

Review

# Application of Deep Learning-Based Object Detection Techniques in Fish Aquaculture: A Review

Hanchi Liu <sup>1</sup>, Xin Ma <sup>1,\*</sup>, Yining Yu <sup>2</sup>, Liang Wang <sup>2</sup> and Lin Hao <sup>3</sup>

<sup>1</sup> School of Control Science and Engineering, Shandong University, Jinan 250061, China; liuhanchi@mail.sdu.edu.cn

<sup>2</sup> Shandong Marine Group Ltd., Jinan 250000, China; yuyn@sdmg.com.cn (Y.Y.); wangla@sdmg.com.cn (L.W.)

<sup>3</sup> Shandong Deep Sea Green Farming Ltd., Qingdao 266000, China; haolin@sdmg.com.cn

\* Correspondence: maxin@sdu.edu.cn

**Abstract:** Automated monitoring and analysis of fish's growth status and behaviors can help scientific aquaculture management and reduce severe losses due to diseases or overfeeding. With developments in machine vision and deep learning (DL) techniques, DL-based object detection techniques have been extensively applied in aquaculture with the advantage of simultaneously classifying and localizing fish of interest in images. This study reviews the relevant research status of DL-based object detection techniques in fish counting, body length measurement, and individual behavior analysis in aquaculture. The research status is summarized from two aspects: image and video analysis. Moreover, the relevant technical details of DL-based object detection techniques applied to aquaculture are also summarized, including the dataset, image preprocessing methods, typical DL-based object detection algorithms, and evaluation metrics. Finally, the challenges and potential trends of DL-based object detection techniques in aquaculture are concluded and discussed. The review shows that generic DL-based object detection architectures have played important roles in aquaculture.

**Keywords:** deep learning; object detection; aquaculture; machine vision



**Citation:** Liu, H.; Ma, X.; Yu, Y.; Wang, L.; Hao, L. Application of Deep Learning-Based Object Detection Techniques in Fish Aquaculture: A Review. *J. Mar. Sci. Eng.* **2023**, *11*, 867. <https://doi.org/10.3390/jmse11040867>

Academic Editor: Rafael Morales

Received: 21 March 2023

Revised: 14 April 2023

Accepted: 16 April 2023

Published: 20 April 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

As an essential aspect of fisheries, aquaculture has become the fastest-growing technology in global food production [1–3]. Along with the rapid growth of the population and increased human demand for fish consumption worldwide, the pressure on the aquaculture industry will continue to increase [4,5]. Therefore, improving the production and quality of aquaculture is urgently needed. Monitoring and analyzing the growth status and behaviors of fish at different growth stages can help managers in aquaculture accurately grasp the changes in fish abundance, body length, and behaviors, thereby optimizing aquaculture management, reducing the risk of severe failure due to fish diseases or overfeeding, and improving the productivity and profitability of aquaculture [6,7]. However, traditional methods for analyzing the growth status and behavior of fish primarily depend on manual observation or measurement [8,9]. Therefore, it is not only time- and labour-intensive, but also may harm the welfare and health of fish.

Machine vision-based non-invasive testing techniques have the advantage of automatically obtaining high-quality data while avoiding damaging the normal growth of fish, which have gradually replaced traditional manual methods and are widely applied in aquaculture [10,11]. However, machine vision-based data analysis related to aquaculture faces various challenges [12,13]. On the one hand, the challenges come from the complex environment, such as illumination changes, low contrast, and water turbidity. On the other hand, the challenges come from the complex fish characteristics, such as scale variation, similar appearance, deformation during movement, and occlusion in the high-density

scenario. Therefore, developing automatic and reliable data analysis methods based on machine vision is essential for fish monitoring in aquaculture.

During the past decades, numerous traditional image-processing methods and machine-learning techniques have been used for vision-based data analysis in aquaculture [14–16]. However, such methods rely heavily on handcrafted feature extraction, which easily leads to the omission of critical features of fish in complex and diverse real-world environments. Recently, due to the powerful adaptive feature extraction and nonlinear mapping capabilities, deep learning (DL) techniques have played an essential role in visual information processing [17,18]. Especially, DL-based object detection techniques, as an essential branch of DL, have the unique advantage of simultaneously identifying the category of the interest object of interest and locating the position of the object using a bounding box, have been widely used in various fields [19], such as face recognition [20,21], text detection [22,23], pedestrian recognition [24,25], and vehicle detection [26,27]. The superiority of DL-based object detection techniques has also made them gain increasing attention in fish monitoring in aquaculture, including the field of fish counting [28,29], body length measurement [30,31], and individual behavior analysis [32,33]. The flow chart of the typical application of DL-based object detection techniques in aquaculture is shown in Figure 1. In addition, image preprocessing techniques are often used in advance to tackle the limitations of the quantity and quality of the data.

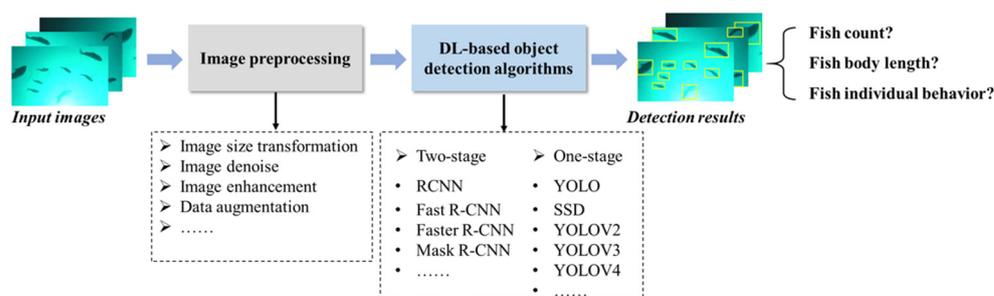


Figure 1. Flow chart of the application of DL-based object detection techniques in aquaculture.

There have been several reviews on the application of DL in aquaculture [12,13], in which they comprehensively concerned with the DL techniques in aquaculture, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs). However, no works have thoroughly explored or studied the application of DL-based object detection techniques in fish aquaculture. Therefore, this paper applies DL-based object detection techniques for fish counting, body length measurement, and individual behavior analysis in aquaculture. In this paper, Section 2 briefly introduces the dataset sources and image preprocessing methods. Section 3 compares the typical DL-based object detection algorithms. Section 4 reviews the research status of DL-based object detection techniques in fish counting, body length measurement, and individual behavior analysis in aquaculture. The research status is summarized from two aspects: image and video analysis. Section 5 summarizes the evaluation metrics of the algorithms. Finally, Section 6 discussed the challenges and potential trends of DL-based object detection techniques in aquaculture to provide a reference for developing DL techniques.

## 2. Datasets and Image Preprocessing

Since deep neural networks (DNNs) generally rely on large-scale and high-quality datasets for training to achieve satisfactory performance, datasets are the basis of the DL-based object detection techniques applied to aquaculture.

### 2.1. Datasets

In studies on applying DL-based object detection techniques in aquaculture, data was mainly acquired from online public and on-site datasets constructed by researchers during the experiment.

### 2.1.1. Public Datasets

The public datasets related to the application of DL-based object detection techniques in aquaculture mainly included Fish4Knowledge [34], LifeCLEF2014 [35], LifeCLEF2015 [36], and NOAA [37]. In addition, many other available datasets have the potential to improve the performance and generality of DL-based object detection models in aquaculture, such as NCFM (The Nature Conservancy Fisheries Monitoring) [38] and ImageNet [39] in aquaculture. The characteristics of the datasets are compared in Table 1.

**Table 1.** The public datasets for fish.

Datasets	Total Videos/Images	Annotation	URL
Fish4Knowledge [34]	700,000 underwater videos with 3000 fish species	-	<a href="https://homepages.inf.ed.ac.uk/rbf/Fish4Knowledge/resources.htm">https://homepages.inf.ed.ac.uk/rbf/Fish4Knowledge/resources.htm</a> (accessed on 15 June 2022)
LifeCLEF2014 [35]	1000 underwater videos with 10 fish species	20,000 labeled fish	<a href="https://www.imageclef.org/2014/lifeclef/fish">https://www.imageclef.org/2014/lifeclef/fish</a> (accessed on 15 June 2022)
LifeCLEF2015 [36]	93 underwater videos with 15 fish species	9000 annotations in videos and 20,000 images with fewer labels	<a href="http://www.imageclef.org/lifeclef/2015/fish">http://www.imageclef.org/lifeclef/2015/fish</a> (accessed on 15 June 2022)
NOAA [37]	929 underwater images	1005 labeled fish	<a href="https://swfscdata.nmfs.noaa.gov/labeled-fishes-in-the-wild/">https://swfscdata.nmfs.noaa.gov/labeled-fishes-in-the-wild/</a> (accessed on 15 June 2022)
NCFM [38]	3777 images	-	<a href="https://www.kaggle.com/c/the-nature-conservancy-fisheries-monitoring">https://www.kaggle.com/c/the-nature-conservancy-fisheries-monitoring</a> (accessed on 5 April 2023)
ImageNet [39]	over 14 million images	-	<a href="http://www.image-net.org/">http://www.image-net.org/</a> (accessed on 5 April 2023)

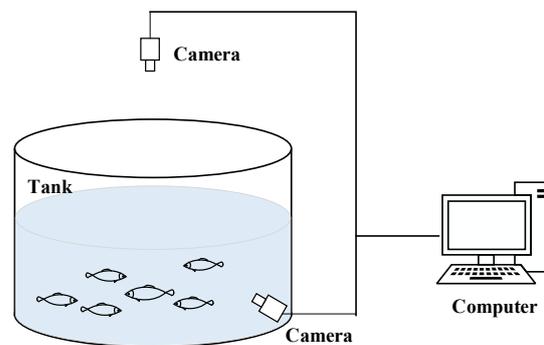
- Fish4Knowledge: Fish4Knowledge dataset was supported by the European Commission and developed jointly by a project team from the University of Edinburgh, Academia Sinica, and other groups, with the main aim of assisting in marine ecosystem research. This dataset contains about 700,000 underwater video clips with 10 min of monitoring Taiwan coral reefs over five years, which can be used for fish identification, detection, and tracking in images and videos. However, this dataset has unbalanced the number of fish categories (over 3000 fish species).
- LifeCLEF2014: LifeCLEF2014 was built based on the Fish4-Knowledge dataset by a project team from the University of Catania and Edinburgh, which contains about 1000 videos with 10 fish species. Approximately 20,000 fish in this dataset were labeled with species in the videos. However, this dataset also suffers from an unbalanced fish population.
- LifeCLEF2015: LifeCLEF2015 was also built based on the Fish4Knowledge dataset by a project team from the University of Catania and Edinburgh, which contains about 93 underwater videos with 15 fish species. This dataset contains about 9000 annotations (bounding boxes and species) in videos and 20,000 images with fewer labels. Compared with LifeCLEF2014, the image and videos in LifeCLEF2015 present a noisier, blurred, and poorer illumination environment.
- NOAA: NOAA dataset was developed by the National Oceanic and Atmospheric Administration (NOAA) during rockfish surveys in the Southern California Bight. It was collected by a digital camera deployed on a remotely operated vehicle (ROV). This dataset contains 929 images and 1005 annotations (locations and bounding rectangles). The challenges of the dataset include variations in the appearance and size of fish, small particles in the water, various swimming speeds and directions of fish, fish hidden behind the rocks or in crevices, and some indistinct fish-like objects.

