Original Research Article: full paper

(2022), «EUREKA: Physics and Engineering» Number 2

DEVELOPMENT OF OBJECT DETECTION AND CLASSIFICATION WITH YOLOV4 FOR SIMILAR AND STRUCTURAL DEFORMED FISH

Ari Kuswantori
Department of Instrumentation and Control Engineering[1]

Taweepol Suesut ✉
Department of Instrumentation and Control Engineering[1]
taweepol.su@kmitl.ac.th

Worapong Tangsrirat
Department of Instrumentation and Control Engineering[1]

Navaphattra Nunak
Department of Food Engineering[1]

[1]School of Engineering
King Mongkut's Institute of Technology Ladkrabang (KMITL)
Chalongkrung Road No. 1, Ladkrabang District, Bangkok Province, Thailand, 10520

✉ **Corresponding author**

Abstract

Food scarcity is an issue of concern due to the continued growth of the human population and the threat of global warming and climate change. Increasing food production is expected to meet the challenges of food needs that will continue to increase in the future. Automation is one of the solutions to increase food productivity, including in the aquaculture industry, where fish recognition is essential to support it. This paper presents fish recognition using YOLO version 4 (YOLOv4) on the «Fish-Pak» dataset, which contains six species of identical and structurally damaged fish, both of which are characteristics of fish processed in the aquaculture industry. Data augmentation was generated to meet the validation criteria and improve the data balance between classes. For fish images on a conveyor, flip, rotation, and translation augmentation techniques are appropriate. YOLOv4 was applied to the whole fish body and then combined with several techniques to determine the impact on the accuracy of the results. These techniques include landmarking, subclassing, adding scale data, adding head data, and class elimination. Performance for each model was evaluated with a confusion matrix, and analysis of the impact of the combination of these techniques was also reviewed. From the experimental test results, the accuracy of YOLOv4 for the whole fish body is only 43.01 %. The result rose to 72.65 % with the landmarking technique, then rose to 76.64 % with the subclassing technique, and finally rose to 77.42 % by adding scale data. The accuracy did not improve to 76.47 % by adding head data, and the accuracy rose to 98.75 % with the class elimination technique. The final result was excellent and acceptable.

Keywords: computer vision, cultured-fish recognition, fish automation, fish classification, YOLO.

DOI: 10.21303/2461-4262.2022.002345

1. Introduction

Global warming and climate change are issues that have received much attention and concern from many scientists and researchers after the COVID-19 pandemic disaster [1]. The biggest problem that is feared to arise from this issue is that it will lead to food scarcity [2]. Food scarcity has also become a threat to the world community because the human population is growing [3]. For this reason, increasing food production is expected as one solution to answer the challenges of food needs that will continue to increase in the future. One of the strategies to increase productivity is automation. Besides increasing production, automation in the food industry will also maintain food quality and safety factors [2, 3].

In the aquaculture industry, especially in the fish industry and its processing, the fish recognition process has been needed to support the automation process during the sorting process, inspection, or another process [4]. Fish recognition is an exciting and challenging topic because it presents various challenges [5, 6]. One of the challenges is that the types of fish are usually very similar to one another. In addition, the condition of the fish outside the water undergoes structural deformation, such as changed or damaged eyes, scales, and fins. These things become a unique challenge in the fish recognition process, especially for fish in the aquaculture industry [7].

Fish classification (FC) using computer vision and machine learning is an exciting research topic and has been continuously developed over the last two decades [7]. In recent years, new methods or approaches have been developed to achieve a high level of accuracy.

Image processing plays a vital role before the recognition algorithm is run. For example, it removes noise with a median filter, detects fish objects, and separates them from the background with a histogram, BLOB analysis, and threshold [7]. The other is image enhancement with contrast enrichment, auto segmentation of fish objects with various techniques [7–9], orientation correction using multi-stage exhaustive enumerative (MSEE) [7], Color Multi-Scale Retinex (MSR) to overcome water turbidity [10], and GMM, Pixel-Wise Posteriors [11], and Optical Flow [12] for fish detection in complex background scenarios.

Features on fish objects are essential for machine learning for the classification process [13]. The literature reported a maximum of 133 features for FC [14]. Features widely used for FC include geometric, statistical, color, and textual features [15]. Then, Artificial Neural Networks (ANN) with supervised learning algorithms were widely chosen for fish classification [9, 16–18].

The algorithms used also vary, such as CNN [4, 19-23], YOLO [10, 12, 24], few-shot learning for limited training images [25], Alex- Net, ResNet-18, ResNet-50, Inception-V3, and GoogLeNet [7]. In addition, modifications were also made to the algorithm to increase performance, such as modifications to Alex-Net [26], modification of CNN [27], and modification of Res-Net [9]. Another approach is combining standard algorithms with decision algorithms such as SVM (Support Vector Machine) [28, 29], Naive Bayesian classifier (NBC) [7, 30], KNN [31], decision tree [29], and backpropagation classifier [18]. Classification performance improves with integrating traditional classifiers such as random forest trees and SVM as layers compared to other standard deep learning networks [32]. In cloudy and blurred underwater images, hyperspectral imaging systems are reported to be more suitable than visual systems [33]. In addition, a transfer learning approach has also been attempted to use a pre-trained network for FC [34].

