EJFA Emirates Journal of Food and Agriculture

Emirates Journal of Food and Agriculture 37: 1–14
doi: 10.3897/ejfa.2025.145632

RESEARCH PAPER FEEDING THE FUTURE

Design A YOLOv5s-based directionally arranged conveyance method for silver carp*

Peng Wan[1,2,3], Ziying Liu[1,2], Shiyang Guo[1,2], Xiang Zhang[1,2], Yuyi Huang[1,2], Jiayi Huang[2,3]

1 *College of Engineering, Huazhong Agricultural University, Wuhan 430070, China*
2 *Key Laboratory of Aquaculture Facilities Engineering, Ministry of Agriculture and Rural Affairs, Wuhan 430070, China*
3 *Key Laboratory of Modern Agricultural Intelligent Equipment in South China, Ministry of Agriculture and Rural Affairs, Guangzhou 510630, China*

Corresponding author: Peng Wan (wanpeng09@mail.hzau.edu.cn)

Academic editor: Zeynep Ünal ♦ **Received** 30 December 2024 ♦ **Accepted** 18 June 2025 ♦ **Published** 23 October 2025

Abstract

Traditional fish body directional conveyance suffers from high labor intensity and low efficiency in the fish processing process. This study employed machine vision technology to achieve directional identification of fish bodies for a special self-developed device. 2000 images of fish in four directions are selected to establish a database. The data set is finally expanded to 3000 images through image processing. A YOLOv5s object detection model was used to determine the head-tail and dorsal-ventral directions of silver carp by analyzing their morphological features and physical properties to achieve directional arrangement of the fish bodies for the device. The identification results were utilized to manage the fish body direction and successfully accomplished head-tail and dorsal-ventral directional arrangement of the fish. The detection model demonstrated an accuracy of 99.76%, a recall rate of 99.59%, an average precision of 99.5%, and a F1-score of 99.66% on the test dataset. The model is compared with other models in the YOLOv5 series as well as the YOLOv8 model. It is found that the accuracy and recall rates are not much different, but the YOLOv5s model used in this paper is only 14.1MB in size, and the average detection speed of the model is as high as 0.029s. At the same time, the performance test of the device was carried out. The results indicated at the conveying speeds of 0.05, 0.45, and 0.6 m/s respectively for the lifting system, separation system, and direction identification and assignment system, the average success rate could reach 97.2% for head-tail directional arrangement and 95.6% for dorsal-ventral directional arrangement. Additionally, the device achieved a directional conveying throughput of up to 15 fish per minute. These findings may facilitate the design of conveyor devices for the arrangement of fish bodies in freshwater fish processing industry.

Keywords

Silver Carp, Head-Tail Orientation, Dorsal-Ventral Orientation, Object Detection, YOLOv5s

Introduction

Silver carp (*Hypophthalmichthys molitrix*) is one of the four major freshwater fish species in China, and is extensively farmed across various regions of the country. In 2022, the farming yield reached approximately 3.87 million tons (Ouyang et al. 2017; Bofafa 2023). Mechanized processing of silver carp can significantly enhance its economic value. Fish body orientation is critical to mechanized and automated preprocessing of silver carp (Li et al. 2010; Wei et al. 2023). It is currently achieved mainly in a manual manner, which generally has a high labor intensity and low efficiency and cannot meet the demands for mechanized and automated preprocessing of silver carp (Zion et al. 2000; Liu 2013). Therefore, it is crucial to develop an automatic device for head-tail and dorsal-ventral orientation and conveyance of silver carp to improve the automation level of processing and economic benefits.

* Topical Collection "Feeding the Future: Advancements in Deep Learning for Sustainable Agriculture and Food Security", edited by Prof. Yonis Gulzar and Zeynep Ünal.

Mechanical orientation devices

There has been ample research regarding the directional conveyance equipment for fish bodies (Zhou and Xu 2000; Borderias and Sanchez-Alonso 2011), with innovations mainly in the two aspects of head-tail orientation and dorsal-ventral orientation.

Regarding head-tail orientation, Xiang (2016) analyzed the torque of silver carp on an inclined vibrating platform. As the fish body slides down the vibrating inclined surface, the frictional force acting on the fish head generates a torque on the mass center which exceeds that generated by the body, thereby achieving head-tail orientation. The orientation success rate can reach 100%.

On the other hand, regarding the ventral-dorsal orientation, Wan et al. (2020) accomplished dorsal-ventral orientation of the fish body by employing rollers rotating in opposite directions to press the dorsal and ventral of the fish based on the morphological characteristics of the fish.

Under optimal conditions, the orientation success rate could reach 100%. The following devices can perform both head-tail and ventral-dorsal orientation simultaneously. Knyszewki et al. (1991) designed a vibrating platform for the head-tail and dorsal-ventral orientation of surgeonfish, using the reciprocating action of the vibrating platform to achieve desired orientation. Gao (2011) achieved both head-tail and dorsal-ventral orientation of silver carp by utilizing an inclined vibrating surface and a gradually changing slide based on the frictional characteristics of the fish. The overall success rate for head-tail orientation reached 97.53%, while that of dorsal-ventral orientation reached 97.5%.