- NCFM: The NCFM dataset was supported by the worldwide competition of “The Nature Conservancy Fisheries Monitoring” hosted by Kaggle, which contains about 3777 fish images. The images were taken by cameras installed on different fishing boats. The light variation, complex background, and occlusion of fish in the dataset make fish recognition very challenging.
- ImageNet: ImageNet was initiated by the team of Fei-Fei Li at Stanford University, which contains over 14 million images. ImageNet is an image dataset organized by the WordNet hierarchy, in which each node is connected to hundreds or thousands of images.

The main advantages of the public dataset are that they are easy to obtain, contain large-scale available data, and are rich in variety. On the other hand, the common shortcomings of these datasets are unbalanced fish species, poor environment, and distortion in the images.

### 2.1.2. On-Site Datasets

To analyze the growth status of fish more accurately in the study area, researchers usually collect data in the field to construct a real dataset. Figure 2 illustrates a typical on-site data acquisition system. The on-site data of the related research papers were primarily derived from the ocean, cages, recirculating aquaculture systems (RAS), and fish tanks. Researchers can choose the type and location of the camera according to their research needs. One primary data acquisition method for fish counting is collecting underwater data using an underwater camera deployed on a remotely operated vehicle (ROV) or other underwater monitoring platforms [40]. Another is to deploy a camera above the water surface of the container for data collection [41]. However, acquiring 2D information by a single camera inevitably suffers from the occlusion of itself or other objects. Compared to 2D imaging systems, 3D imaging systems can obtain more comprehensive information and resolve occlusion problems [42]. In addition, autonomous underwater vehicles (AUVs) could be used for information collection on fish in the future [43,44]. For fish body length measurement, some researchers placed the monocular camera at a fixed distance above the captured fish to obtain its 2D information.



**Figure 2.** A typical on-site data acquisition system.

Moreover, binocular cameras have been widely used to collect 3D information on underwater fish without constraining their movement. For fish behavior identification, researchers commonly fix the camera above the water surface of the container to collect the movement videos of fish [32,33]. However, visible light cameras are susceptible to light in aquaculture. On the other hand, the image quality of the infrared reflection system is not affected by light intensity, which allows for the monitoring of fish behavior in dark, poorly lit, and obscured conditions [45,46].

The main advantages of the on-site dataset are that the researchers can choose the appropriate research objects, collection equipment, collection location, and collection time according to their research requirements. Moreover, researchers can adjust the quality and quantity of collected data to avoid poor-quality images and imbalances in the dataset. However, there are also some problems. For example, the total number of on-site datasets

and species categories tends to be relatively small and varied due to the time data collection takes and the required resources.

### 2.2. Image Preprocessing

Since the collected dataset is usually affected by issues related to quality and quantity, image preprocessing has become a critical step in applying DL-based object detection techniques to aquaculture. Common image preprocessing approaches include image size transformation, image denoise, image enhancement, and data augmentation.

- Image size transformation: Image size transformation (such as image cropping and resizing) is the most common image preprocessing method, which can reduce the computation or meet the input requirements of DNN models by adjusting images of different sizes to a uniform size [47].
- Image denoise: Strong noise in images may obscure the discriminative features of the object and disturb the feature extraction of DL-based algorithms. In some related studies, gaussian filtering [48] and median filtering [32] have removed image noise.
- Image enhancement: Blurred and low-contrast images will lose some detail in the target. Image enhancement strategies such as linearization, contrast-limited adaptive histogram equalization (CLAHE), Retinex, and discrete wavelet transform (DWT) can recover high-quality images from low-resolution data [29,49]. In addition, the DL-based image enhancement approaches have received increasing attention in the aquatic field [50,51].
- Data augmentation: Data augmentation techniques can extend the number of training samples to avoid overfitting DNN models to small amounts of training samples. Data augmentation methods include rotation, cropping, flipping, and Cutmix [52]. In recent years, generative adversarial networks (GAN) that can generate pseudo-images based on input noise have been widely used in data augmentation [53].

### 3. Typical DL-Based Object Detection Algorithms

DL-based object detection techniques have the unique advantage of simultaneously identifying the category of the object of interest and locating the object’s position using a bounding box, which has been widely used in fish detection. Mainstream DL-based object detection algorithms can be roughly divided into two categories: two-stage object detection (region-based) and one-stage object detection (regression-based) algorithms. The main difference is whether they include a step for the region proposal. This section discusses some typical object detection algorithms widely used in aquaculture. Figure 3 shows the basic flow chart of two-stage and one-stage object detection algorithms. Table 2 summarizes the advantages and disadvantages of the typical DL-based object algorithms.

**Table 2.** The advantages and disadvantages of typical two-stage and one-stage object detection algorithms.

Object Detection Algorithms	Advantages	Disadvantages
R-CNN [54]	Introduced DL to object detection for the first time	Slow training process; computer resource heavy
Two-stage object algorithms	Use ROI pooling to change the scale of the feature	Time-consuming selective search for region proposals
Faster R-CNN [56]	End-to-end training	Low detection accuracy for multi-scale and small object
Mask R-CNN [57]	Accurate instance segmentation and high detection accuracy	Expensive instance segmentation

Table 2. Cont.

Object Detection Algorithms	Advantages	Disadvantages
YOLO [58]	A novel one-stage detection algorithm, the detection speed is fast	Low detection accuracy; weak generalization ability
SSD [59]	Combines regression and anchor mechanisms	Loss of small object features
YOLOV2 [60]	Further improved detection speed and improved the recall rate	Poor detection accuracy for small objects.
YOLOV3 [61]	Improves the detection accuracy of small objects	Low recall rate
YOLOV4 [62]	It incorporates a variety of tuning techniques	Largely unchanged detection model

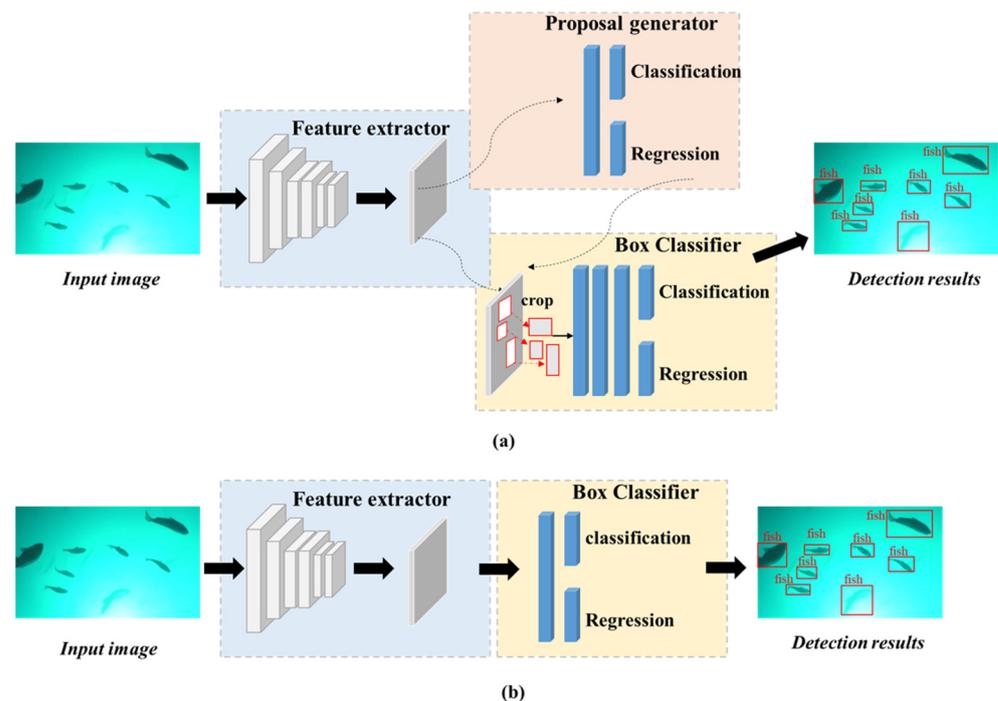


Figure 3. The basic flow chart of the two-stage (a) and one-stage (b) object detection algorithms.

### 3.1. Two-Stage Object Detection Algorithms

Two-stage object detection algorithms (region-based) generate a region proposal based on the input image and then classify the regions using the classifier. Two-stage object detection algorithms have high accuracy, while their common disadvantage is their slow detection speed. Typical two-stage object detection algorithms mainly include R-CNN [54], Fast R-CNN [55], Faster R-CNN [56], and Mask R-CNN [57].