This paper presents the results of fish recognition using the viral YOLO algorithm with a relatively new version (version 4) which is still rarely used in fish. The condition of very similar fish between species and fish with structural deformation becomes a difficult challenge. In addition to testing the recognition of fish under these conditions using YOLOv4, a significant contribution of this work is the trial of combining YOLOv4 with various techniques to determine its impact on accuracy results. Applying landmarking, subclassing, adding fish scale data, adding fish head data, and class elimination strategies are among the techniques mentioned.

2. Material and methods

2. 1. Fish dataset

In this work, the «Fish-Pak» dataset has been used, in which its dataset consists of six fish species, namely:

1) *Catla* (Thala);
2) *Hypophthalmichthys molitrix* (Silver);
3) *Labeo rohita* (Rohu);
4) *Cirrhinus mrigala* (Mori);
5) *Cyprinus carpio* (Common carpio);
6) *Ctenopharyngodon Idella* (Grass carp) [35].

These fish are commonly bred in South Asia, such as India and Pakistan [36], and some types can also be found in other regions, such as Southeast Asia. Image data on each fish species, consisting of images of the whole body, head only, and scales only. There are 271 images of the body, 254 images of the scale, and 390 images of the head, so that the total fish images in this dataset are 915 images.

Some types of fish are very similar and very difficult to distinguish by ordinary people. In addition, the pictures of the fish were taken in out of water conditions with misalignment and structural deformation conditions such as the eyes and scales that were damaged lightly, moderately, and severely until the contents of the stomach came out. This causes this dataset to be considered quite ideal in this work. Examples of fish images in this dataset can be considered in **Fig. 1**.

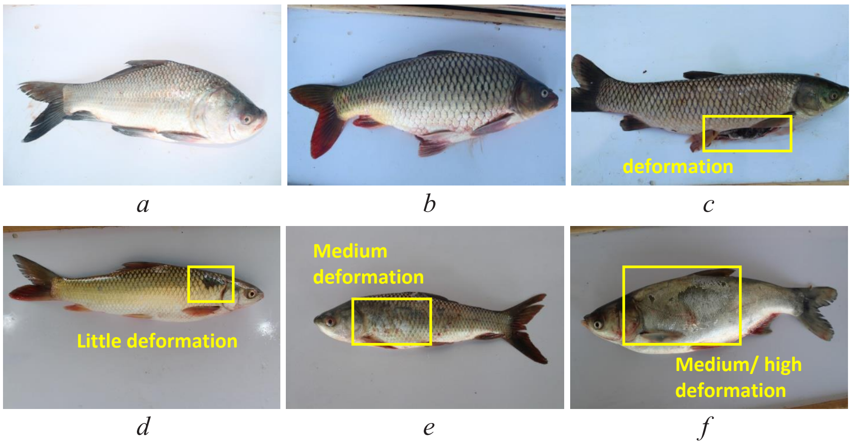


Fig. 1. Example images from Fish-Pak dataset:
a – Catla; *b* – C. Carpio; *c* – G. Carp; *d* – Mori; *e* – Rohu; *f* – Silver

2. 2. Image augmentation

The number of images for each fish species in the Fish-Pak dataset is minimal (less than 100) and not balanced between one species and another. Little image data for each class type makes validation low, and unbalanced data makes the algorithm get unequal opportunities in the training process for each class type. For this purpose, image augmentation is required [7]. The augmentation technique used in this work is flip, rotation, and translation, which is suitable for fish objects on the conveyor [32].

(1) is the number of original images in each class of fish (I_f^C), where $\left\{I_f^{1C}, I_f^{2C}, \ldots, I_f^{N_cC}\right\}$ indicates each fish image. The multiplier factor (m_f) is obtained through (2). It is obtained by comparing the number of images in each class (N_C) with the target images (N_T) where $N_C<N_T$. In this work, the fish body's target images are set to 100 [7]. By repeating this procedure for each class of fish, the set of multiplication factors can be obtained for all classes (m_f^C, $C\in[1,2,\ldots,C]$). The multiplier factor indicates the number of augmented images for each class to be generated (I_{af}^{mc}). The augmentation techniques of flip (F_a), rotation (θ_a), and translation (T_a) were performed randomly for each image in each class (I_f^{nc}) according to the multiplier factor ($a\in\left[1,2,\ldots,m_f^C\right]$) as in (3). Finally, all augmented images $\left\{I_{af}^{1C}, I_{af}^{2C}, \ldots, I_{af}^{N_cC}\right\}$ were collected with the original images so that a new dataset was formed (I_{Nf}^{c}).

As a consequence, the image is randomly divided into 80 % for training and 20 % for testing [26]. A description of this dataset can be summarized in **Table 1**.