In summary, some relatively advanced fish body orientation technologies require high implementation costs, while some other technologies primarily rely on mechanical devices, and there have been few specific devices designed for head-tail and dorsal-ventral orientation of fish bodies. The existing methods mainly include vibration-based head-tail orientation, gradient slide-based dorsal-ventral orientation, and roller pressing-based dorsal-ventral orientation. However, these methods often cause damage to the surface of the fish body.

Deep Learning and YOLO in Aquaculture

Recent years have seen the emergence of computer vision and deep learning in aquaculture automation. Dan et al. (2021) utilized the Harris corner detection algorithm in image processing to assist identification of the head and tail of fish, and combined this algorithm with an adjustment device to achieve accurate head-tail orientation of the fish body. Similarly, Ai et al. (2022) designed a fish body orientation device based on the ResNet-18 classification network to recognize the head-tail direction and employed image processing to identify the dorsal-ventral direction. Aziz et al. (2023) introduced a deep neural network for fish body classification, optimized via genetic and cuckoo search algorithms for improved accuracy. Other related works applied deep learning to adjacent domains. For example, Gulzar (2024) used an improved InceptionV3 model for soybean seed classification.

Meanwhile, advanced object detectors like YOLO have been adopted in related tasks. Sun et al. (2023) developed an optimized YOLOv5s detector for real-time river crab identification. Ye et al. (2023) applied an improved YOLOv5 for automated sorting of live crayfish by size and maturity, achieving both high accuracy and real-time inference. Kuswantori and Suroso (2024) used YOLOv8 to detect and classify fish on a moving conveyor belt; their system accurately identified eight fish species at a belt speed of 505.08 m/h. Zhang et al. (2024) developed 'BSS-FISH-YOLOv8', an enhanced YOLOv8n model for underwater fish detection, which achieved higher mAP than the baseline YOLOv8n. Wang et al. (2023) improved the YOLOv5 network for diseased fish detection in intensive aquaculture, highlighting the importance of robust detection models. In general, YOLOv8 incorporates architectural advances (e.g. C2f backbone modules) that make it faster and more accurate than YOLOv5. Transformer-enhanced versions of YOLOv8 have been shown to surpass previous YOLO variants on underwater fish recognition tasks. These results underline the strong potential of YOLO-series models for aquatic object detection and sorting.

In summary, deep learning–driven computer vision has been successfully applied across a range of aquaculture tasks. Despite these advances, the specific challenge of fully automated fish orientation—critical for mechanical processing—remains under-addressed, as most existing methods still depend on simple visual cues that lack robustness and generalization in complex production environments.

To effectively address the challenges of automatic fish body alignment and head-tail or dorsal-ventral orientation, we propose a method that integrates a fish body orientation and conveyance device with machine vision. The method utilizes deep learning algorithms to accurately recognize the head-tail and dorsal-ventral directions of the fish body, which greatly improves the direction identification speed and accuracy of the algorithms. YOLOv5s is a single-stage object detection algorithm, offering high detection accuracy, rapid processing speed, and real-time object detection (Lv et al. 2019; Wan et al. 2021; Yang et al. 2023). Therefore, this device employs YOLOv5s to develop a detection model for the head-tail and dorsal-ventral orientation of fish bodies, aiming to achieve automatic and efficient directionally arranged conveyance of fish bodies.

The main contributions of this study are summarized as follows:

- We propose a novel vision-based orientation detection method for silver carp. A multi-angle image dataset of silver carp was constructed and augmented, and a lightweight YOLOv5s model was trained to recognize head–tail and dorsal–ventral direction. The resulting model achieves 99.76% precision and 99.59% recall on test images, demonstrating strong generalization and real-time detection capability.

- We design and implement a compact directional-conveyance device that integrates this vision system with mechanical actuators. The system automatically lifts fish, separates them, and assigns orientation, enabling continuous processing. This combined hardware–software solution effectively advances the automation of fish-body conveyance beyond prior work.
- We conduct systematic experiments to optimize system performance. By varying conveyor speeds in each section, we identify the optimal parameters (lifting: 0.05 m/s; separation: 0.45 m/s; orientation: 0.60 m/s) that maximize orientation success. Under these settings, the device achieves an average 97.2% success rate for head–tail orientation and 95.6% for dorsal–ventral orientation, with a throughput of 15 fish per minute. These results validate the efficiency and robustness of the proposed system.

Together, the above innovations in algorithm and device demonstrate significant improvements in automated fish orientation and conveyance, offering a practical solution for mechanized processing of freshwater fish.

Materials and methods

Physical characteristics of silver carp

Silver carp is a freshwater fish characterized by a laterally compressed, slightly elevated, and fusiform body shape. This study used commercially available silver carp samples (fresh weight between 1.0 and 1.2 kg).

During the test, 100 silver carp samples were randomly selected and placed on a flat plate. The external dimensions of the fish, including the length (L), width (W), and thickness (H), were measured using a vernier caliper (range: 150 mm, precision: 0.1 mm) and a steel ruler (range: 600 mm, precision: 0.1 mm). The test results are presented in Table 1.