- R-CNN: Girshick et al. proposed the R-CNN algorithm in 2014, introducing DL into object detection. R-CNN first uses the selective search to generate region proposals of the input image, then inputs the region proposals into a convolutional neural network (CNN) to extract features. Finally, it classifies the features using SVM and fine-tunes the bounding regions via bounding-box regression and greedy non-maximal suppression (NMS). Although R-CNN pushes object detection into the era of DNN, it occupies a lot of computing resources. Moreover, it easily leads to missing data information caused by the norm of the region proposals.
- Fast R-CNN: To further improve the training speed of the algorithm and reduce the occupation of computing resources, Fast R-CNN inputs the whole image into a

CNN for feature extraction and introduces a region of interest (ROI) pooling layer for scale transformation of the region proposal features with different sizes. However, the region proposal generation mechanism based on the selective search is still the bottleneck restricting the further improvement of the detection speed.

- **Faster R-CNN:** Faster R-CNN enables end-to-end detection, which innovatively proposes a regional proposal network (RPN) to replace the selective search, thereby significantly improving the generation speed of the detection bounding box. The performance of the Faster R-CNN for fish detection with an accuracy of 82.7% exceeds Fast R-CNN on the ImageCLEF dataset [63,64]. In addition, compared with ZF Net and CNN-M, the Faster R-CNN algorithm based on VGG-16 achieved the best results for fish detection with a mean average precision (mAP) of 82.4% on underwater images obtained from remote underwater video stations [65].
- **Mask R-CNN:** Mask R-CNN is an extension of Faster R-CNN by adding an object segmentation branch parallel to object classification and bounding box regression branches, which can perform object detection and instance segmentation simultaneously using one network structure.

### 3.2. One-Stage Object Detection Algorithms

One-stage object detection algorithms (regression-based) abandon the time-consuming region proposal process and use a DNN to directly generate objects' class probability and position coordinates based on the input image. The main advantage of one-stage object detection algorithms is the fast detection speed, but their detection accuracy is generally lower than the two-stage object detection algorithms. Typical one-stage object detection algorithms include the YOLO series (Your Only Look Once) and the SSD (single-shot multi-box detector).

- **YOLO:** Redmon et al. proposed YOLO in 2016 [58], an end-to-end one-stage DNN algorithm. It divides the whole input image into  $S \times S$  grids. Then, it performs classification and bounding box detection of the target on each grid, which can directly regress the location and type of the target from the input image. YOLO achieved a fish detection accuracy of 93% and a detection speed of 16.7 frames per second (FPS) on the NOAA dataset, which can process noisy, dim-light, and hazy underwater images and outperformed the HOG classifier-based algorithm and SVM classifier [66].
- **SSD:** Aiming to overcome the low detection accuracy of YOLO for small objects, Liu et al. proposed the SSD algorithm in 2016 [59]. SSD combines multi-scale feature maps with the anchor mechanism in the Faster R-CNN and replaces the fully connected layer in YOLO with a convolutional layer, ensuring the detection speed while meeting the detection accuracy.
- **YOLOV2:** Although YOLO achieves real-time object detection, it suffers from many localization errors. To obtain higher detection accuracy, YOLOV2 [60] introduces some new technologies based on YOLOV1, including batch normalization, a high-resolution classifier, and bounding box prediction based on K-Means clustering and multi-scale training.
- **YOLOV3:** Redmon et al. utilized the residual network, feature pyramid network (FPN), and binary cross-entropy loss to upgrade YOLOV2 to YOLOV3 [61], making it suitable for multi-size objects. YOLOV3 achieved a mAP of 53.92% on MHK and hydropower underwater dataset, which can distinguish bubbles, debris, and fish [67].
- **YOLOV4:** YOLOV4 [62] applies a new backbone network and combines spatial pyramid pooling and path aggregation network (PAN) for feature fusion, which achieves higher detection performance.

## 4. Application of DL-Based Object Detection Techniques in Fish Aquaculture

Accurate monitoring and analysis of fish growth conditions and behaviors are essential for scientific fish management, breeding density control, and water quality monitoring in aquaculture [12,13]. The rapid development of machine vision and DL-based object

detection techniques has made it possible to quickly classify and localize critical features of fish from images, enabling more accurate information for control decisions of aquaculture. This paper reviewed the application of DL-based object detection techniques in aquaculture, including fish counting, body length measurement, and individual behavior analysis, to guide the researcher in aquaculture.

#### 4.1. Fish Counting

In intensive aquaculture, the regular number statistics of fish can provide aquaculture managers with reliable quantitative information to observe fish population changes over time, thereby optimizing breeding strategies, controlling breeding density, and establishing optimizing marketing plans [68]. Numerous DL approaches have been studied to fish counting in images, mainly divided into detection-based and density estimation-based [69,70]. Density estimation-based fish counting methods maps can obtain the total number of fish in the image by integrating the predicted density map. However, the main limitation of these methods is the inability to distinguish individual fish. DL-based object detection techniques can automatically extract discriminative features of the object of interest in images and classify and localize objects simultaneously. These have been the focus of many researchers developing fish counting algorithms and offering new opportunities for fish aquaculture. Table 3 details the relevant fish counting studies using DL-based object techniques. Image and video analysis are the two main fish-counting methods based on object detection techniques. This paper divides the related studies according to whether the processed data comes only from individual images or the continuous video sequence.

**Table 3.** Summary of DL-based object detection techniques applied in fish counting.

Data	References	Approaches	Fish Species/ Public Dataset	Data Preprocessing	Results
Image	Li et al. [64]	Faster R-CNN	LifeCLEF2014	N/A	mAP = 82.7%; Time = 0.102 s/im
	Mandal et al. [65]	Faster R-CNN	50 species of fish and crustaceans	N/A	mAP = 82.4% FPS = 5
	Sung et al. [66]	YOLO	NOAA dataset	N/A	Sensitivity = 93%; IOU = 65.2%; FPS = 16.7
	Xu et al. [67]	YOLOV3	Unknown species;	N/A	mAP = 53.9%
	Ditria et al. [71]	Mask R-CNN	Luderick	N/A	F1-Score = 92.4% mAP50 = 92.5%
	Labao and Naval Jr [72]	Improved Faster R-CNN	Unknown species	N/A	Precision = 53.29% Recall = 37.77% F1-Score = 44.21%
	Li et al. [73]	Improved YOLOV5	Takifugu rubripes	Resize	Precision = 97.53% Recall = 98.09%
	Li et al. [40]	Improved YOLO	Unknown species	CLAHE; Rotation; Brightness transformation	Precision = 89% Recall = 73% IOU = 66% FPS = 122
	Cai et al. [47]	Improved YOLOV3	Takifugu rub ripe	Resize	AP = 78.63%

Table 3. Cont.

Data	References	Approaches	Fish Species/ Public Dataset	Data Preprocessing	Results
Video	Salman et al. [74]	Improved Faster R-CNN	Fish4Knowledge with Complex Scenes (FCS); LifeCLEF 2015	N/A	F1-Score = 87.44%(FCS) F1-Score = 80.02% (LifeCLEF 2015)
	Ben et al. [28]	Improved Faster R-CNN	LifeCLEF 2015	N/A	F1-Score = 83.16% mAP = 73.69%
	Levy et al. [75]	RetnaNet + SORT	Unknown species	Resize	Precision = 74%
	Arvind et al. [41]	Mask R-CNN + GOTURN	Ornamental fish	Resize	Precision = 99.81% Recall = 83.112% F1-Score = 90.70% FPS = 16
	Mohamed et al. [76]	YOLO + optical flow	Golden fish	Multi-Scale Retinex	Detected an average of 8 fish from above and 3 fish from underwater
	Liu et al. [29]	Improved YOLOV4 + Kalman filter	Sebastodes fuscenscens; Asteroidea. Hexagrammos otakii	Color compensation. CLAHE; Resize	ACC = 95.6% Recall = 93.3% IOU = 83% FPS = 33 MOTA = 83.6% IDF1 = 83.2% ID Switches = 59%

#### 4.1.1. Image-Based Fish Counting

Intuitively, the number of fish in the image can be calculated by counting the bounding boxes predicted by object detection algorithms. At present, typical two-stage, and one-stage DL-based object detection algorithms such as Faster R-CNN [64,65], YOLO [66], and YOLOV3 [67] have been widely utilized to detect fish in the ocean. In addition, mask R-CNN [71] has been used to assess the abundance of luderick in the estuary. Although the above studies have achieved satisfactory results, these generic detection algorithms are initially designed for objects on land, and their application in actual underwater scenes has certain limitations. Therefore, researchers further considered the special characteristics of fish detection and have made various improvements to the generic detection algorithms to improve the detection accuracy and speed.

Due to the complex background interference, blurring and occlusion in the underwater image, it is challenging for the DNN to extract generalized fish features to detect fish accurately. Some researchers have incorporated attention mechanisms and deformable convolution into typical object detection architectures to improve the algorithm's performance. Labao and Naval Jr [72] proposed a multi-cascade R-CNN object detection network linked by sequential LSTMs to accurately count fish under various benthic backgrounds and illumination conditions in the ocean. The automatic correction of cascade structure and attention mechanism of LSTM links makes it robust to small fish, the multi-scale distorted and cropped images, and varying background environments. They achieved a precision of 53.29% on the test data taken in the wild containing many fish objects having small pixel areas, which is superior to Faster R-CNN. However, the accuracy of this method cannot meet the actual demand, and the amount of computation will increase with the number of cascaded R-CNN structures. Li et al. [73] integrated the deformable convolution module (DCM) and adaptive threshold module (ATM) into the YOLOv5 framework to accurately detect and count Takifugu rubripe in aquaculture ponds, in which the DCM was used to reduce the blurred background interference, and ATM was used to mitigate missed detection in densely occluded scenarios. They achieved a precision of 97.53% and a recall of 98.09% on the dataset taken in the real aquaculture ponds.

