Preliminary Study on Fish Tracking in Indoor Aquaculture through Deep Learning

Nguyen-Ngoc Huynh[1], Myoungjae Jun[1,*], Hang Thi Phuong Nguyen[1], Choonsung Shin[2] and Hieyong Jeong[1,*]

[1] Department of Artificial Intelligence Convergence, Chonnam National University
77 Yongbong-ro, Buk-gu, Gwangju 61186, Republic of Korea
[2] Graduate School of Culture, Chonnam National University
77 Yongbong-ro, Buk-gu, Gwangju 61186, Republic of Korea
{217019, 208258, cshin, h.jeong}@jnu.ac.kr, 21cnehemiah@gmail.com

Abstract. Due to food shortages, areas such as aquaculture intelligent farms are getting a lot of attention. However, many technologies still need to be solved to automatically control the growth process of randomly and infinitely moving marine organisms. Therefore, the economics of aquaculture requires the creation of systems that can do this automatic tracking of fish behavior. This paper demonstrates a software approach to fish tracking with deep learning. The result of experiments have nevertheless shown that our process is simply efficient and above all easy to implement in fish farms, although there were some restrictions.

Keywords: YOLO · DeepSORT · object detection · multi-object tracking · fish tracking.

1 Introduction

An important area of computer vision is target tracking. It is extensively employed in other industries, such as video monitoring, and aquaculture may soon follow suit. Fish tracking technology is a crucial tool for behavior observation in aquaculture. Monitoring fish behavior and growth through tracking enables the aquaculture sector to use its resources most. Fish tracking and behavior analysis can also be used to keep an eye on the environment that supports fish growth. Additionally, it can more effectively manage water quality, evaluate the health of cultured fish, and promptly address abnormal behavior. Additionally, precise feeding, disease detection, counting, and tracking of fish can all be done using fish-tracking techniques.

With the increasing demand for seafood resources and the rapid development of deep learning technology, the scale of aquaculture is constantly expanding. Therefore, the application of technology in practice is very necessary. The development of an application with the ability to detect and track underwater targets in real time greatly supports the aquaculture industry. However, in reality, it is almost difficult for humans to distinguish the fish due to its small size, transparent body, easy deformation, motion blur, and similar shapes [14]. Since then, analyzing fish behavior from videos collected from fish farms is difficult even it can be called "impossible" in some ways.

In addition, the other three complex problems in fish tracking are occlusion, background interference, and morphology [7]. The first two problems frequently lead to disrupted trajectories by conflicting with the fish's identification [15], [13]. The third circumstance results in poor tracking precision. (shown in Figure. 5).

Technologies for manual tracking are time-consuming and ineffective. To describe the actions of a single experimental object, extensive manual observations and labeling of image features are required. With improvements in computer performance and the effective use of convolutional neural networks in computer vision, learning-based computer vision technology offers a promising means of automating manual tracking [8]. The technique has most recently been used to automate various item-tracking systems.

Numerous manual observations are needed to completely characterize the actions of a single experimental object, labeling image features, and image features as well. Due to improvements in computer performance and the successful use of convolutional neural networks in computer vision, learning-based computer vision technology gives a workable method to automate manual tracking. The approach was most recently used to automate various item-tracking processes [3], [12], [2]. As object detection technology has advanced, tracking-by-detection technology has become the industry standard for multiple objects online monitoring because of its simple design and robust implementation. In the present study, we simulate a deep learning-based fish monitoring system. Finally, we assign the unique id to each fish during observation using deepSORT [11].

2 Related Works

Fish have been detected using a variety of techniques in order to keep track of their number and size. Various object detection techniques have been used to detect fish [1]. And to track fish size and count image processing and computer vision systems are considered. In order to track fish movement, tracking algorithms like optical flow and frame subtraction are done.

Many methods have been reported to detect fish to maintain tabs on their abundance and size. Fish detection was performed using a variety of object detection methods [1]. Technologies based on computer vision were considered to recognize fish size and number. In addition, tracking technologies like optical flow and frame subtraction were used to monitor the movement of fish.