Table 1. Measurement results of silver carp external dimensions (mm).

Parameter	Length	Width	Height
Average	474.6±7.3	113.0±4.1	48.3±2.7
Maximum	481.2	115.7	49.2
Minimum	464.0	99.3	47.8

As shown in Table 1, the selected samples exhibited L, W, and H of 464.0–481.2, 99.3–115.7, and 47.8–49.2 mm, respectively. Notably, the average W is more than twice larger than the average H.

We employed a friction coefficient measurement device to measure the friction coefficients of samples on various materials of the conveyor (Simonyan and Zisserman 2015). The friction coefficients of silver carp samples were determined on lawn-pattern conveyor belts, stripe-pattern conveyor belts, and inverted triangular-pattern conveyor belts, as well as between the surfaces of different silver carp samples. For the measurements, 20 samples were randomly selected, and the measured results are presented in Table 2.

Table 2. Friction coefficients of silver carp bodies on different material surfaces.

Surface structure	Coefficient of friction with scale/μ_1	Coefficient of friction for inverse scales/μ_2
Lawn pattern	1.14±0.09	1.63±0.07
Stripe pattern	0.81±0.08	1.07±0.09
Inverted triangular pattern	0.68±0.05	1.29±0.13

Table 2 clearly shows that the small scales of silver carp result in relatively high friction coefficients on various conveyor belt surfaces. Furthermore, there are notable differences in friction coefficient between the directions of the scales and against the scales.

Mechanical design of directionally arranged conveyance device for silver carp

Overall structure of the device

Based on the spindle-shaped body of silver carp and the characteristics of its rough surface, a directionally arranged conveyance device for silver carp was designed based on the physical property measurements of silver carp. The device primarily consists of a lifting system, a separation system, a direction identification and assignment system, and a return system of fish bodies as well as a control system. The overall structure is illustrated in Fig. 1.

The fish body lifting system primarily consists of a stripe conveyor belt and brush rollers. In contrast, the fish body separation system comprises a conveyor belt, a fish body sliding apparatus, and brush rollers. The fish body sliding apparatus is positioned at one end of the fish body separation system, while the brush roller is positioned at the other end. The primary function of this system is to guide the fish bodies from the lifting system onto the conveyor belt and effectively separate the overlapping fish bodies for subsequent identification and orientation. The fish body direction identification and assignment system primarily comprises a conveyor belt, a machine vision acquisition system, a head-tail actuator, and a dorsal-ventral actuator. The machine vision acquisition system includes a CMOS industrial camera (MV-VS120-10GC, 12 megapixels), a light source, and a photoelectric sensor (Omron E3JK-SD30M1). This system is designed to capture images of fish bodies on the conveyor belt and transmit them to the control system. The head-tail actuator and the dorsal-ventral actuator manipulate the fish body by recognizing its head-tail and dorsal-ventral directions. This ensures that the fish body is arranged and conveyed in alignment with the predetermined head-tail and dorsal-ventral directions. The fish body return system consists of a fish body sliding apparatus, a conveyor, and fish slipway. The fish body sliding apparatus is positioned directly beneath the head-tail actuator and is designed.

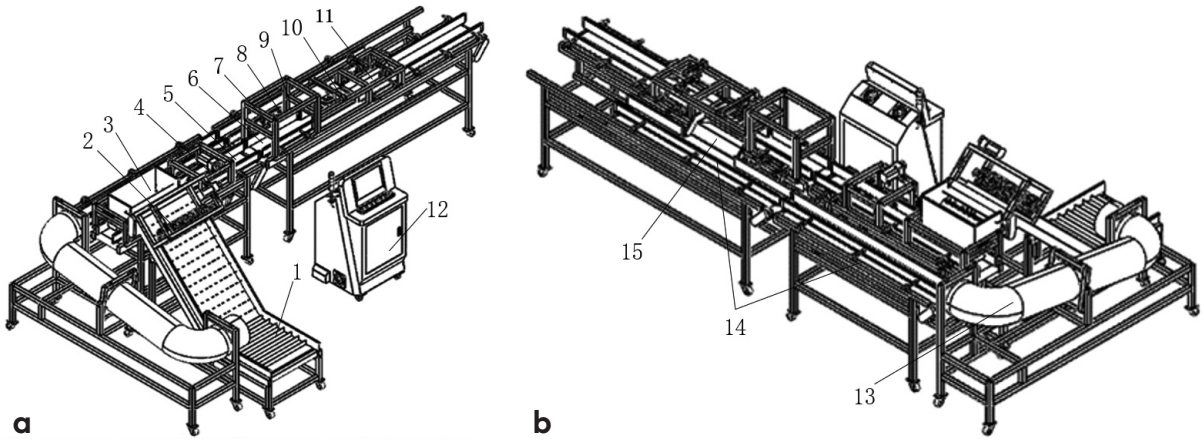

Figure 1. Structure diagram of the directionally arranged conveyance device of silver carp body. **a.** Front axonometric view; **b.** Back axonometric view. 1. Stripe conveyor belt; 2. Brush roller; 3. Fish body sliding apparatus; 4. Conveyor belt; 5. Brush roller; 6. Conveyor belt; 7. Photoelectric sensor; 8. Industrial camera; 9. Light source; 10. Head-tail actuator; 11. Dorsal-ventral actuator; 12. Control system; 13. Fish slipway; 14. Conveyor; 15. Sliding apparatus.

to direct fish with a head-backward direction into the fish body return system. Subsequently, the conveyor and the fish slipway facilitate the convey of the fish back to the fish body lifting system. The control system comprises a computer, a fish body image recognition program, and control circuits for the head-tail and dorsal-ventral actuators. Its primary function is to analyze and recognize the acquired fish body images, and subsequently control the head-tail and dorsal-ventral actuators to effectively manipulate the fish body.