Embedding lightweight network models with relatively few parameters and high accuracy into edge devices (e.g., Jetson Nona, Raspberry Pi) to enable real-time and online fish monitoring has received increasing attention. The one-stage object detection algorithms with fast detection speed have become the preferred method in related research. Li et al. [40] designed a lightweight Underwater-YOLO architecture for an underwater robotic embedded system to perform real-time fish detection in an unconstrained marine environment. They improved the original YOLO architecture by feature fusion and reducing the number of convolutional layers and kernels and combined it with video frame selection to improve fish detection speed. It achieved a detection speed of 28 FPS with 93% detection accuracy on 500 marine images collected by ROV. However, the improvement in the detection speed of the video frame selection algorithm on the GPU platform is not significant. Cai et al. [47] developed a lightweight YOLO V3 detection framework using MobileNetV1 as a backbone to detect and count *Takifugu rubripe* in aquaculture ponds in real-time, which replaced the regular convolution operation in CNN with depth-separable convolution to reduce the number of model parameters. In addition, they further reselected the feature maps of MobileNet according to the object scale and receptive field for better fish detection. As a result, they achieved a mAP of 78.63% and a detection speed of 13 FPS.

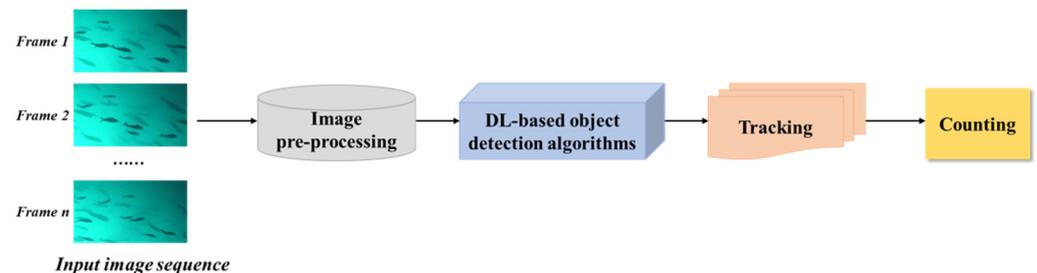
Judging from the literature, for relatively simple underwater scenes with sparse fish populations and slight background interference, the researchers' one-stage and two-stage object detection algorithms have shown promising results in detecting and locating fish in images. However, for more complex scenes, especially with issues such as complex background interference, frequent occlusions, and appearance variation of fish, there is still room for improvement in the detection accuracy of the detection algorithm. DL-based preprocessing methods of images may effectively alleviate the adverse effects of the environment. In the future, a combination of DL-based object detection and density estimation algorithms may be suitable for accurate fish counting in intensive aquaculture. Furthermore, performing real-time detection with high accuracy is more practical for fish moving rapidly in intensive aquaculture.

#### 4.1.2. Video-Based Fish Counting

In aquaculture, it is inefficient to count fish in a single image. The video contains the fish's appearance information (such as color, shape, and texture) in the static image and the motion information of the fish in the temporal dimension. Therefore, on the one hand, it is effective to develop a hybrid system that fuses the appearance information (spatial information) and the motion information (temporal information) of fish to improve the detection accuracy of fish in complex underwater environments and occlusion conditions. On the other hand, compared with object detection in images, the association and tracking of the same object between different video frames can prevent duplicate counts of the same object.

Regarding spatial-temporal information fusion, some studies have been proposed to combine DL and handcrafted feature extraction methods to fuse spatial-temporal information from videos for better fish detection. For example, Salman et al. [74] used the Gaussian mixture model (GMM) and optical flow to extract motion information of fish in unconstrained underwater videos, which was then combined with grayscale images as the input to a Faster R-CNN to generate fish candidate regions. As a result, they achieved an F1-Score of 87.44% and 80.02% on the FCS and LifeCLEF 2015 datasets, respectively. The experimental results demonstrate that the method is superior to various image-based fish detection methods. However, the GMM algorithm is sensitive to noise and other environmental factors variations. To solve this problem, Ben et al. [28] proposed multi-stream fusion methods consisting of two Faster R-CNN models sharing the same RPN or the same classifier. This method used two streams to merge appearance information from the RGB image and motion information from the optical flow to detect the moving fish in the ocean, achieving a mAP of 73.69% and F1-Score of 83.16% on the LifeCLEF2015 dataset.

Methods for associating individual fish in videos to reduce duplicate counts generally obey a detection-tracking framework, as shown in Figure 4, where an object detection network is utilized to detect individual fish in different video frames. Then a tracking algorithm is used to correlate detection results in the video to generate the tracking results. Finally, the number of fish can be calculated by counting the number of trajectories. Currently, some studies mainly focus on single-class multi-target fish detection and tracking. Levy et al. [75] adopted RetinaNet that is a one-stage CNN detector as the object detector to detect the fish in underwater video frames, then used the simple online real-time tracker (SORT) as the object tracker to connect detections of the same fish, which achieved an AP of 74%. Arvind et al. [41] applied Mask R-CNN and the GOTURN tracking algorithm to detect and track fish in videos taken by unmanned aerial vehicles (UAV) above the tank, reaching an F1 score of 91% and 16 FPS. Mohamed et al. [76] used YOLO and an optical flow algorithm to track the goldfish in the fishpond. However, tracking after all detected results cannot meet the real-time requirements. Regarding multi-class multi-target fish detection and tracking, Liu et al. [29] proposed a two-branch DNN based on YOLOV4 to simultaneously detect and track underwater multi-class fish in a real marine ranch environment real-time. The detection speed of this method reached 33 FPS, which met the real-time requirements and achieved an accuracy of fish stock statistics of 95.6%.



**Figure 4.** Flow chart of the fish counting based on detection-tracking.

Judging from the aforementioned literature, considering both spatial and temporal information has been proven to be an effective way to improve the detection accuracy of fish, while how to fuse these two types of information and apply it to real aquaculture environments needs further research. Moreover, lightweight one-stage object detection algorithms have become the preferred method for the detection-tracking framework of fish and have achieved satisfactory results. However, the challenge of tracking loss and errors due to fish deformation, mutual occlusion, and sudden movements still persists. In the future, enhancing the accuracy of the detection branch can facilitate the tracking of accurate targets. Additionally, Correcting missed and false detection in trajectories can be used to ensure error-free trajectory tracking. Furthermore, most current studies focus on tracking a single category of the target, while the issue of target category imbalance in multi-class target tracking merits carefully investigated in the future.

#### 4.2. Fish Body Length Measurement

The body length of fish in aquaculture is not only crucial basic information to represent the growth status of fish but also an important indicator of product grading [77–79]. Since the longitudinal length of the ideal object detection bounding box predicted by the DL-based object detection algorithms corresponds to the distal extremes of the morphological features representing the body length of the fish, which has become the preferred method for measuring the body length of fish, especially, Mask R-CNN can perform fish detection and body area segmentation simultaneously using one network structure, which has been widely used in fish body length measurement. A typical Mask R-CNN-based fish detection architecture is shown in Figure 5. Existing studies related to fish body length measurement mainly considered image data. According to the image acquisition method, existing fish body length measurement methods can be divided into monocular and stereo vision-based.

Table 4 details the relevant fish body length measurement studies using DL-based object techniques.

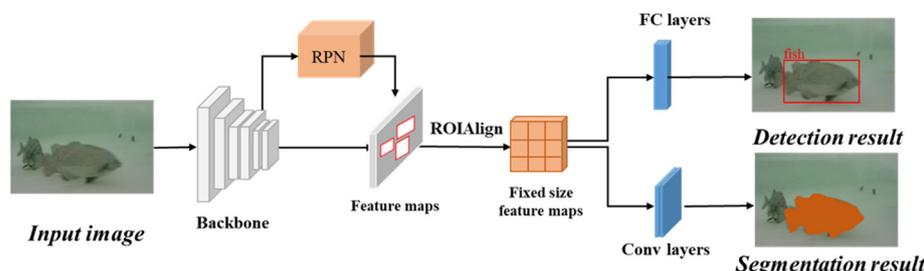


Figure 5. A typical Mask R-CNN-based fish detection architecture.

Table 4. Summary of DL-based object detection techniques applied in fish body length measurement.

Camera	References	Approaches	Fish Species	Data Preprocessing	Results
Monocular camera	Monkman et al. [30]	R-CNN	European sea bass	N/A	Mean bias error = 2.2%
	Álvarez-Ellacuría et al. [31]	Mask R-CNN + Statistical models	European hake	N/A	Root-mean-square deviation = 1.9 cm
	Palmer et al. [80]	Mask R-CNN + Statistical model	Dolphinfish	N/A	The square root of the mean squared deviation = 2.4 cm
Stereo camera	Huang et al. [81]	Mask R-CNN + GrabCut + 3D cloud points + Coordinate transformation	Porphyry seabream	N/A	Average error = 5.5 mm (length); Average error = 2.7 mm (width)
	Garcia et al. [49]	Mask R-CNN + Local gradient + Morphological operations + Curve fitting	Saithe; Blue whiting; Redfish; Atlantic mackerel; Velvet belly lanternshark; Norway pout; Atlantic herring	Image linearization. Correction of non-uniform lighting	The average IOU = 0.89 (single fish); The average IOU = 0.79 (Overlapping fish)

#### 4.2.1. Monocular Vision-Based Fish Body Length Measurement

Monocular cameras can be combined with DL-based object detection techniques to measure the body length of harvested fish from images collected in diverse environments with limited space, such as the decks of vessels and fish boxes. Monkman et al. [30] utilized three R-CNN architectures based on NASNet, ResNet-101, and MobileNet to detect the fusiform fish on the shore and calculate the total length of the fusiform fish based on the real-world length per pixel provided by the ArUco fiducial markers. They achieved a percent mean bias error of 2.2%. However, the fish are obscured from each other in most cases, and only a few fish get a full view. In contrast, many complete fish heads are visible in the images. Therefore, Álvarez-Ellacuría et al. [31] utilized Mask R-CNN to detect European hake heads from images of fish boxes and developed different statistical models that permit the estimation of the entire length of European based on the detected heads' length. For fish with average lengths ranging from 20–40 cm, the maximum deviation between the predicted value of this method and the real measured mean body length value of fish was 4.0 cm. However, it remains difficult for this method to address cases where the complete fish head is not visible. To overcome this issue, Palmer et al. [80] used

MaskR-CNN to count the number of dolphinfish (*Coryphaena hippurus*) in each fish box, then developed a statistical model to predict the mean length of the dolphinfish based on the weight of the box and the number of dolphinfishes.