With augmentation, the number of images becomes more prosperous, and now the data is more balanced, as indicated by the STD (Standard Deviation) value, which was initially 23.16 and then decreased to 19.87.

$$I_f^C=\left\{I_f^{1C}, I_f^{2C}, \ldots, I_f^{N_cC}\right\}, \tag{1}$$

$$m_f=\left\{m_f^c=\left|1-\frac{N_C}{N_T}\right|, C\in[1,2,\ldots C], N_C<N_T\right\}, \tag{2}$$

$$I_{af}^{mc}=\left\{H\left(I_f^{nc}, F_a, \theta_a, T_a\right)\middle| a\in\left[1,2,\ldots,m_f^c\right]\right\}, \tag{3}$$

$$I_{Nf}^{C}=\left\{I_f^{1C}, I_f^{2C}, \ldots, I_f^{N_cC}, I_{af}^{1C}, I_{af}^{2C}, \ldots, I_{af}^{N_cC}\right\}. \tag{4}$$

Table 1

Dataset image detail and augmentation (body)

Fish	Number of Images (body)	Multiplication Factor	Number of Augmented Images	New dataset (Body)	Training (80 %)	Testing (20 %)
Catla	20	4	80	100	80	20
C. Carpio	50	1	50	100	80	20
G. Carp	11	9	99	110	88	22
Mori	70	1	70	140	112	28
Rohu	73	1	73	146	117	29
Silver	47	2	94	141	113	28
Total	271	–	466	737	590	147
Standard Deviation Average	23.16	–	–	19.87	–	–

In this work, the effect of adding scale and head data to the training process on the algorithm used is also tested. For this reason, scale and head image data in the dataset are also used. However, the scale and head data augmented and used for training are only limited to 3 classes, namely Catla, C. Carpio, and G. Carp, because only these three classes are challenging to detect and classify by the algorithm (low detection and classification results). Augmentation is also done so that the data is balanced for each class. The multiplication factor is determined in the same way that new datasets can be created. For scale image data, N_T is set to 27, and the head image is set to 48. All new datasets for scale and head in the three classes are used as training data. It can be observed in detail in **Table 2**.

Table 2

Scale and head image data and augmentation

Fish	Scale Image	Scale & Augmented	For Training (Scale)	Head Image	Head & Augmented	For Training (Head)	Total Training w/Scale & Head
Catla	11	33	33	25	75	75	188
C. Carpio	44	44	44	64	64	64	188
G. Carp	9	27	27	16	48	48	163
Mori	71	71	0	100	100	0	112
Rohu	62	62	0	114	114	0	117
Silver	57	57	0	71	71	0	113
Total	254	294	104	390	472	187	881
Standard Deviation Average (3 classes)	16.05	7.04	–	20.83	11.09	–	–

2. 3. YOLO and training process

In this work, YOLO version 4 (YOLOv4) is used. YOLO (You Only Look Once) is a very popular, widely used, and widely developed algorithm for object detection because of its ability to quickly and accurately detect objects. It is the reason why YOLO is used in this work. YOLO was built from 24 convolution layers and two fully connected layers at the beginning of its development, as shown in **Fig. 2**. Until now, YOLO has continued to be developed to improve accuracy and detection time, including up to YOLOv4 used in this work [37]. The YOLO training process in each test was carried out with a batch size of 32 with 32 subdivisions. Data enhancement was performed by rotation with a threshold between a minimum of –180° and a maximum of 180° (with 90 steps), and contrast with a threshold between a minimum of 0.4 and a maximum of 1.1 (with 0.2 steps). Data enhancement during the training process is not done by simulating noise and blur. Each training process is carried out until the average loss value becomes acceptable, i. e., up to 0.01 to a maximum of 0.03.

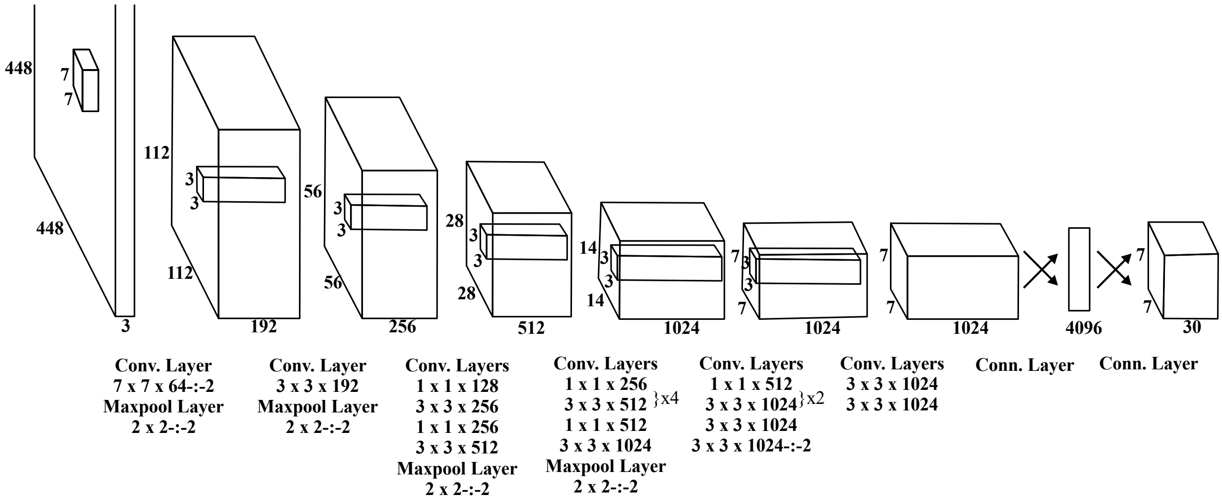
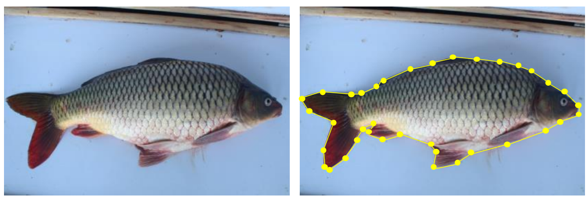


Fig. 2. YOLO Architecture [38]

2. 4. Experimental testing

YOLOv4, which is expected to work well to detect and classify similar and deformed fish, is applied to the selected dataset. Then several experiments were carried out to determine the impact on the results. The experiment combined the YOLOv4 with several techniques: landmarking, subclassing, adding scale data in the training process, adding scale and head data, and class elimination.