The primary technical specifications of the directionally arranged conveyance device are presented in Table 3.

Table 3. Main technical parameters of the directionally arranged conveyance device of silver carp.

Category	Overall dimension (Length×Width×Height) /mm	Motor Power rating/W	Conveying speed/ ms^{-1}
Fish body lifting system	1 600×580×1300	250	0~0.3
Fish body separation system	2 000×580×950	250	0~0.6
Fish body direction identification and assignment system	3 000×580×950	250	0~0.6
Fish body return system	4 000×1000×750	250	0~0.6

Working principle of the device

In the operation process of the device, the fish is first placed into the fish body lifting system. The fish is initially placed on the stripe conveyor belt, followed by individual lifting facilitated by the brush roller. Upon reaching the apex of the fish body lifting system, the fish descends via a sliding apparatus under the action of gravitational forces, and subsequently enters the fish body separation system. Once entering the separation system, the reverse-rotating brush systematically separates the individual fish and transport them along the conveyor belt. As the fish progresses into the direction identification and assignment system, a photoelectric sensor activates the machine vision acquisition system. This system captures images of the fish during transit and transmits them to the control system. The control system analyzes the acquired images utilizing the YOLOv5s deep learning object detection model to ascertain the head-tail and dorsal-ventral directions of the fish. It then sends instructions to the assignment system via the control circuit. For fish exhibiting a tail-forward and head-backward direction, the assignment system redirects them to the return system, which subsequently conveys them back to the fish body lifting system for re-elevation to correct their head-tail direction. Simultaneously, for fish with a head-forward and tail-backward direction, the fish is further assigned according to their dorsal-ventral directions.

Control system

The control system primarily facilitates the identification of the head-tail and dorsal-ventral directions of fish body images collected during the conveying process. It also regulates the actuators to execute directional assignment based on the identified directions of the fish body.

The control system primarily consists of an Arduino UNO controller, a computer equipped with a detection program, a CMOS industrial camera, actuator cylinders for head-tail orientation, actuator cylinders for dorsal-ventral orientation, solenoid valves (specifically, the 24V210-08 two-position five-way type), opto-coupler isolation relays, photoelectric sensors, and a power supply, among other components. The components of the control system are illustrated in Fig. 2a.

The main controller uses an Arduino development board, which is a microcontroller based on the ATMega328. It includes multiple input/output pins and a development environment, making it suitable for connecting various sensors, actuators, and other electronic components to implement different electronic projects and applications. Additionally, it is cost-effective and connects to a computer via USB, making it convenient for writing and uploading code. This fully meets the requirements of the device.

The CMOS industrial camera can be used in conjunction with the device to capture real-time images of fish

bodies in transit. The CMOS camera supports burst mode and global shutter functions, which give it an advantage in capturing high-speed moving objects and enabling rapid image acquisition.

The workflow of the control system is as follows. first, the software program is initiated, and the relevant parameters are set. Then, the entire device is started. When a silver carp passes the photoelectric sensor of the machine vision acquisition subsystem, it triggers the industrial camera to capture image data of the fish body and sends it to such subsystem. The subsystem identifies the orientation category of the fish and sends corresponding instructions to the Arduino controller. When the Arduino controller receives fish orientation data indicating "The fish with the head-backward and the back-left direction" or " The fish with the head-backward and the back-right direction " and receives the position feedback from the photoelectric sensor in the head-tail actuator, it controls the actuator cylinder to extend. Upon receiving the same fish orientation data again, along with position feedback from the head-tail orientation sensor, the Arduino controller commands the actuator cylinder to retract, then immediately retracts it. When the Arduino controller receives fish orientation data indicating " The fish with the head-forward and the back-left direction " or " The fish with the head-forward and the back-right direction " and receives position feedback from the photoelectric sensor of the dorsal-ventral actuator, it controls the corresponding cylinder in the dorsal-ventral actuator to execute the required action. The whole control process is illustrated in Fig. 2b.