Judging from the aforementioned literature, typical two-stage target detection architectures can effectively identify and locate the head and body of harvested fish. Nevertheless, there is still room for refinement regarding the precision of fish length measurement and detection speed. The main challenges lie in the pose of fish, occlusion, and complex background environment. In addition to human intervention, image preprocessing methods based on DL techniques may effectively solve these problems. Lightweight instances segment networks, such as YOLAT [82] and SOLO [83], merit further investigation. In addition, recent studies have indicated a strong correlation between the body weights of fish and their biometric characteristics, such as length, width, and height [84,85].

#### 4.2.2. Stereo Vision-Based Fish Body Length Measurement

The binocular camera can reconstruct the 3D information of the image and be combined with the DL-based object detection techniques to perform automatic measurements of underwater fish without constraining their movement. Huang [81] extracted the 3D cloud points of fish based on the segmentation results in stereo images obtained by Make R-CNN and Grabcut, then applied the coordinate transformations to get the dimensions of multiple fish with various orientations and positions in the image. They conducted experiments on a dataset collected from a lab. The average error of the method for fish length and width measurement was around 5.5 mm and 2.7 mm, respectively. However, this method assumes that the outline of the fish is co-plane, which imposes restrictions on the measured position of the fish. For a more complex and uncontrolled real-world marine environment, Garcia [49] utilized Mask R-CNN and local gradient to automatically segment individual fish in stereo images obtained by the deep vision imaging system placed in the trawl. Then, they used morphological operations, curve fitting, and stereo information to estimate the fish skeleton and the total length of the fish. It achieved the average IoU values of 0.79 and 0.89 for overlapping and non-overlapping fish, respectively. However, verifying the accuracy of the predicted fish length is difficult.

Judging from the aforementioned literature, existing studies utilizing stereo vision technology and Mask R-CNN for measuring the body length of fish have yielded satisfactory results in controlled and ideal experimental scenarios. However, measuring the dimensions of fish in real-world unconstrained aquaculture remains challenging, mainly due to occlusion and fish bending. Therefore, in the future, it is necessary to develop more advanced object detection algorithms that possess powerful fish discriminative feature extraction capabilities, to measure the length of fish and the weight of the fish accurately and quickly in actual intensive aquaculture. Moreover, the potential of DL-based object detection techniques for biomass prediction of fish needs to be carefully researched [86].

#### 4.3. Individual Fish Behavior Analysis

Fish typically respond to and interact with the environment through various behaviors. Monitoring and analyzing fish behavior changes in aquaculture are vital for evaluating fish health and welfare, guiding feeding, and reducing water pollution [87,88]. Typical fish behavior includes abnormal feeding, swimming, and reproductive behavior, which can be classified into individual and group behaviors. Regarding fish population behavior analysis, DL-based image classification methods have been widely used to classify fish aggregation and dispersion states in images to identify the different behaviors of fish populations [89,90]. In this section, we mainly focused on individual fish behavior analysis. Locating and identifying specific behaviors of individuals in a fish population is challenging for high-density commercial aquaculture. In recent years, DL-based object detection algorithms have been increasingly applied to individual behavior analysis of fish. Table 5 details the relevant individual fish behavior analysis studies using DL-based object techniques.

Image and video processing are the two primary means of vision-based individual behavior analysis of fish.

**Table 5.** Summary of DL-based object detection techniques applied in individual fish behavior analysis.

Data	References	Approaches	Fish Species /Feed	Behaviors	Data Preprocessing	Results
Image	Hu et al. [32]	Improved YOLOv3	Crucian carp; catfish	Hunger and lack of oxygen behavior	CLAHE; DWT; Median filter; Flipping; Rotation; Gaussian blurring; Resize	Precision = 89.7%; Recall = 88.4%; IOU = 89.2%; FPS = 240
	Hu et al. [52]	Improved YOLOV4	Uneaten feed	Feeding status	CLAHE; Mosaic	Precision = 94%; Recall = 89%; F1 Score = 91%; AP50 = 92.61%
Video	Xu et al. [91]	Faster R-CNN; YOLOV3	Red goldfish	Fish behavior in different concentrations of ammonia	Random cutting	ACC = 98.13% (Faster R-CNN); ACC = 95.66% (YOLOV3)
	Wang et al. [33]	Improved YOLOV5 + SiamRPN++	Porphyry seabream	Turning-over behavior	N/A	Detection: AP50 = 99.4%; Tracking Precision = 76.7%

#### 4.3.1. Image-Based Individual Fish Behavior Analysis

DL-based object detection algorithms with a classifier can describe and quantify the behavior patterns of individual fish from images based on their visual characteristics, such as appearance and morphology information. Hunger or lack of oxygen affects the feeding and stress behavior of fish. Hu et al. [32] developed a lightweight DL architecture (YOLOv3-Lite) to automatically detect the hunger and oxygenation behavior of crucian carp and catfish in tanks. It used a lightweight MobileNetv2 as the backbone to simplify detection. In addition, it introduced an improved spatial pyramid pooling block structure and bounding box regression loss function to better identify the behaviors of multi-scale fish. They achieved a processing speed of 240 FPS and a precision of 89.7%. However, the research was conducted in a controlled indoor environment. The actual outdoor aquaculture environments usually are more complex, including complex backgrounds and dense fish populations, which is challenging for accurate individual fish behavior detection in aquaculture. In addition, monitoring uneaten feed in underwater images can indirectly assess the feeding status of fish. To overcome the detection challenges posed by low-quality underwater images, large amounts of, and highly small-sized feed pellets, Hu et al. [52] designed an improved YOLOV4 to accurately detect underwater feed pellets in maricultural net cages with Adult Atlantic salmon. The improved YOLO-V4 used a finer-grained feature map, dense connection, and a de-redundancy operation to improve the detection accuracy of uneaten feed pellets, reaching a precision of 94%.

Judging from the aforementioned literature, further investigation is required to analyze the individual feeding behavior of fish in the actual intensive breeding environment. The challenges encountered by DL-based object detection algorithms for image-based individual fish behavior analysis in aquaculture mainly come from complex background interference, frequent occlusions, and appearance variation of fish. Therefore, more advanced methods suitable for achieving a balance between detection speed and detection

accuracy must be proposed for various scenarios. Additionally, due to the constraints posed by dataset acquisition conditions and fish behavior, data imbalance across different fish behaviors is common. Therefore, in the future, the methods that are suitable for small sample datasets are worth researching.

#### 4.3.2. Video-Based Individual Fish Behavior Analysis

The behaviors of fish in the videos are continuous and temporally correlated. Therefore, extracting and analyzing individual fish movement trajectories from underwater videos can effectively evaluate fish behavior patterns and underlying regularities. To analyze the individual behavior of red goldfish under different ammonia concentrations, Xu et al. [91] used Faster R-CNN and YOLOV3 to locate and identify red goldfish in underwater images. Then they plotted the three-dimensional spatial trajectories of red goldfish in different concentrations of ammonia. The average recognition accuracy of Faster R-CNN is 98.13%, which is superior to YOLO-V3. However, they are limited to the trajectory description of a single fish in a container, which cannot identify different individuals in multiple fish. Thus, individual fish associations across frames in multiple fish can be achieved by object tracking. Wang et al. [33] used the improved YOLOV5 to detect the turning-over behavior of porphyry seabream and used SiamRPN++ to track them. They modified the YOLOV5 architecture to improve the detection precision of the small targets by incorporating multi-level features and feature mapping. The detection precision AP50 of this method is 99.4%, and the tracking precision is 76.7%. However, they used a single-target tracking algorithm to track multiple underwater fish, and the algorithm gets slower as the number of fish increases.

Judging from the aforementioned literature, lightweight one-stage object detection algorithms and trajectory tracking have been employed to effectively analyze the abnormal behavior of fish. However, individual fish behavior analysis based on detection and tracking in aquaculture still faces the challenges of high density and complex behavior of fish. Furthermore, identifying fish behavior from individual video frames ignores the correlation information of images acquired before and after fish movements. The integration of CNN and RNN can provide a powerful framework for extracting spatiotemporal features that can facilitate improved identification and tracking of fish behavior in video frames [92,93]. Moreover, these studies considered only one kind of abnormal behavior, which is insufficient for practical application.

### 5. Performance Evaluation Metrics

Different evaluation metrics can quantitatively evaluate the performance of the algorithm from different perspectives as well as guide the design of the algorithm. The quantitative metrics employed in the related literature are listed in Table 6. Accuracy (ACC), precision, recall, F1-score, mAP, and intersection over union (IOU) are the primary evaluation metrics for object detection techniques [32,69,71]. In image-based fish counting and body length measurement tasks, mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE) and mean absolute percentage error (MAPE) is used to evaluate the performance of the algorithms [74,81]. ID switch and multiple objects tracking accuracy (MOTA) are commonly used evaluation metrics when fish counting based on object tracking in the video is performed [29]. Additionally, the detection speed of the algorithm is also an important evaluation metric measured by FPS. Better results are measured using larger ACC, precision, recall, F1-Score, IOU, and MOTA values and smaller MAE, MSE, MAPE, RMSE, and ID switch values.

