to the establishment of the fish auto-sorting machine, thereby meeting our research objectives. Our future work will focus on investigating glare effects and developing a mechanism to reduce them.

ACKNOWLEDGEMENTS

Authors would like to thank Dept. of Instrumentation and Control Engineering, Dept. of Food Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand, for providing the KMITLFish public dataset. Authors also thank AMI-KMITL for granting the CiRA-Core software license. Special thanks to Mr. Agustami Sitorus, from the National Research and Innovation Agency (BRIN), Indonesia, who has assisted in compiling the KMITLFish dataset.

REFERENCES

[1] V. M. Hazzard, K. A. Loth, L. Hooper, and C. B. Becker, "Food insecurity and eating disorders: A review of emerging evidence," *Current psychiatry reports,* vol. 22, pp. 1-9, 2020.
[2] M. Thomas, D. P. Miller, and T. W. Morrissey, "Food insecurity and child health," *Pediatrics,* vol. 144, 2019.
[3] D. Schanzenbach and A. Pitts, "How much has food insecurity risen? Evidence from the Census Household Pulse Survey," *Institute for Policy Research Rapid Research Report,* pp. 1-10, 2020.

[4] A. Hossain, T. J. Krupnik, J. Timsina, M. G. Mahboob, A. K. Chaki, M. Farooq, *et al.*, "Agricultural land degradation: processes and problems undermining future food security," in *Environment, climate, plant and vegetation growth*, ed: Springer, 2020, pp. 17-61.

[5] A. D. Tripathi, R. Mishra, K. K. Maurya, R. B. Singh, and D. W. Wilson, "Estimates for world population and global food availability for global health," in *The role of functional food security in global health*, ed: Elsevier, 2019, pp. 3-24.

[6] T. Hasegawa, G. Sakurai, S. Fujimori, K. Takahashi, Y. Hijioka, and T. Masui, "Extreme climate events increase risk of global food insecurity and adaptation needs," *Nature Food,* vol. 2, pp. 587-595, 2021.

[7] H. El Bilali, I. H. N. Bassole, L. Dambo, and S. Berjan, "Climate change and food security," *Agriculture & Forestry/Poljoprivreda i Sumarstvo,* vol. 66, 2020.

[8] D. Paudel, R. C. Neupane, S. Sigdel, P. Poudel, and A. R. Khanal, "COVID-19 Pandemic, Climate Change, and Conflicts on Agriculture: A Trio of Challenges to Global Food Security," *Sustainability,* vol. 15, p. 8280, 2023.

[9] C. Wang, Z. Li, T. Wang, X. Xu, X. Zhang, and D. Li, "Intelligent fish farm—the future of aquaculture," *Aquaculture International,* pp. 1-31, 2021.

[10] W.-S. Kim, W.-S. Lee, and Y.-J. Kim, "A review of the applications of the internet of things (IoT) for agricultural automation," *Journal of Biosystems Engineering,* vol. 45, pp. 385-400, 2020.

[11] V. Puranik, A. Ranjan, and A. Kumari, "Automation in agriculture and IoT," in *2019 4th international conference on internet of things: smart innovation and usages (IoT-SIU)*, 2019, pp. 1-6.

[12] M. S. A. Mahmud, M. S. Z. Abidin, A. A. Emmanuel, and H. S. Hasan, "Robotics and automation in agriculture: present and future applications," *Applications of Modelling and Simulation,* vol. 4, pp. 130-140, 2020.

[13] A. Sartal, D. Carou, R. Dorado-Vicente, and L. Mandayo, "Facing the challenges of the food industry: Might additive manufacturing be the answer?," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture,* vol. 233, pp. 1902-1906, 2018.

[14] 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, pp. 66-90, 2021.

[15] G. Kaur, N. Adhikari, S. Krishnapriya, S. G. Wawale, R. Malik, A. S. Zamani, *et al.*, "Recent Advancements in Deep Learning Frameworks for Precision Fish Farming Opportunities, Challenges, and Applications," *Journal of Food Quality,* vol. 2023, 2023.

[16] D. Li and L. Du, "Recent advances of deep learning algorithms for aquacultural machine vision systems with emphasis on fish," *Artificial Intelligence Review,* vol. 55, pp. 4077-4116, 2022.

[17] 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, p. 105075, 2019/12/01/ 2019.