Youssed Wageeh [10] proposed using euclidean tracking in fish farms to detect fish. Firstly, he used the Image Enhancement algorithm to improve unclear images. Then, the object detection algorithm and euclidean for tracking were used. However, pond fish videos had many disadvantages, such as blurred, low-quality videos resulting in poor recognition and switched ID (identification). Like Weiran Li [5], he introduced four sub-branches for object detection and Re-ID object extraction, extending ResNet-101 as the framework for object map ex-

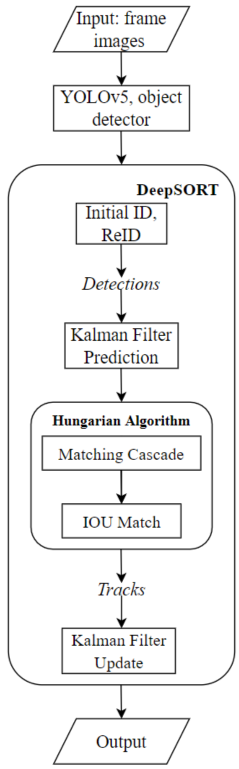


Fig. 1. The mechanism YOLOv5-DeepSORT operates. To find the target, YOLOv5 is provided input images (predict the bounding box of objects). Then, from the output of YOLOv5, REID requires a well-distinguishing feature embedding. Finally, the Kalman filter predicts the trajectory of each target. For the target's trajectory, the outputs of the YOLOv5 and Kalman filters are compared with the Hungarian algorithm for matching.

traction. However, long-term tracking performance still needed to be improved while swimming toward bright spots or pond shores.

According to our survey, improving the method of tracking multiple fish simultaneously through the software approach is necessary [6].

3 Methods

This section describes the general organization of the suggested fish-tracking method based on deep learning. First, the YOLO model is used for fish detection, and the DeepSort model is applied for fish tracking. Then, we mainly concentrate on presenting the DeepSORT tracking component, which includes using the ID appearance feature model, cascade matching, and IoU matching.

Fig. 2. Fish data set samples.

The difficulty of detection and tracking caused by dense targets and the re-identification of targets in fish farm chores led to the development of an intelligent fish recognition and tracking system. In Figure 1, the overall frame diagram is displayed. First, the location information (i.e., the center's x, y, width, and height), categorization details (class), and confidence of each target frame are acquired using YOLOv5 as the detector to extract feature information. Next, the detected results are entered into DeepSORT, where a previously predicted trajectory T_{pre} is obtained using the Kalman Filter predict module. The Hungarian Algorithm is then used to determine the degree to which the detected result of the current frame (D_t) matches the predicted track (T_{pre}). Finally, incorrect tracks are eliminated to complete fish tracking during tracking, and corrected tracks ($Dmatch/Tmatch$) are updated via the Kalman Filter update module.

4 Results

In this experiment, the Goldfish images are collected from YouTube videos. First, it has split frame by frame, and then we select high-quality images manually, selecting from blurry photos that do not show the subject well. Finally, the automatic decision algorithm helps us choose images from some candidates.

The 2,835 images that compose this research's data set were collected. The 595 images among the total images were chosen (ratio: 8:2) to make the test set, and the remaining images served as the training set. In Figure 2, an example of the dataset is shown.

Whole images with a resolution of 640 × 640 pixels are annotated by the Make-sense [9], the annotation tool, which is a free online tool for labeling photos in Computer Vision projects. The TXT annotation file in YOLO format is made. Moreover, it is used to feed into the YOLOv5 for training object detection and DeepSORT for tracking. A sample example of this tool is shown in Figure 3.

Table 1 truly configures the experiment as below: It is described in terms of hardware configuration software and libraries version.

The videos are tested with the trained model and show the effectiveness of YOLOv5 [4] and DeepSORT [11] for detection and tracking. The results are

Fig. 3. Labelled ground truth.

shown in Figure 4. The object prediction process is continuous in each frame. Furthermore, it helps DeepSORT get the coordinate information of the object, which is very useful in initializing the object's id because when the object detector misses, it will initialize multiple ids for one object.