**Table 6.** The evaluation metrics for DL-based models.

Evaluation Metrics	Better Results	Description
ACC	Larger	The ratio of the number of correctly identified samples to the total number of identified samples
Precision	Larger	The ratio of correctly identified fish to all identified fish
Recall	Larger	The ratio of correctly identified fish to all fish in the sample
F1-Score	Larger	The harmonic means of precision and recall
mAP	Larger	Takes both precision and recall into consideration
IOU	Larger	The overlap rate between the candidate area and the ground truth area
MAE	Smaller	The expected value of the absolute difference between the predicted value and the ground truth
MAPE	Smaller	Considers not only the error between the predicted value and the ground truth but also the ratio between the error and the ground truth
MSE	Smaller	The expected value of the square of the difference between the predicted value and the ground truth
RMSE	Smaller	The square root of the MSE
ID switch	Smaller	The average total number of times that a resulting trajectory switches its matched ground-truth identity with another trajectory
MOTA	Larger	Combines false positives and missed targets and identifies switches
FPS	Larger	The number of images processed by the algorithm per second

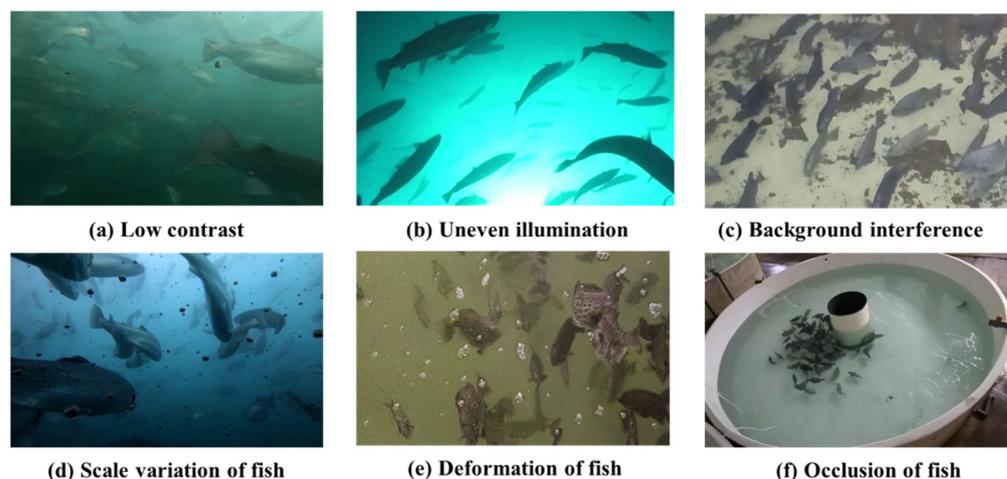
## 6. Challenges and Future Perspectives

DL-based object detection, with its unique advantage of simultaneously identifying fish categories and locations, has become a key technology in fish aquaculture. Generic DL-based object detection architectures have yielded satisfactory fish counting, body length measurement, and behavior analysis results. However, DL-based object detection techniques in aquaculture practice still face various challenges.

(1) Complex environment and fish characteristics in real aquaculture make it difficult for the DL-based object detection models to discriminate key information about fish, as shown in Figure 6. For instance, in terms of complex environments, vision-based underwater images often show low contrast (Figure 6a) and uneven illumination (Figure 6b) due to the scattering and absorption of light during underwater propagation. Moreover, the images may contain multiple non-fish background interferences similar to fish (Figure 6c). Regarding complex fish characteristics, the scale of the fish in the image may vary from a few hundred pixels to several thousand pixels due to the camera perspective (Figure 6d). Furthermore, the pose and shape of the fish may be distorted when it swims freely (Figure 6e). In addition, fish may easily obscure each other in high-density scenes (Figure 6f).

(2) The DL-based object detection models with high detection accuracy are usually limited by many operations and often require massive computing resources. Therefore, embedding the lightweight DL-based object detection model into fish detection systems is urgently needed for real-time fish monitoring in actual aquaculture environments.

(3) Most of the DL-based object detection models proposed by existing aquaculture-related studies rely on large-scale labeled datasets for training to achieve outstanding performance. However, annotating fish with dense distributions in images is time-consuming and labour-intensive.



**Figure 6.** Challenges of complex environments and complex fish characteristics in aquaculture. These images are from real aquaculture farms.

(4) Other application fields that may have great potential include fish disease diagnosis and biomass prediction. However, no research on DL-based object detection techniques has been reported.

On this basis, we forecast the development trends of DL-based object detection techniques in aquaculture, including:

(1) Advanced algorithms specifically for real aquaculture environments will serve fish detection in complex aquaculture scenarios more effectively, which is a helpful research direction. Incorporating both the spatial-temporal information of fish and various contextual information into the models, or utilizing background suppression techniques, are promising approaches that can improve the effectiveness of fish detection in real-world aquaculture settings.

(2) DNNs that balance detection speed and detection accuracy are worth investigating. Optimizing the architecture and parameters of the model or implementing model compression techniques is practical.

(3) To alleviate the need for large-scale labeled data for DL-based object detection models, it is practical to study semi-supervised or unsupervised learning methods to perform fish counting, body length measurement, and behavior analysis tasks in aquaculture [94].

(4) Besides the fish counting, body length measurement, and behavior analysis reviewed in this paper, DL-based object techniques applied to fish disease diagnosis [95,96] and fish biomass prediction are also worth investigating. GAN can augment the sample size of the fish to be detected, e.g., diseased fish and normal fish.

## 7. Conclusions

This study reviewed the application of DL-based object detection techniques in fish aquaculture, focusing on fish counting, body length measurement, and individual behavior analysis. The research status was concluded from two aspects: image and video analysis. The technical details of the DL-based object detection techniques in aquaculture were also summarized, including the dataset, image preprocessing methods, typical DL-based object detection algorithms, and evaluation metrics. Finally, the challenges and potential trends of DL-based object detection techniques in aquaculture were concluded and discussed to provide a reference for developing DL techniques in fish monitoring. The review shows that generic DL-based object detection architectures have played important roles in aquaculture.

However, fish detection in real aquaculture practice still faces various challenges, such as the complex background environment and fish characteristics, real-time detection requirements, the lack of large-scale labeled datasets, and unexplored research areas. In the future, it is practical to develop more advanced algorithms with high accuracy and real-time

detection speed designed explicitly for real aquaculture environments. To alleviate the need for large-scale labeled data for DL-based object detection models, semi-supervised or unsupervised algorithms are feasible. In addition, it is worthwhile to further apply DL-based object technology in fish disease diagnosis and biomass prediction.

**Author Contributions:** Conceptualization, investigation, writing—original draft preparation, H.L.; supervision, investigation, writing—review and editing, X.M.; project administration, Y.Y., L.W. and L.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded in part by the Key Research and Development Program of Shandong Province, grant number 2021SFGC0701, and in part by the Marine Science and Technology Innovation Project of Qingdao City, grant number 22-3-3-hygg-4-hy.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Lauder, G.V. Fish Locomotion: Recent Advances and New Directions. *Annu. Rev. Mar. Sci.* **2015**, *7*, 521–545. [[CrossRef](#)] [[PubMed](#)]
- Monkman, G.G.; Hyder, K.; Kaiser, M.J.; Vidal, F.P. Application of machine vision systems in aquaculture with emphasis on fish: State-of-the-art and key issues. *Rev. Aquac.* **2017**, *9*, 369–387. [[CrossRef](#)]
- FAO. *The State of World Fisheries and Aquaculture 2020: Sustainability in Action*; FAO: Rome, Italy, 2020; p. 244.
- Bossier, P.; Ekasari, J. Biofloc technology application in aquaculture to support sustainable development goals. *Microb. Biotechnol.* **2017**, *10*, 1012–1016. [[CrossRef](#)] [[PubMed](#)]
- Zhao, S.; Zhang, S.; Liu, J.; Wang, H.; Zhu, J.; Li, D.; Zhao, R. Application of machine learning in intelligent fish aquaculture: A review. *Aquaculture* **2021**, *540*, 736724. [[CrossRef](#)]
- 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. *Arch. Comput. Methods Eng.* **2020**, *28*, 2785–2816. [[CrossRef](#)]
- Mei, Y.; Sun, B.; Li, D. Recent advances of target tracking applications in aquaculture with emphasis on fish. *Comput. Electron. Agric.* **2022**, *201*, 107335. [[CrossRef](#)]
- Sutterlin, A.M.; Jokola, K.J.; Holte, B. Swimming Behavior of Salmonid Fish in Ocean Pens. *J. Fish. Res. Board Can.* **1979**, *36*, 948–954. [[CrossRef](#)]
- Yada, S.; Chen, H. Weighing Type Counting System for Seedling Fry. *Nihon-Suisan-Gakkai-Shi* **1997**, *63*, 178–183. [[CrossRef](#)]
- Li, D.; Hao, Y.; Duan, Y. Nonintrusive methods for biomass estimation in aquaculture with emphasis on fish: A review. *Rev. Aquac.* **2019**, *12*, 1390–1411. [[CrossRef](#)]
- An, D.; Hao, J.; Wei, Y.; Wang, Y.; Yu, X. Application of computer vision in fish intelligent feeding system—A review. *Aquac. Res.* **2020**, *52*, 423–437. [[CrossRef](#)]
- Yang, X.; Zhang, S.; Liu, J.; Gao, Q.; Dong, S.; Zhou, C. Deep learning for smart fish farming: Applications, opportunities and challenges. *Rev. Aquac.* **2020**, *13*, 66–90. [[CrossRef](#)]
- Li, D.; Du, L. Recent advances of deep learning algorithms for aquacultural machine vision systems with emphasis on fish. *Artif. Intell. Rev.* **2021**, *55*, 4077–4116. [[CrossRef](#)]
- Kutlu, Y.; Iscimen, B.; Turan, C. Multi-stage fish classification system using morphometry. *Fresenius Environ. Bull.* **2017**, *26*, 1910–1916.
- Lalabadi, H.M.; Sadeghi, M.; Mireei, S.A. Fish freshness categorization from eyes and gills color features using multi-class artificial neural network and support vector machines. *Aquac. Eng.* **2020**, *90*, 102076. [[CrossRef](#)]
- Zhao, Y.-P.; Sun, Z.-Y.; Du, H.; Bi, C.-W.; Meng, J.; Cheng, Y. A novel centerline extraction method for overlapping fish body length measurement in aquaculture images. *Aquac. Eng.* **2022**, *99*, 102302. [[CrossRef](#)]
- LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
- Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaria, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 1–74. [[CrossRef](#)]
- Zhao, Z.-Q.; Zheng, P.; Xu, S.-T.; Wu, X. Object Detection With Deep Learning: A Review. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3212–3232. [[CrossRef](#)]
- Ranjan, R.; Patel, V.M.; Chellappa, R. HyperFace: A Deep Multi-Task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *41*, 121–135. [[CrossRef](#)]
- Liu, W.; Hasan, I.; Liao, S. Center and Scale Prediction: Anchor-free Approach for Pedestrian and Face Detection. *Pattern Recognit.* **2023**, *135*, 109071. [[CrossRef](#)]