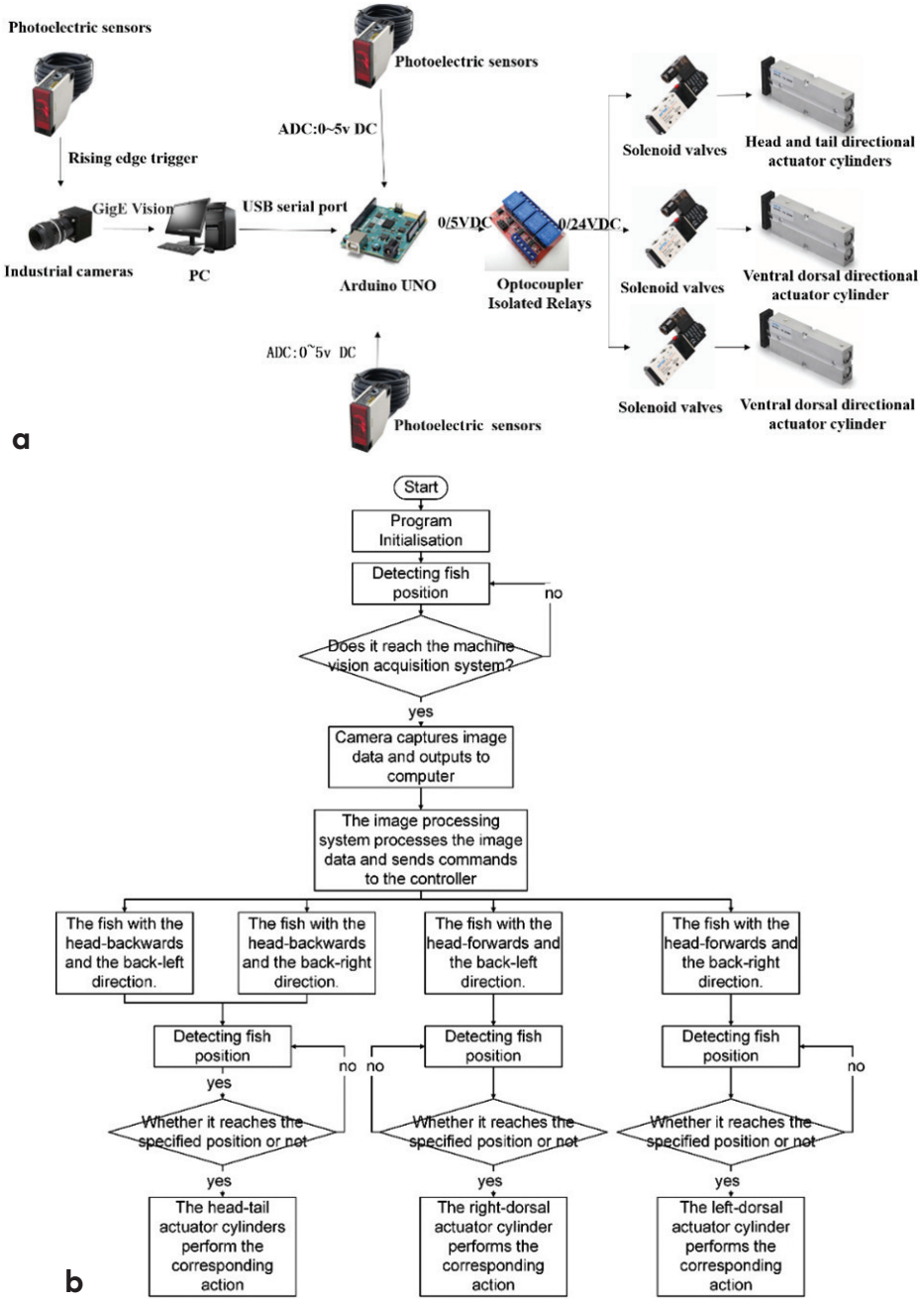

Figure 2. Control system components and control flow diagram: **a.** Composition of the control system; **b.** Flow chart of the control program of the control system.

A YOLOv5s-based model for detecting the head-tail and dorsal-ventral direction of fish bodies

To accurately and quickly detect the head, tail, dorsal and ventral orientation of the fish body, a YOLOv5s detection model was constructed to identify the fish body.

Model training environment

The detection model for fish body head-tail and dorsal-ventral directions is trained under the Windows 10 operating system. The system is equipped with an Intel(R) Xeon(R) Gold 6248R CPU and an NVIDIA GeForce RTX 3090 GPU. The deep learning framework used is PyTorch, with a software environment of Python 3.6 and CUDA 11.4.

Construction of an image dataset for silver carp

To construct a dataset comprising images of silver carp, we employed the directionally arranged conveyance device for fish body to facilitate image collection. As the silver carp traverses the conveyor belt of the machine vision acquisition system, a photoelectric sensor identifies the fish and activates an industrial camera to capture an image of the fish body. Subsequently, the captured image is transmitted to a computer for storage. Initially, the conveyance speed in the fish body direction identification and assignment system was randomly established within the range of 0 to 0.6 m/s. A total of 500 silver carp samples, with each weighing between 1 and 1.2 kg, were utilized for the experiment. Images of the silver carp were captured in four distinct directions: head forward with the dorsal oriented to the left, head forward with the dorsal oriented to the right, head backward with the dorsal oriented to the left, and head backward with the dorsal oriented to the right. In total, 2,000 images were collected, comprising 500 images for each direction. The resolution of the images was 4,024 × 3,036 pixels, and the images were saved in JPEG format. The images depicting these four directions are presented in Fig. 3.