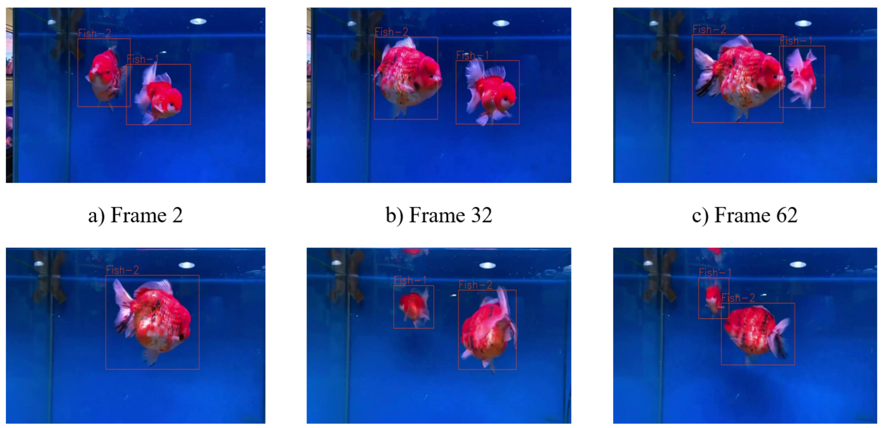
As the results show, the unique id for each object remains constant every 30 frames (the number of frames that DeepSORT tracks the object, if it disappears, more than 30 frames continue to be tracked and initialize a new id when it appears).

This study successfully recognized and tracked fish with unique identifiers, although the suggested method still had certain drawbacks.

1. Compared to genuine footage from the aquaculture farm, the data set used in this study is less varied in terms of posture, background, and image quality.
2. In the future, software that incorporates our algorithm for live video viewing of farms should be developed. Fish farmers have benefited immensely from the software's ability to tell them when fish are acting abnormally and to monitor the situation.
3. The number of fish in the video is limited to less than three fish to achieve good results with only one unique id during tracking.

Table 1. Experimental configuration.

Configuration	Parameters
Processor	Intel(R) Core(TM) i9-10900K
GPU	NVIDIA GeForce RTX 2070 SUPER
Operating system	Windows 10
Accelerated environment	CUDA 11.3 & CUDNN 8.2.1
Code editor	Visual Studio Code 1.72.1
Libraries	Opencv 4.6.0, torch 1.12.1, numpy 1.21.6



a) Frame 2 b) Frame 32 c) Frame 62

e) Frame 92 f) Frame 122 e) Frame 152

Fig. 4. The efficiency of DeepSORT in different frames.

5 Consclusion and future works

To predict as well as follow fish in this study, we could put forward for consideration in deep learning technological solution. YOLOv5 was an object detector to predict the coordinate of fish in the frame by frame, and DeepSORT, as a tracker, gave the unique id for each fish. We used the Goldfish data set to demonstrate the effectiveness of our method. This study motivated us to do future research with videos from fish farms. The proposed method was simple and easy to deploy on the farm. Through this method, fish detection and tracking could be monitored conveniently.

Acknowledgements This research was supported by Korea Institute of Marine Science & Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries, Korea (No. 20220596, Development of Digital Flow - through Aquaculture System).