22. Ma, J.; Shao, W.; Ye, H.; Wang, L.; Wang, H.; Zheng, Y.; Xue, X. Arbitrary-Oriented Scene Text Detection via Rotation Proposals. *IEEE Trans. Multimed.* **2018**, *20*, 3111–3122. [[CrossRef](#)]
23. Xu, Y.; Fu, M.; Wang, Q.; Wang, Y.; Chen, K.; Xia, G.-S.; Bai, X. Gliding Vertex on the Horizontal Bounding Box for Multi-Oriented Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 1452–1459. [[CrossRef](#)] [[PubMed](#)]
24. Li, J.; Liang, X.; Shen, S.; Xu, T.; Feng, J.; Yan, S. Scale-aware Fast R-CNN for Pedestrian Detection. *IEEE Trans. Multimed.* **2017**, *20*, 985–996. [[CrossRef](#)]
25. Islam, M.M.; Newaz, A.A.R.; Karimodini, A. Pedestrian Detection for Autonomous Cars: Inference Fusion of Deep Neural Networks. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 23358–23368. [[CrossRef](#)]
26. Wang, H.; Yu, Y.; Cai, Y.; Chen, X.; Chen, L.; Liu, Q. A Comparative Study of State-of-the-Art Deep Learning Algorithms for Vehicle Detection. *IEEE Intell. Transp. Syst. Mag.* **2019**, *11*, 82–95. [[CrossRef](#)]
27. Li, G.; Ji, Z.; Qu, X. Stepwise Domain Adaptation (SDA) for Object Detection in Autonomous Vehicles Using an Adaptive CenterNet. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 17729–17743. [[CrossRef](#)]
28. Ben Tamou, A.; Benzinou, A.; Nasreddine, K. Multi-stream fish detection in unconstrained underwater videos by the fusion of two convolutional neural network detectors. *Appl. Intell.* **2021**, *51*, 5809–5821. [[CrossRef](#)]
29. Liu, T.; Li, P.; Liu, H.; Deng, X.; Liu, H.; Zhai, F. Multi-class fish stock statistics technology based on object classification and tracking algorithm. *Ecol. Inform.* **2021**, *63*, 101240. [[CrossRef](#)]
30. Monkman, G.G.; Hyder, K.; Kaiser, M.J.; Vidal, F.P. Using machine vision to estimate fish length from images using regional convolutional neural networks. *Methods Ecol. Evol.* **2019**, *10*, 2045–2056. [[CrossRef](#)]
31. Álvarez-Ellacuría, A.; Palmer, M.; Catalán, I.A.; Lisani, J.-L. Image-based, unsupervised estimation of fish size from commercial landings using deep learning. *ICES J. Mar. Sci.* **2019**, *77*, 1330–1339. [[CrossRef](#)]
32. Hu, J.; Zhao, D.; Zhang, Y.; Zhou, C.; Chen, W. Real-time nondestructive fish behavior detecting in mixed polyculture system using deep-learning and low-cost devices. *Expert Syst. Appl.* **2021**, *178*, 115051. [[CrossRef](#)]
33. Wang, H.; Zhang, S.; Zhao, S.; Wang, Q.; Li, D.; Zhao, R. Real-time detection and tracking of fish abnormal behavior based on improved YOLOV5 and SiamRPN++. *Comput. Electron. Agric.* **2021**, *192*, 106512. [[CrossRef](#)]
34. Fisher, R.B.; Chen-Burger, Y.-H.; Giordano, D.; Hardman, L.; Lin, F.-P. *Fish4Knowledge: Collecting and Analyzing Massive Coral Reef Fish Video Data*; Springer: Berlin/Heidelberg, Germany, 2016. [[CrossRef](#)]
35. Joly, A.; Goëau, H.; Glotin, H.; Spampinato, C.; Bonnet, P.; Vellinga, W.P.; Planque, R.; Rauber, A.; Fisher, R.; Müller, H. LifeCLEF 2014: Multimedia life species identification challenges. In *Information Access Evaluation. Multilinguality, Multimodality, and Interaction, Proceedings of the 5th International Conference of the CLEF Initiative, CLEF 2014, Sheffield, UK, 15–18 September 2014*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 229–249. [[CrossRef](#)]
36. Joly, A.; Goëau, H.; Glotin, H.; Spampinato, C.; Bonnet, P.; Vellinga, W.-P.; Planqué, R.; Rauber, A.; Palazzo, S.; Fisher, B. LifeCLEF 2015: Multimedia life species identification challenges. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction, Proceedings of the 6th International Conference of the CLEF Association, CLEF'15, Toulouse, France, 8–11 September 2015*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 462–483. [[CrossRef](#)]
37. Cutter, G.; Stierhoff, K.; Zeng, J. Automated detection of rockfish in unconstrained underwater videos using haar cascades and a new image dataset: Labeled fishes in the wild. In *Proceedings of the 2015 IEEE Winter Applications and Computer Vision Workshops, Waikoloa, HI, USA, 6–9 January 2015*; pp. 57–62.
38. Ali-Gombe, A.; Elyan, E.; Jayne, C. Fish classification in context of noisy images. In *Engineering Applications of Neural Networks, Proceedings of the 18th International Conference, EANN 2017, Athens, Greece, 25–27 August 2017*; Springer International Publishing: Cham, Switzerland, 2017; pp. 216–226. [[CrossRef](#)]
39. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Li, F.-F. Imagenet: A large-scale hierarchical image database. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June*; pp. 248–255. [[CrossRef](#)]
40. Li, Q.; Li, Y.; Niu, J. Real-time detection of underwater fish based on improved Yolo and transfer learning. *Pattern Recognit. Artif. Intell.* **2019**, *32*, 193–203. [[CrossRef](#)]
41. Arvind, C.; Prajwal, R.; Bhat, P.N.; Sreedevi, A.; Prabhudeva, K. Fish detection and tracking in pisciculture environment using deep instance segmentation. In *Proceedings of the TENCON 2019—2019 IEEE Region 10 Conference (TENCON), Kochi, India, 17–20 October 2019*; pp. 778–783. [[CrossRef](#)]
42. Costa, C.; Scardi, M.; Vitalini, V.; Cataudella, S. A dual camera system for counting and sizing Northern Bluefin Tuna (*Thunnus thynnus*; Linnaeus, 1758) stock, during transfer to aquaculture cages, with a semi automatic Artificial Neural Network tool. *Aquaculture* **2009**, *291*, 161–167. [[CrossRef](#)]
43. Petritoli, E.; Cagnetti, M.; Leccese, F. Simulation of Autonomous Underwater Vehicles (AUVs) Swarm Diffusion. *Sensors* **2020**, *20*, 4950. [[CrossRef](#)]
44. Wu, Y.; Duan, Y.; Wei, Y.; An, D.; Liu, J. Application of intelligent and unmanned equipment in aquaculture: A review. *Comput. Electron. Agric.* **2022**, *199*, 107201. [[CrossRef](#)]
45. Zhou, C.; Zhang, B.; Lin, K.; Xu, D.; Chen, C.; Yang, X.; Sun, C. Near-infrared imaging to quantify the feeding behavior of fish in aquaculture. *Comput. Electron. Agric.* **2017**, *135*, 233–241. [[CrossRef](#)]
46. Lin, K.; Zhou, C.; Xu, D.; Guo, Q.; Yang, X.; Sun, C. Three-dimensional location of target fish by monocular infrared imaging sensor based on a L-z correlation model. *Infrared Phys. Technol.* **2018**, *88*, 106–113. [[CrossRef](#)]