Recognition optimization using the landmarking technique (LM technique): this method makes the occupancy ratio of objects in the image effective during the training process. The occupancy ratio is the ratio between the object that wants to be recognized and the total area in the image or bounding box, including the background [7]. A higher excellent occupancy ratio (1 or close to 1) means expected more effectiveness for the algorithm to extract the features of the object during the training process, so that it is expected that the recognition result will be more accurate. The occupancy ratio of the image or bounding box (OR_{bb}) is obtained from the comparison between the object area in the image or bounding box (I_{bb}) and the total area of the row (M) and column (N) in the image or bounding box as in (5). In this work, let's compared the YOLOv4 training process with the whole image input and the landmarking method to determine the impact of the occupancy ratio on objects during training on the recognition output, as shown in **Fig. 3**.



a *b*

Fig. 3. Object segmentation method for YOLOv4 training: *a* – insert the whole image; *b* – use the landmarking technique

This landmarking technique is carried out with the CiRA-Core software:

$$OR_{bb} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} I_{bb}(i,j)}{M \times N}. \tag{5}$$

Subclassing: the subclassing technique is used when different subspecies within a class, such as different colors, patterns, or shapes, can be distinguished. As in this work, class C. Carpio has C. Carpio brown and C. Carpio red (**Fig. 4**), which are in the same class (C. Carpio). For that, the class C. Carpio red will be added in this subclassing technique. This method is applied to see the impact on the recognition ability in the class and as a whole.

a *b*

Fig. 4. The same class (C. Carpio) consists of different subspecies:
a – C. Carpio Brown; *b* – C. Carpio Red

Trials were carried out by incorporating images of heads and scales into the training process. It is intended to enrich object feature references and focus on algorithm training. Head and scale images are used because each type of fish in the dataset can be distinguished from the head and scales. The head and scale images have been augmented as previously described.

Class elimination technique: algorithm limitations in classification, are evaluated by class elimination techniques.

The number of classes eliminated (E_C) is determined by N_C-1, where N_C is the number of classes whose accuracy level cannot be accepted, as shown in (6):

$$E_C = N_C - 1. \tag{6}$$

2. 5. Validation matrix

A confusion matrix is used to validate the experimental results. The confusion matrix is built from 4 building blocks, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). True Positive (TP) is described as when the model detects the object correctly. True Negative (TN) is described when the model does not detect an object because the object does not exist. False Positive (FP) is described when the model detects an object but is wrong, including when it detects a double, even though it has correct predictions (double prediction with the wrong class). Furthermore, False Negative (FN) is described when the model does not detect an object even though it exists [8].

Accuracy is one of the evaluation metrics, which is denoted in (7). It is the ratio of the total number of correct predictions to the total of all predictions:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\,\%, \tag{7}$$

where, TP – True Positives, TN – True Negatives, FP – False Positives, and FN – False Negatives.

Alternatively, the level of model accuracy can also be expressed by:

$$Accuracy = \frac{\sum_i^N Pi}{\sum_i^N |Qi|} \times 100\,\%, \tag{8}$$

where $\sum_i^N Pi$ is the number of correct predictions, and $\sum_i^N |Qi|$ is the total number of predictions.

Precision is the ratio between correctly classified fish (TP) and positive detection (number of TP and FP). It calculates the percentage of fish classified accurately as:

$$Precision = \frac{TP}{TP + FP} \times 100\,\%. \tag{9}$$

Recall or sensitivity is the ratio between the correctly classified fish (*TP*) and the fundamental truth fish (total number of *TP* and *FN*), which can be defined as:

$$Recall\ /\ Sensitivity = \frac{TP}{TP+FN} \times 100\ \%. \tag{10}$$

Specificity is determined by the ratio of *TN* to the sum of *FP* and *TN*, as described:

$$Specivity = \frac{TN}{TN+FP} \times 100\ \%. \tag{11}$$

*F*1*Score* (Measure *F*) is a metric calculated as the average of precision symphony and memory [39], as denoted by the following equation:

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall} \times 100\ \%. \tag{12}$$

3. Results and discussions

The experimental results were obtained from testing 20 % of the test data from the augmented dataset as described previously. First of all, YOLOv4 is trained with raw images as a whole, and the results are evaluated. From the test results, the results are not good. Only one class (Mori) achieved good accuracy (96.43 %), while the other classes only achieved 65 % accuracy (Catla) or below (Rohu; 48.28 % and Silver; 39.29 %). Class G. Carp achieves very low accuracy (9.09 %), and even class C. Carpio achieves 0 % accuracy. The overall accuracy was only 43.01 %.

Then YOLOv4 is trained by applying the landmarking technique to each image data set. The result is that the accuracy performance has increased by 72.65 %. However, in class C. Carpio and G. Carp still got poor accuracy scores at this step and were still not acceptable (15 and 40.91 %).