Preprocessing of the image dataset of silver carp fish bodies

To enhance the generalization capability of the detection model of fish body directions, the detection dataset was enriched, and the risk of overfitting was mitigated through the application of various data augmentation techniques. For a subset of the images, brightness enhancement was applied to 300 images, Gaussian noise was added to another 300 images, and rotation was applied to 400 images, so that the number of images in each category is as equal as possible. The structure is shown in Table 4. The brightness enhancement range was set between 0.2 and 1.5, the amount of Gaussian noise ranged from 0.5 to 1.2, and the rotation angle was set at 20°, as shown in Fig. 4. Ultimately, these techniques ultimately expanded the dataset to a total of 3,000 images.

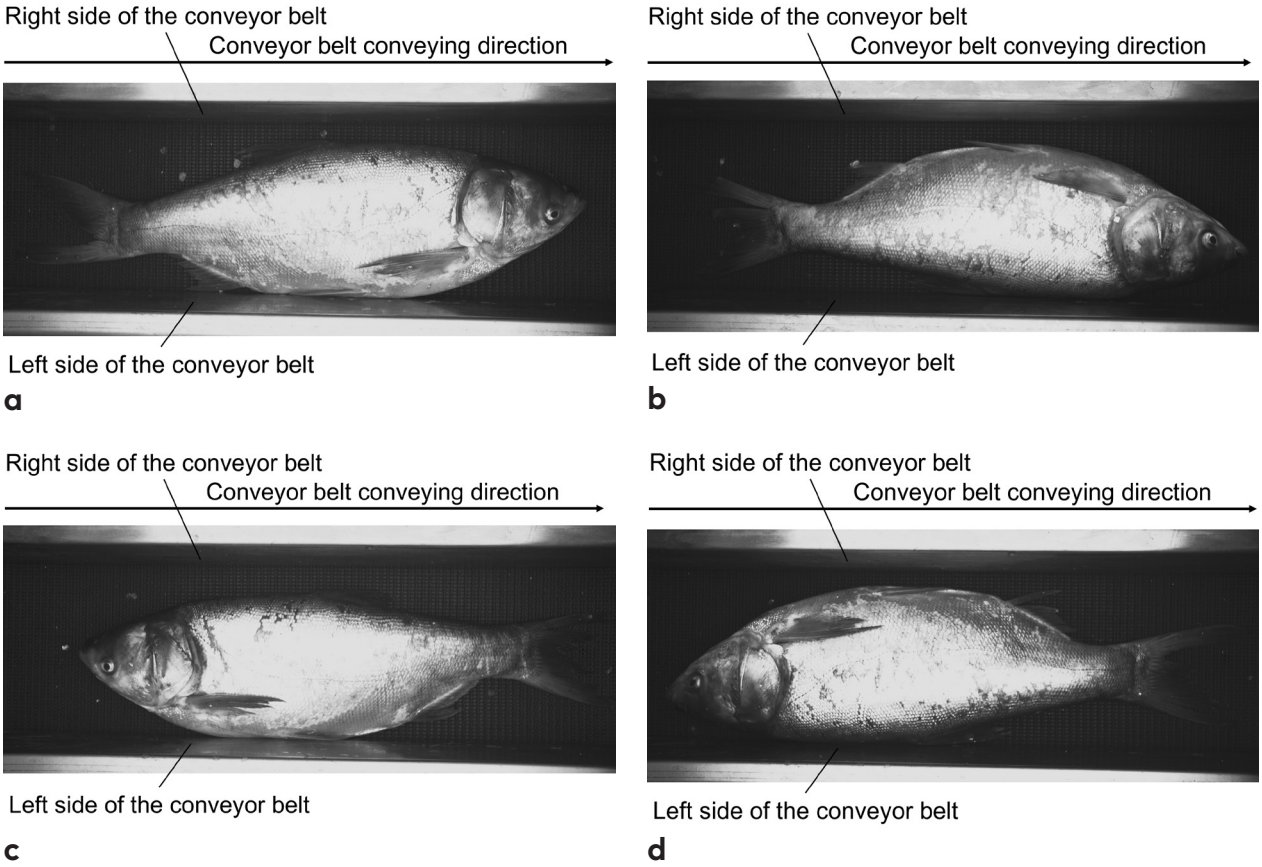

Figure 3. Image of fish body: **a.** The fish with the head-backwards and the back-right direction; **b.** The fish with the head-forwards and the back-left direction; **c.** The fish with the head-forwards and the back-right direction; **d.** The fish with the head-forwards and the back-left direction.

Figure 4. Data processing: **a.** The brightness enhancement; **b.** The Gaussian noise; **c.** The Rotation.

Table 4. Silver carp fish body image data structure.

Category	The head-forwards and the back-right direction	The head-forwards and the back-left direction	The head-backward and the back-right direction	The head-backward and the back-left direction
Source Image	500	500	500	500
Data Augmentation	125	125	125	125
Total quantity	725	725	725	725

The constructed fish body image dataset was labeled utilizing the LabelImg tool. As shown in Fig. 5, the head and back of the silver carp are selected using a rectangular frame. When marking, the rectangular frame is ensured to contain as little background as possible. Subsequently, the labeled dataset was randomly divided into training and testing sets at an 8: 2 ratio.