47. Cai, K.; Miao, X.; Wang, W.; Pang, H.; Liu, Y.; Song, J. A modified YOLOv3 model for fish detection based on MobileNetv1 as backbone. *Aquac. Eng.* **2020**, *91*, 102117. [[CrossRef](#)]
48. Salman, A.; Jalal, A.; Shafait, F.; Mian, A.; Shortis, M.; Seager, J.; Harvey, E. Fish species classification in unconstrained underwater environments based on deep learning. *Limnol. Oceanogr. Methods* **2016**, *14*, 570–585. [[CrossRef](#)]
49. Garcia, R.; Prados, R.; Quintana, J.; Tempelaar, A.; Gracias, N.; Rosen, S.; Vågstøl, H.; Løvall, K. Automatic segmentation of fish using deep learning with application to fish size measurement. *ICES J. Mar. Sci.* **2019**, *77*, 1354–1366. [[CrossRef](#)]
50. Zhou, W.-H.; Zhu, D.-M.; Shi, M.; Li, Z.-X.; Duan, M.; Wang, Z.-Q.; Zhao, G.-L.; Zheng, C.-D. Deep images enhancement for turbid underwater images based on unsupervised learning. *Comput. Electron. Agric.* **2022**, *202*, 107372. [[CrossRef](#)]
51. Ranjan, R.; Sharrer, K.; Tsukuda, S.; Good, C. Effects of image data quality on a convolutional neural network trained in-tank fish detection model for recirculating aquaculture systems. *Comput. Electron. Agric.* **2023**, *205*, 107644. [[CrossRef](#)]
52. Hu, X.; Liu, Y.; Zhao, Z.; Liu, J.; Yang, X.; Sun, C.; Chen, S.; Li, B.; Zhou, C. Real-time detection of uneaten feed pellets in underwater images for aquaculture using an improved YOLO-V4 network. *Comput. Electron. Agric.* **2021**, *185*, 106135. [[CrossRef](#)]
53. Lu, Y.; Chen, D.; Olaniyi, E.; Huang, Y. Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review. *Comput. Electron. Agric.* **2022**, *200*, 107208. [[CrossRef](#)]
54. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587. [[CrossRef](#)]
55. Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1440–1448. [[CrossRef](#)]
56. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 1137–1149. [[CrossRef](#)]
57. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2961–2969. [[CrossRef](#)]
58. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788. [[CrossRef](#)]
59. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016. [[CrossRef](#)]
60. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271. [[CrossRef](#)]
61. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767. [[CrossRef](#)]
62. Bochkovskiy, A.; Wang, C.; Liao, H. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934. [[CrossRef](#)]
63. Li, X.; Shang, M.; Qin, H.; Chen, L. Fast accurate fish detection and recognition of underwater images with fast r-cnn. In Proceedings of the OCEANS 2015-MTS/IEEE Washington, Washington, DC, USA, 19–22 October 2015; pp. 1–5. [[CrossRef](#)]
64. Li, X.; Shang, M.; Hao, J.; Yang, Z. Accelerating fish detection and recognition by sharing CNNs with objectness learning. In Proceedings of the OCEANS 2016—Shanghai, Shanghai, China, 10–13 April 2016. [[CrossRef](#)]
65. Mandal, R.; Connolly, R.M.; Schlacher, T.A.; Stantic, B. Assessing fish abundance from underwater video using deep neural networks. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–6.
66. Sung, M.; Yu, S.-C.; Girdhar, Y. Vision based real-time fish detection using convolutional neural network. In Proceedings of the OCEANS 2017—Aberdeen, Aberdeen, UK, 19–22 June 2017. [[CrossRef](#)]
67. Xu, W.; Matzner, S. Underwater fish detection using deep learning for water power applications. In Proceedings of the 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 12–14 December 2018; pp. 313–318. [[CrossRef](#)]
68. Li, D.; Miao, Z.; Peng, F.; Wang, L.; Hao, Y.; Wang, Z.; Chen, T.; Li, H.; Zheng, Y. Automatic counting methods in aquaculture: A review. *J. World Aquac. Soc.* **2020**, *52*, 269–283. [[CrossRef](#)]
69. Yu, X.; Wang, Y.; An, D.; Wei, Y. Counting method for cultured fishes based on multi-modules and attention mechanism. *Aquac. Eng.* **2021**, *96*, 102215. [[CrossRef](#)]
70. Zhao, Y.; Li, W.; Li, Y.; Qi, Y.; Li, Z.; Yue, J. LFCNet: A lightweight fish counting model based on density map regression. *Comput. Electron. Agric.* **2022**, *203*, 107496. [[CrossRef](#)]
71. Ditría, E.M.; Lopez-Marcano, S.; Sievers, M.; Jinks, E.L.; Brown, C.J.; Connolly, R.M. Automating the Analysis of Fish Abundance Using Object Detection: Optimizing Animal Ecology With Deep Learning. *Front. Mar. Sci.* **2020**, *7*, 429. [[CrossRef](#)]
72. Labao, A.B.; Naval, P.C., Jr. Cascaded deep network systems with linked ensemble components for underwater fish detection in the wild. *Ecol. Inform.* **2019**, *52*, 103–121. [[CrossRef](#)]
73. Li, H.; Yu, H.; Gao, H.; Zhang, P.; Wei, S.; Xu, J.; Cheng, S.; Wu, J. Robust detection of farmed fish by fusing YOLOv5 with DCM and ATM. *Aquac. Eng.* **2022**, *99*, 102301. [[CrossRef](#)]

74. Salman, A.; Siddiqui, S.A.; Shafait, F.; Mian, A.; Shortis, M.R.; Khurshid, K.; Ulges, A.; Schwanecke, U. Automatic fish detection in underwater videos by a deep neural network-based hybrid motion learning system. *ICES J. Mar. Sci.* **2019**, *77*, 1295–1307. [[CrossRef](#)]
75. Levy, D.; Belfer, Y.; Osherov, E.; Bigal, E.; Scheinin, A.P.; Nativ, H.; Tchernov, D.; Treibitz, T. Automated analysis of marine video with limited data. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1385–1393. [[CrossRef](#)]
76. Mohamed, H.E.-D.; Fadl, A.; Anas, O.; Wageeh, Y.; ElMasry, N.; Nabil, A.; Atia, A. MSR-YOLO: Method to Enhance Fish Detection and Tracking in Fish Farms. *Procedia Comput. Sci.* **2020**, *170*, 539–546. [[CrossRef](#)]
77. White, D.; Svellingen, C.; Strachan, N. Automated measurement of species and length of fish by computer vision. *Fish. Res.* **2006**, *80*, 203–210. [[CrossRef](#)]
78. Shafry, M.R.M. FiLeDI framework for measuring fish length from digital images. *Int. J. Phys. Sci.* **2012**, *7*, 607–618. [[CrossRef](#)]
79. Muñoz-Benavent, P.; Andreu-García, G.; Valiente-González, J.M.; Atienza-Vanacloig, V.; Puig-Pons, V.; Espinosa, V. Enhanced fish bending model for automatic tuna sizing using computer vision. *Comput. Electron. Agric.* **2018**, *150*, 52–61. [[CrossRef](#)]
80. Palmer, M.; Álvarez-Ellacuría, A.; Moltó, V.; Catalán, I.A. Automatic, operational, high-resolution monitoring of fish length and catch numbers from landings using deep learning. *Fish. Res.* **2021**, *246*, 106166. [[CrossRef](#)]
81. Huang, K.; Li, Y.; Suo, F.; Xiang, J. Stereo vision and mask-rcnn segmentation-based 3D points cloud matching for fish dimension measurement. In Proceedings of the 2020 39th Chinese Control Conference (CCC), Shenyang, China, 27–29 July 2020; pp. 6345–6350. [[CrossRef](#)]
82. Bolya, D.; Zhou, C.; Xiao, F.; Lee, Y.J. Yolact: Real-time instance segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9157–9166. [[CrossRef](#)]
83. Wang, X.; Kong, T.; Shen, C.; Jiang, Y.; Li, L. Solo: Segmenting objects by locations. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; pp. 649–665. [[CrossRef](#)]
84. Fernandes, A.F.; Turra, E.; de Alvarenga, R.; Passafaro, T.L.; Lopes, F.B.; Alves, G.F.; Singh, V.; Rosa, G.J. Deep Learning image segmentation for extraction of fish body measurements and prediction of body weight and carcass traits in Nile tilapia. *Comput. Electron. Agric.* **2020**, *170*, 105274. [[CrossRef](#)]
85. Yu, X.; Wang, Y.; Liu, J.; Wang, J.; An, D.; Wei, Y. Non-contact weight estimation system for fish based on instance segmentation. *Expert Syst. Appl.* **2022**, *210*, 118403. [[CrossRef](#)]
86. Chen, F.; Sun, M.; Du, Y.; Xu, J.; Zhou, L.; Qiu, T.; Sun, J. Intelligent feeding technique based on predicting shrimp growth in recirculating aquaculture system. *Aquac. Res.* **2022**, *53*, 4401–4413. [[CrossRef](#)]
87. Liu, J.; Bienvenido, F.; Yang, X.; Zhao, Z.; Feng, S.; Zhou, C. Nonintrusive and automatic quantitative analysis methods for fish behaviour in aquaculture. *Aquac. Res.* **2022**, *53*, 2985–3000. [[CrossRef](#)]
88. Zhou, C.; Xu, D.; Chen, L.; Zhang, S.; Sun, C.; Yang, X.; Wang, Y. Evaluation of fish feeding intensity in aquaculture using a convolutional neural network and machine vision. *Aquaculture* **2019**, *507*, 457–465. [[CrossRef](#)]
89. Sun, L.; Wang, B.; Yang, P.; Wang, X.; Li, D.; Wang, J. Water quality parameter analysis model based on fish behavior. *Comput. Electron. Agric.* **2022**, *203*, 107500. [[CrossRef](#)]
90. Måløy, H.; Aamodt, A.; Misimi, E. A spatio-temporal recurrent network for salmon feeding action recognition from underwater videos in aquaculture. *Comput. Electron. Agric.* **2019**, *167*, 105087. [[CrossRef](#)]
91. Xu, W.; Zhu, Z.; Ge, F.; Han, Z.; Fengli, G. Analysis of Behavior Trajectory Based on Deep Learning in Ammonia Environment for Fish. *Sensors* **2020**, *20*, 4425. [[CrossRef](#)]
92. Han, F.; Zhu, J.; Liu, B.; Zhang, B.; Xie, F. Fish shoals behavior detection based on convolutional neural network and spatio-temporal information. *IEEE Access* **2020**, *8*, 126907–126926. [[CrossRef](#)]
93. Wang, G.; Muhammad, A.; Liu, C.; Du, L.; Li, D. Automatic Recognition of Fish Behavior with a Fusion of RGB and Optical Flow Data Based on Deep Learning. *Animals* **2021**, *11*, 2774. [[CrossRef](#)]
94. Wang, M.; Deng, W.; Wang, M.; Deng, W. Deep visual domain adaptation: A survey. *Neurocomputing* **2018**, *312*, 135–153. [[CrossRef](#)]
95. Chen, J.C.; Chen, T.-L.; Wang, H.-L.; Chang, P.-C. Underwater abnormal classification system based on deep learning: A case study on aquaculture fish farm in Taiwan. *Aquac. Eng.* **2022**, *99*, 102290. [[CrossRef](#)]
96. Darapaneni, N.; Sreekanth, S.; Paduri, A.R.; Roche, A.S.; Murugappan, V.; Singha, K.K.; Shenwai, A.V. AI Based Farm Fish Disease Detection System to Help Micro and Small Fish Farmers. In Proceedings of the 2022 Interdisciplinary Research in Technology and Management (IRTM), Kolkata, India, 24–26 February 2022; pp. 1–5. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.