Then the subclassing technique is applied. At the final evaluation, C. Carpio red was separated and became a separate class, but it was still grouped as C. Carpio. As a result, the overall accuracy increased to 76.64 %. The accuracy results in class C. Carpio also increased significantly, from 15 to 60 %, and Catla from 80 to 100 %. However, it turns out that this also affects the level of accuracy of other classes that are lower. The Rohu class, originally 100, fell to 86.21 %, and the G. Carp, fell significantly from 40.91 to 13.64 %.

Then the data scale is added during the training process. As a result, the final average accuracy increased to 77.42 %. However, several classes produced low and unacceptable accuracy, namely C. Carpio (40 %), G. Carp (54.55 %), and Catla (70 %). Then the head data is also added. The result turned out to be unexpected because it did not improve accuracy, but slightly lowered it. The overall accuracy average dropped to 76.47 %. Some classes still produce a low and unacceptable level of accuracy, namely G. Carp (27.27 %), C. Carpio (50 %), and Catla (85 %).

The last is the data training trial by applying the class elimination technique. From a series of trials that have been carried out, three classes always get poor and unacceptable accuracy scores; G. Carp, C. Carpio, and Catla. By applying (6), which has been described previously, two classes will be eliminated. Those classes are two classes with the lowest accuracy results (G. Carp and C. Carpio). From the test results, the final accuracy value is excellent and could reach 98.75 %. Even so, the average value of accuracy for each class reaches 95 % or more. The accuracy results of this series of experiments are summarized in **Table 3** and **Fig. 5**. **Fig. 6** shows the results of the evaluation by the confusion matrix, and **Fig. 7** shows the other parameters' results (precision, recall/sensitivity, specificity, and F1 score).

A limitation should also be reported in this work. This study used CiRA-Core software as the main tool to run YOLO as a recognition algorithm. This software has many advantages, such as being easy to use and being integrated (supports ready-to-use technology). However, the main limitation related to this work is that the YOLO algorithm provided is patent/cannot be modified.

Because of that, it is not possible to do a study to modify the YOLO algorithm to improve recognition performance.

Table 3
Experimental results (accuracy)

Class	Detection Accuracy (%)					
	Body (Whole Image)	**Body (with land-marking technique)**	**Body (after subclassing)**	**Body+ Scale**	**Body+ Scale+Head**	**Body (4 classes)**
Catla	65.00	80.00	100.00	70.00	85.00	95.00
C. Carpio (+C. Carpio Red)	0.00	15.00	60.00	40.00	50.00	–
G. Carp	9.09	40.91	13.64	54.55	27.27	–
Mori	96.43	100.00	100.00	100.00	100.00	100.00
Rohu	48.28	100.00	86.21	100.00	96.55	100.00
Silver	39.29	100.00	100.00	100.00	100.00	100.00
Average	43.01	72.65	76.64	77.42	76.47	98.75

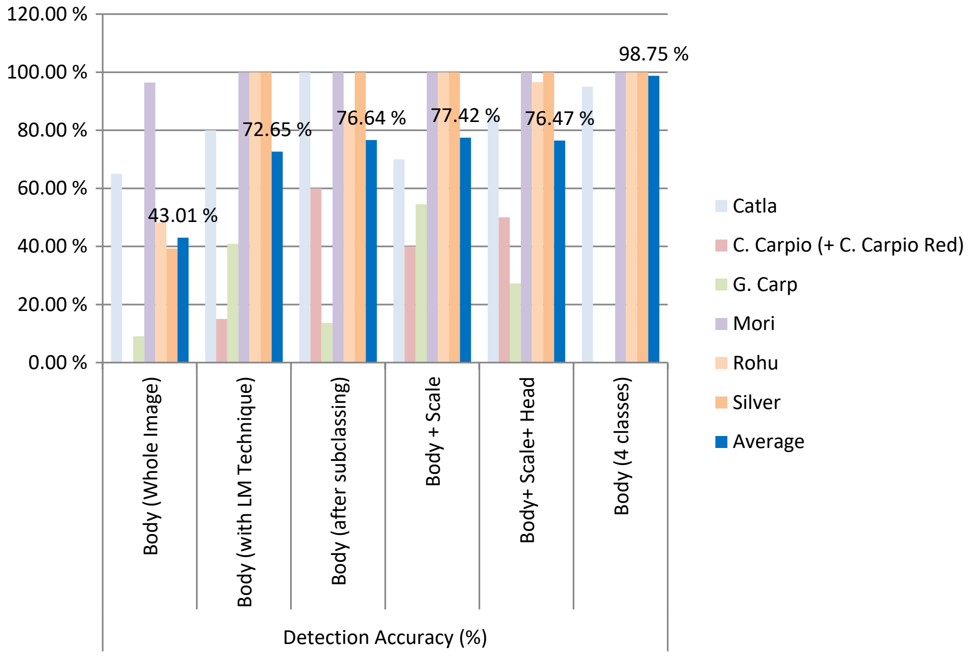


Fig. 5. Experimental results (accuracy)

Another point that should also be reported is the weakness. Suppose referring to the report of the results of this work, it is possible to agree that this study is very good to find out the impact and increase in the use of the techniques used in combination with YOLOv4 on similar and deformed fish objects on the level of accuracy or other performance parameters. However, the only major disadvantage in this study is that the best results could be achieved by eliminating (sacrificing) two classes of fish with low recognition performance. This matter makes this work incapable of fully answering the challenge.

For this reason, this study can be further developed in the future to improve accuracy and other performance parameters for an entire class. It may be done by modifying the YOLO algorithm used (by another tool) or applying image processing techniques that have not been performed in this study.