Detection model of fish body head-tail and dorsal-ventral directions

This paper presents an algorithmic model designed for the real-time detection of the head-tail and dorsal-ventral directions of fish bodies on a directionally arranged conveyance device. Given the relatively uniform backgrounds of fish body images, it is feasible to achieve real-time detection of fish body directions. YOLOv5s is a single-stage object detection algorithm with both high detection speed and accuracy. Our preliminary experiments indicated that YOLOv5s satisfies the performance requirements necessary for the directionally arranged conveyance device, and there is no need to improve the model for the time being. Therefore, we proposed the model for detecting the head-tail and dorsal-ventral directions of fish bodies based on YOLOv5s. The model comprises four components, including Input, Backbone, Neck, and Head (Liu et al. 2018; Tan et al. 2020). The structure of the model is illustrated in Fig. 6.

The Input utilizes a training set derived from a fish body image dataset, where Mosaic data augmentation is applied to the incoming fish body images. Specifically, four fish body images are randomly selected and subjected to various operations, including random scaling, cropping, and stitching, to generate a new composite image, which is subsequently fed into the Backbone. This new image contains information from four smaller images, helping the model to better learn the contextual and background information of the target. This technique increases the diversity of the

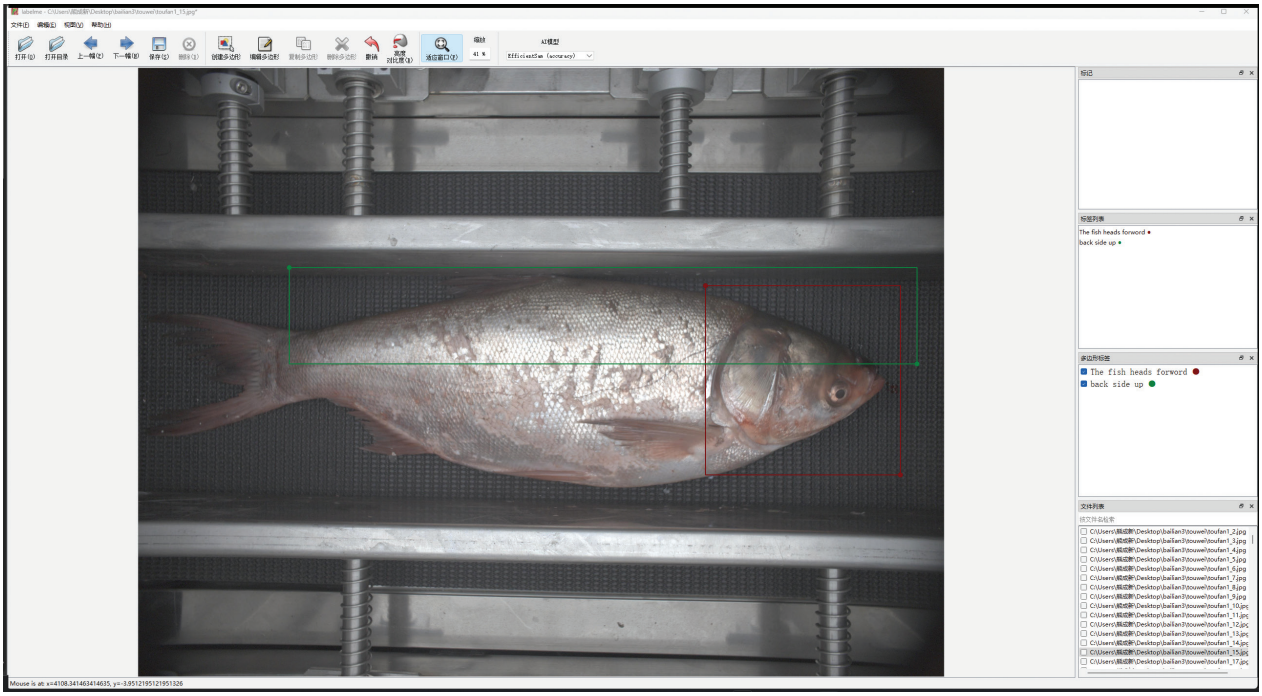

Figure 5. Image annotation effect diagram.