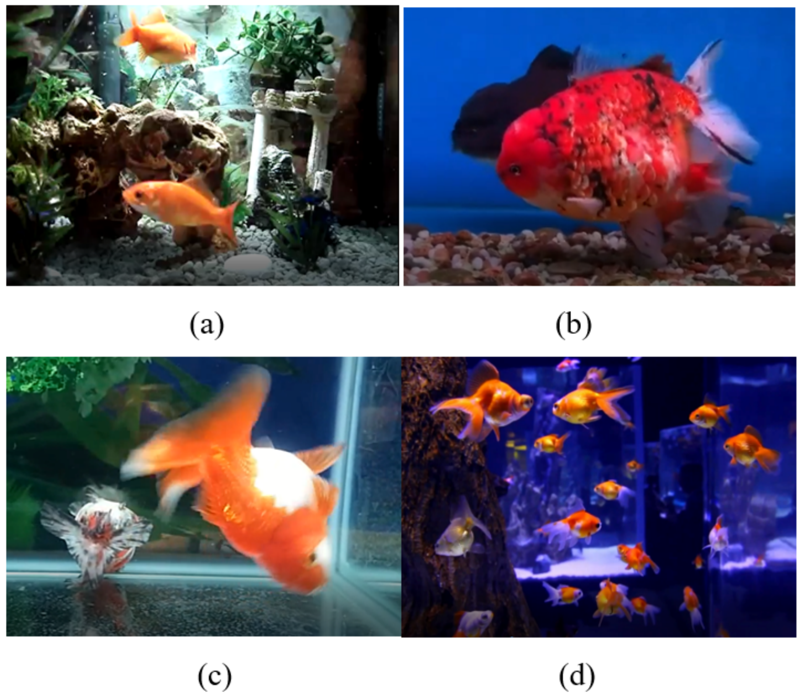


Fig. 5. Example of challenges. **(a)** Background complicated, **(b)** Overlap objects. **(c)** Inverted poses. **(d)** Similar objects.

The authors would like to express a huge thank you to the BK21 Plus program at Chonnam National University, through the National Research Foundation, funded by the Ministry of Education of Korea.

References

1. Duggal, S., Manik, S., Ghai, M.: Amalgamation of video description and multiple object localization using single deep learning model. In: Proceedings of the 9th International Conference on Signal Processing Systems. pp. 109–115 (2017)
2. Egi, Y., Hajyzadeh, M., Eyceyurt, E.: Drone-computer communication based tomato generative organ counting model using yolo v5 and deep-sort. Agriculture **12**(9), 1290 (2022)
3. Hou, X., Wang, Y., Chau, L.P.: Vehicle tracking using deep sort with low confidence track filtering. In: 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). pp. 1–6. IEEE (2019)
4. Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, Kwon, Y., Michael, K., TaoXie, Fang, J., imyhxy, Lorna, Yifu), Wong, C., V, A., Montes, D., Wang, Z., Fati, C., Nadar, J., Laughing, UnglvK-

itDe, Sonck, V., tkianai, yxNONG, Skalski, P., Hogan, A., Nair, D., Strobel, M., Jain, M.: ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation (Nov 2022). https://doi.org/10.5281/zenodo.7347926, https://doi.org/10.5281/zenodo.7347926
5. Li, W., Li, F., Li, Z.: Cmftnet: Multiple fish tracking based on counterpoised jointnet. Computers and Electronics in Agriculture **198**, 107018 (2022)
6. Luo, J., Han, Y., Fan, L.: Underwater acoustic target tracking: A review. Sensors **18**(1), 112 (2018)
7. Mei, Y., Sun, B., Li, D., Yu, H., Qin, H., Liu, H., Yan, N., Chen, Y.: Recent advances of target tracking applications in aquaculture with emphasis on fish. Computers and Electronics in Agriculture **201**, 107335 (2022)
8. Ravoor, P.C., Sudarshan, T.: Deep learning methods for multi-species animal reidentification and tracking–a survey. Computer Science Review **38**, 100289 (2020)
9. Skalski, P.: Make Sense. https://github.com/SkalskiP/make-sense/ (2019)
10. Wageeh, Y., Mohamed, H.E.D., Fadl, A., Anas, O., ElMasry, N., Nabil, A., Atia, A.: Yolo fish detection with euclidean tracking in fish farms. Journal of Ambient Intelligence and Humanized Computing **12**(1), 5–12 (2021)
11. Wojke, N., Bewley, A., Paulus, D.: Simple online and realtime tracking with a deep association metric. CoRR **abs/1703.07402** (2017), http://arxiv.org/abs/1703.07402
12. Wu, F., Jin, G., Gao, M., HE, Z., Yang, Y.: Helmet detection based on improved yolo v3 deep model. In: 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC). pp. 363–368 (2019). https://doi.org/10.1109/ICNSC.2019.8743246
13. Xu, W., Matzner, S.: Underwater fish detection using deep learning for water power applications. In: 2018 International conference on computational science and computational intelligence (CSCI). pp. 313–318. IEEE (2018)
14. Yang, L., Liu, Y., Yu, H., Fang, X., Song, L., Li, D., Chen, Y.: Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior analysis: a review. Archives of Computational Methods in Engineering **28**(4), 2785–2816 (2021)
15. Yang, X., Zhang, S., Liu, J., Gao, Q., Dong, S., Zhou, C.: Deep learning for smart fish farming: applications, opportunities and challenges. Reviews in Aquaculture **13**(1), 66–90 (2021)