True Class \ Prediction Class	Catla	C. Carpio	G. Carp	Mori	Rohu	Silver
Catla	13					
C. Carpio		0				
G. Carp			2			
Mori			15	27	1	
Rohu					14	
Silver						11

a

True Class \ Prediction Class	Catla	C. Carpio	G. Carp	Mori	Rohu	Silver
Catla	18					
C. Carpio		4				
G. Carp			9			
Mori		16	13	28		
Rohu					29	
Silver	3	1				28

b

True Class \ Prediction Class	Catla	C. Carpio	G. Carp	Mori	Rohu	Silver
Catla	20					
C. Carpio		12				
G. Carp			3			
Mori		7	19	28	2	
Rohu					25	
Silver						28

c

True Class \ Prediction Class	Catla	C. Carpio	G. Carp	Mori	Rohu	Silver
Catla	14					
C. Carpio		9				
G. Carp		1	12			
Mori		11	10	28		
Rohu					29	
Silver	5					28

d

True Class \ Prediction Class	Catla	C. Carpio	G. Carp	Mori	Rohu	Silver
Catla	18					
C. Carpio		10				
G. Carp			6			
Mori		9	14	28		
Rohu					28	
Silver	2					28

e

True Class \ Prediction Class	Catla	Mori	Rohu	Silver
Catla	19			
Mori		28		
Rohu			29	
Silver				28

f

Total data	
Catla	20
C. Carpio	20
G. Carp	22
Mori	28
Rohu	29
Silver	28

g

Fig. 6. Confusion matrix results for: *a* – training with the whole image; *b* – with landmarking technique; *c* – with subclassing technique; *d* – with the addition of scale image; *e* – with the addition of scale and head image; *f* – with only four classes (after class elimination); *g* – total data

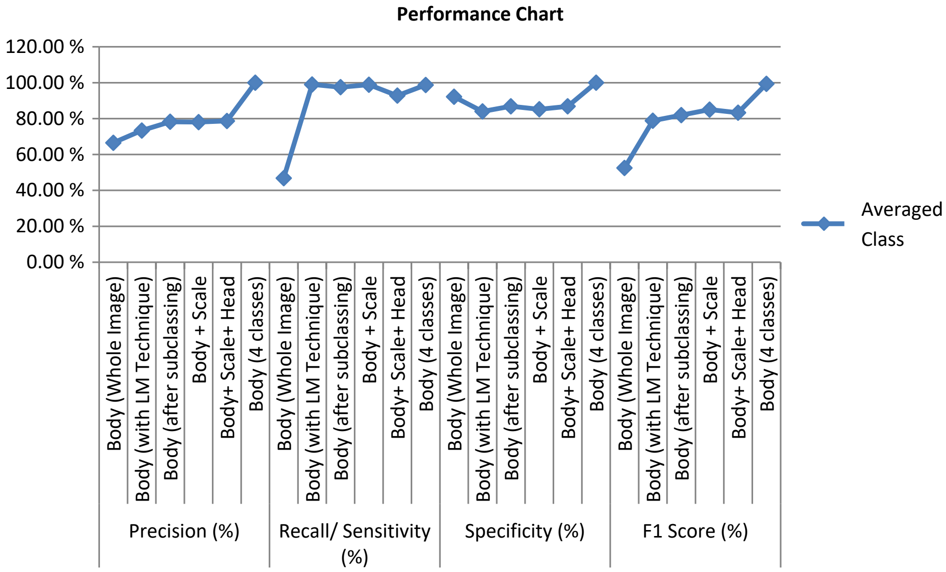


Fig. 7. Experimental results (precision, recall/sensitivity, specificity, and F1 score)

4. Conclusions

A series of experiments shows that YOLOv4 is promising for detecting and classifying fish species with similar and deformed conditions, both of which are characteristic of fish in the aquaculture industry. On the «Fish-Pak» dataset, which contains six species of fish, the accuracy of YOLOv4 is only 43.01 %, but the result rose to 72.65 % with the landmarking technique, and finally rose to 76.64 % with the subclassing technique, then rose to 77.42 % by adding scale data. The accuracy did not improve to 76.47 % by adding head data, and finally, the accuracy rose to 98.75 % with the class elimination technique. However, the accuracy rate can be further improved with complete classes (without class elimination) in future work. Image processing techniques may be improved for pre-processing or also by modifying the recognition algorithm.

Acknowledgments

This work was fully supported by King Mongkut's Institute of Technology Ladkrabang. The authors would also like to thank the College of Advanced Manufacturing Innovation (AMI-KMITL Thailand) for granting the CiRA-Core software license.