[18] S. Z. H. Shah, H. T. Rauf, M. IkramUllah, M. S. Khalid, M. Farooq, M. Fatima, *et al.*, "Fish-Pak: Fish species dataset from Pakistan for visual features based classification," in *Data in Brief* vol. 27, D. i. Brief, Ed., ed: Data in Brief, 2019, p. 104565.

[19] A. N.S, S. D, and R. K. S, "Naive Bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries," *Ecological Informatics,* vol. 61, p. 101248, 2021/03/01/ 2021.

[20] X. Xu, W. Li, and Q. Duan, "Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification," *Computers and Electronics in Agriculture,* vol. 180, p. 105878, 2021/01/01/ 2021.

[21] S. A. Shammi, S. Das, M. Hasan, and S. R. H. Noori, "FishNet: Fish Classification using Convolutional Neural Network," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2021, pp. 1-5.

[22] A. Banerjee, A. Das, S. Behra, D. Bhattacharjee, N. T. Srinivasan, M. Nasipuri, *et al.*, "Carp-DCAE: Deep convolutional autoencoder for carp fish classification," *Computers and Electronics in Agriculture,* vol. 196, p. 106810, 2022/05/01/ 2022.

[23] A. Kuswantori, T. Suesut, W. Tangsrirat, and N. Nunak, "Development of object detection and classification with YOLOv4 for similar and structural deformed fish," *EUREKA: Physics and Engineering,* vol. 2, pp. 154–165, 2022.

[24] M. A. Ahmed, M. S. Hossain, W. Rahman, A. H. Uddin, and M. T. Islam, "An advanced Bangladeshi local fish classification system based on the combination of deep learning and the internet of things (IoT)," *Journal of Agriculture and Food Research,* vol. 14, p. 100663, 2023/12/01/ 2023.

[25] B. Gong, K. Dai, J. Shao, L. Jing, and Y. Chen, "Fish-TViT: A novel fish species classification method in multi water areas based on transfer learning and vision transformer," *Heliyon,* 2023.

[26] 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, p. 3812, 2023.

[27] A. Kuswantori, T. Suesut, W. Tangsrirat, and S. Satthamsakul, "Fish Recognition Optimization in Various Backgrounds Using Landmarking Technique and YOLOv4," presented at the The 37th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC), Phuket, Thailand, 05-08 July 2022.

[28] J. Austen. (2024, 20/6/2024). *What is YOLOv8? Exploring its Cutting-Edge Features*. Available: https://yolov8.org/what-is-yolov8/

[29] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications," *Multimedia Tools and Applications,* pp. 1-33, 2022.

[30] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo algorithm developments," *Procedia Computer Science,* vol. 199, pp. 1066-1073, 2022.

[31] G. Jocher, A. Chaurasia, and J. Qiu. (2023, March, 24th). *YOLO by Ultralytics*. Available: https://github.com/ultralytics/ultralytics
[32] M. Sohan, T. Sai Ram, and C. V. Rami Reddy, "A Review on YOLOv8 and Its Advancements," in *Data Intelligence and Cognitive Informatics*, Singapore, 2024, pp. 529-545.
[33] Github. (2023, March, 24th). *YOLOv8 by MMYOLO*. Available: https://github.com/open-mmlab/mmyolo/tree/main/configs/yolov8
[34] R. Jomtarak, C. Faikhamta, T. Prasoplarb, and K. Lertdechapat, "CiRA-Core: The Connector for Developer Teachers and User Teachers to Artificial Intelligence," in *International Conference on Technology in Education*, 2023, pp. 13-19.

BIBLIOGRAPHY OF AUTHORS

Ari Kuswantori holds Doctor of Engineering (Dr. Eng) in instrumentation and control engineering from King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand in 2023. Now, he is involved and active as a lecturer, researcher, and technology development engineer in the electronic engineering department of Politeknik Gajah Tunggal, a private non-profit polytechnic located in Tangerang, Banten, Indonesia, which provides free education and guaranteed employment after graduation which focuses on children from underprivileged families. His research areas of interest include image processing, deep learning, image classification, object detection, electrical control systems, robotics, and instrumentation.

Dwi Joko Suroso received B.Eng from Universitas Gadjah Mada (UGM), Indonesia in 2010 majoring Engineering Physics, both M.Eng and D.Eng from King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand in 2012 and 2023, respectively. He is faculty member at Dept. Nuclear Engineering and Engineering Physics, Faculty of Engineering, UGM. His research interest including smart sensor system, short-range wireless communication, robotics and machine/deep learning implementation.