References

[1] Ranney, M. A., Velautham, L. (2021). Climate change cognition and education: Given no silver bullet for denial, diverse information-hunks increase global warming acceptance. Current Opinion in Behavioral Sciences, 42, 139–146. doi: https://doi.org/10.1016/j.cobeha.2021.08.001

[2] Bader, F., Rahimifard, S. (2018). Challenges for industrial robot applications in food manufacturing. ISCSIC'18: Proceedings of the 2nd International Symposium on Computer Science and Intelligent Control, 1–8. doi: https://doi.org/10.1145/3284557.3284723

[3] Goncharuk, A. (2015). Food business and food security challenges in research. Journal of Applied Management and Investments, 4 (4), 223–230. Available at: http://www.jami.org.ua/Papers/JAMI_4_4_223-230.pdf

[4] Dos Santos, A. A., Gonçalves, W. N. (2019). Improving pantanal fish species recognition through taxonomic ranks in convolutional neural networks. Ecological Informatics, 53, 100977. doi: https://doi.org/10.1016/j.ecoinf.2019.100977

[5] Alsmadi, M. K., Almarashdeh, I. (2020). A survey on fish classification techniques. Journal of King Saud University – Computer and Information Sciences. doi: https://doi.org/10.1016/j.jksuci.2020.07.005

[6] Zhao, S., Zhang, S., Liu, J., Wang, H., Zhu, J., Li, D., Zhao, R. (2021). Application of machine learning in intelligent fish aquaculture: A review. Aquaculture, 540, 736724. doi: https://doi.org/10.1016/j.aquaculture.2021.736724

[7] Abinaya, N. S. M., Susan, D., Rakesh Kumar, S. (2021). Naive bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries. Ecological Informatics, 61, 101248. doi: https://doi.org/10.1016/j.ecoinf.2021.101248

[8] Ahmed, M. S., Aurpa, T. T., Azad, M. A. K. (2021). Fish disease detection using image based machine learning technique in aquaculture. Journal of King Saud University – Computer and Information Sciences. doi: https://doi.org/10.1016/j.jksuci.2021.05.003

[9] Alshdaifat, N. F. F., Talib, A. Z., Osman, M. A. (2020). Improved deep learning framework for fish segmentation in underwater videos. Ecological Informatics, 59, 101121. doi: https://doi.org/10.1016/j.ecoinf.2020.101121

[10] Mohamed, H. E.-D., Fadl, A., Anas, O., Wageeh, Y., ElMasry, N., Nabil, A., Atia, A. (2020). Msr-yolo: Method to enhance fish detection and tracking in fish farms. Procedia Computer Science, 170, 539–546. doi: https://doi.org/10.1016/j.procs.2020.03.123

[11] Salman, A., Maqbool, S., Khan, A. H., Jalal, A., Shafait, F. (2019). Real-time fish detection in complex backgrounds using probabilistic background modelling. Ecological Informatics, 51, 44–51. doi: https://doi.org/10.1016/j.ecoinf.2019.02.011

[12] Jalal, A., Salman, A., Mian, A., Shortis, M., Shafait, F. (2020). Fish detection and species classification in underwater environments using deep learning with temporal information. Ecological Informatics, 57, 101088. doi: https://doi.org/10.1016/j.ecoinf.2020.101088

[13] Fouad, M. M. M., Zawbaa, H. M., El-Bendary, N., Hassanien, A. E. (2013). Automatic nile tilapia fish classification approach using machine learning techniques. 13th International Conference on Hybrid Intelligent Systems (HIS 2013). doi: https://doi.org/10.1109/HIS.2013.6920477

[14] Kutlu, Y., Iscimen, B., Turan, C. (2017). Multi-stage fish classification system using morphometry. Fresenius Environmental Bulletin, 26 (3), 1910–1916. Available at: https://www.researchgate.net/publication/314284234_MULTI-STAGE_FISH_CLASSIFICATION_SYSTEM_USING_MORPHOMETRY

[15] Hu, J., Li, D., Duan, Q., Han, Y., Chen, G., Si, X. (2012). Fish species classification by color, texture and multi-class support vector machine using computer vision. Computers and Electronics in Agriculture, 88, 133–140. doi: https://doi.org/10.1016/j.compag.2012.07.008

[16] Andayani, U., Wijaya, A., Rahmat, R. F., Siregar, B., Syahputra, M. F. (2019). Fish species classification using probabilistic neural network. Journal of Physics: Conference Series, 1235, 012094. doi: https://doi.org/10.1088/1742-6596/1235/1/012094

[17] Mohammadi Lalabadi, H., Sadeghi, M., Mireei, S. A. (2020). Fish freshness categorization from eyes and gills color features using multi-class artificial neural network and support vector machines. Aquacultural Engineering, 90, 102076. doi: https://doi.org/10.1016/j.aquaeng.2020.102076

[18] Pornpanomchai, C., Lurstwut, B., Leerasakultham, P., Kitiyanan, W. (2013). Shape- and texture-based fish image recognition system. Kasetsart Journal - Natural Science, 47 (4), 624–634. Available at: https://www.researchgate.net/publication/289604551_Shape-_and_texture-based_fish_image_recognition_system

[19] Miyazono, T., Saitoh, T. (2017). Fish species recognition based on cnn using annotated image. IT Convergence and Security 2017, 156–163. doi: https://doi.org/10.1007/978-981-10-6451-7_19

[20] Rekha, B. S., Srinivasan, G. N., Reddy, S. K., Kakwani, D., Bhattad, N. (2020). Fish detection and classification using convolutional neural networks. Computational Vision and Bio-Inspired Computing, 1221–1231. doi: https://doi.org/10.1007/978-3-030-37218-7_128

[21] Taheri-Garavand, A., Nasiri, A., Banan, A., Zhang, Y.-D. (2020). Smart deep learning-based approach for non-destructive freshness diagnosis of common carp fish. Journal of Food Engineering, 278, 109930. doi: https://doi.org/10.1016/j.jfoodeng.2020.109930

[22] Villon, S., Mouillot, D., Chaumont, M., Darling, E. S., Subsol, G., Claverie, T., Villéger, S. (2018). A deep learning method for accurate and fast identification of coral reef fishes in underwater images. Ecological Informatics, 48, 238–244. doi: https://doi.org/10.1016/j.ecoinf.2018.09.007

[23] Labao, A. B., Naval, P. C. (2019). Cascaded deep network systems with linked ensemble components for underwater fish detection in the wild. Ecological Informatics, 52, 103–121. doi: https://doi.org/10.1016/j.ecoinf.2019.05.004

[24] Cai, K., Miao, X., Wang, W., Pang, H., Liu, Y., Song, J. (2020). A modified YOLOv3 model for fish detection based on MobileNetv1 as backbone. Aquacultural Engineering, 91, 102117. doi: https://doi.org/10.1016/j.aquaeng.2020.102117

[25] Villon, S., Iovan, C., Mangeas, M., Claverie, T., Mouillot, D., Villéger, S., Vigliola, L. (2021). Automatic underwater fish species classification with limited data using few-shot learning. Ecological Informatics, 63, 101320. doi: https://doi.org/10.1016/j.ecoinf.2021.101320

[26] Ju, Z., Xue, Y. (2020). Fish species recognition using an improved alexnet model. Optik, 223, 165499. doi: https://doi.org/10.1016/j.ijleo.2020.165499

[27] Qin, H., Li, X., Liang, J., Peng, Y., Zhang, C. (2016). DeepFish: Accurate underwater live fish recognition with a deep architecture. Neurocomputing, 187, 49–58. doi: https://doi.org/10.1016/j.neucom.2015.10.122

[28] Islam, M. A., Howlader, M. R., Habiba, U., Faisal, R. H., Rahman, M. M. (2019). Indigenous fish classification of bangladesh using hybrid features with svm classifier. 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2). doi: https://doi.org/10.1109/IC4ME247184.2019.9036679

[29] Robotham, H., Castillo, J., Bosch, P., Perez-Kallens, J. (2011). A comparison of multi-class support vector machine and classification tree methods for hydroacoustic classification of fish-schools in Chile. Fisheries Research, 111 (3), 170–176. doi: https://doi.org/10.1016/j.fishres.2011.07.010

[30] Kutlu, Y., Reyhaniye, A. N., Turan, C. (2014). Image analysis methods on fish recognition. 2014 22nd Signal Processing and Communications Applications Conference (SIU). doi: https://doi.org/10.1109/SIU.2014.6830503

[31] Badawi, U., Alsmadi, M. (2014). A general fish classification methodology using meta-heuristic algorithm with back propagation classifier. Journal of Theoretical and Applied Information Technology, 66 (3), 803–812. Available at: http://www.jatit.org/volumes/Vol66No3/18Vol66No3.pdf

[32] Liu, Z., Jia, X., Xu, X. (2019). Study of shrimp recognition methods using smart networks. Computers and Electronics in Agriculture, 165, 104926. doi: https://doi.org/10.1016/j.compag.2019.104926

[33] Pettersen, R., Braa, H., Gaweł, B., Letnes, P., Sæther, K., Aas, L. (2019). Detection and classification of Lepeophterius salmonis (Krøyer, 1837) using underwater hyperspectral imaging. Aquacultural Engineering, 87, 102025. doi: https://doi.org/10.1016/j.aquaeng.2019.102025

[34] Liawatimena, S., Heryadi, Y., Lukas, Trisetyarso, A., Wibowo, A., Abbas, B. S., Barlian, E. (2018). A Fish Classification on Images using Transfer Learning and Matlab. 2018 Indonesian Association for Pattern Recognition International Conference (INAPR), 108–112. doi: https://doi.org/10.1109/INAPR.2018.8627007

[35] Shah, S. Z. H., Rauf, H. T., IkramUllah, M., Khalid, M. S., Farooq, M., Fatima, M., Bukhari, S. A. C. (2019). Fish-pak: Fish species dataset from pakistan for visual features based classification. Data in Brief, 27, 104565. doi: https://doi.org/10.1016/j.dib.2019.104565

[36] Rauf, H. T., Lali, M. I. U., Zahoor, S., Shah, S. Z. H., Rehman, A. U., Bukhari, S. A. C. (2019). Visual features based automated identification of fish species using deep convolutional neural networks. Computers and Electronics in Agriculture, 167, 105075. doi: https://doi.org/10.1016/j.compag.2019.105075

[37] Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv. doi: https://doi.org/10.48550/arXiv.2004.10934

[38] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. arXiv. doi: https://doi.org/10.48550/arXiv.1506.02640

[39] Minh, D. H. T., Ienco, D., Gaetano, R., Lalande, N., Ndikumana, E., Osman, F., Maurel, P. (2018). Deep Recurrent Neural Networks for Winter Vegetation Quality Mapping via Multitemporal SAR Sentinel-1. IEEE Geoscience and Remote Sensing Letters, 15 (3), 464–468. doi: https://doi.org/10.1109/LGRS.2018.2794581

Received date 04.11.2021
Accepted date 20.03.2022
Published date 31.03.2022

How to cite: *Kuswantori, A., Suesut, T., Tangsrirat, W., Nunak, N. (2022). Development of object detection and classification with YOLOv4 for similar and structural deformed fish. EUREKA: Physics and Engineering, 2, 154–165. doi: https://doi.org/10.21303/2461-4262.2022.002345*