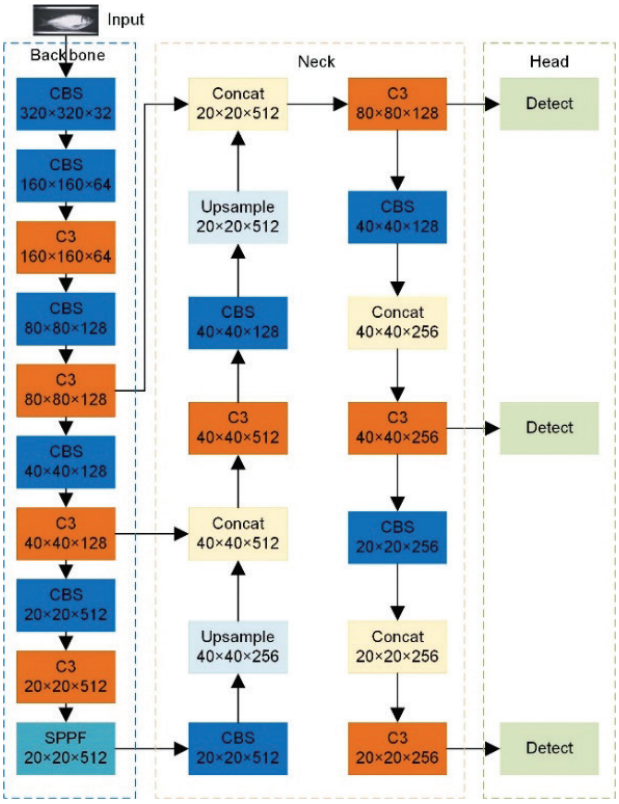

Figure 6. Network structure of the fish body direction detection model. Note: CBS is the convolution unit; the number after C3 indicates the number of modules; SPPF is the spatial pyramid pooling module; Concat is data merging; Up sample is feature up sampling.

training data and enhances the model's ability to detect targets in different scenarios. Additionally, random scaling of the images improves the detection of small targets.

The Backbone includes CBS layer, C3 layer and SPPF layer, which mainly realizes the extraction of fish body features (Tan et al. 2020). The model replaces the CSP module with the C3 module, removing the residual module and convolution layer from the CSP module. Additionally, the activation function of the convolution layer following the concat module is replaced with SiLU, effectively reducing the computational load. The model also uses the SPPF module instead of the SPP module to further improve the network's operational efficiency.

The Neck (Duan et al. 2022) is constructed using a feature pyramid network (FPN) in conjunction with a path aggregation network (PAN). The FPN processes the input image in a top-down manner, generating a set of feature maps at different resolutions. These feature maps are then fed into an upward path, where information from high-resolution feature maps is combined with information from low-resolution feature maps to create a feature pyramid structure. However, FPN can only pass deep-layer feature map information to shallow layers and cannot transfer shallow-layer positional information to deeper layers. PAN addresses this limitation by introducing a bottom-up pathway, supplementing the FPN. By combining the top-down methodology of FPN with the bottom-up methodology of PAN, the network achieves better multi-scale feature fusion, enhancing the detection of fish images at various scales.

Ultimately, the Head outputs the fish body direction category, confidence level, and the location of the detection box. In object detection, an algorithm typically generates numerous candidate boxes around objects in an image. Some of these boxes may overlap or contain redundant information. In such cases, Non-Maximum Suppression (NMS) ranks these candidate boxes based on their confidence scores. NMS is a technique used in computer vision to filter out duplicate or overlapping detections of objects in an image. It selects the candidate box with the highest confidence score and removes all other boxes that overlap with it above a certain threshold. By using NMS, object detection algorithms can improve accuracy and reduce false-positive detection.

The loss function

The loss function of YOLOv5s consists of three parts: classification loss, bounding box loss, and objectness loss. These three parts work together to guide the model to continuously optimize during training by measuring the difference between the model's prediction results and the true labels. The total loss function combines the three parts of the loss by weighted summation to ensure that the model can simultaneously take into account the accuracy of target classification, positioning, and confidence in the target detection task.

Classification loss is used to measure the model's prediction accuracy for the target category. YOLOv5s uses binary cross entropy loss (BCE Loss) as the classification loss function, and calculates the category probability predicted by each anchor box. Classification loss improves the model's ability to recognize the target category by minimizing the difference between the predicted probability and the true label. The formula is as follows:

$$L_{\text{class}} = -\sum_{i=1}^{S^2}\sum_{j=1}^{B} 1_{ij}^{\text{obj}} \left[\hat{p}_j \log\left(p_j\right) + \left(1-\hat{p}_j\right)\log\left(1-p_j\right)\right] \quad (1)$$

In this formulas, S^2 represents the number of grids in the feature map, B represents the number of anchor boxes predicted by each grid, and 1_{ij}^{obj} is an indicator function used to determine whether the current anchor box contains the target.

The positioning loss is used to optimize the model's prediction accuracy for the target bounding box. YOLOv5s uses CIoU loss (Complete Intersection over Union) as the positioning loss function, which comprehensively considers the overlapping area, center point distance, and aspect ratio consistency of the bounding box. CIoU loss can more accurately optimize the regression of the bounding box by introducing geometric factors, especially in scenes with large target overlap and scale changes. Its formula is as follows:

$$L_{\text{box}} = 1 - \text{CIoU} = 1 - \left(\text{IoU} - \frac{\rho^2\left(b, b^{\text{gt}}\right)}{c^2} - \alpha v\right) \quad (2)$$

In this formulas, ρ represents the Euclidean distance between the center point of the predicted box and the

real box, c represents the diagonal length of the minimum bounding box, v is the aspect ratio consistency factor, and α is the weight coefficient.

Confidence loss is used to measure the model's prediction accuracy of whether the target exists or not. YOLOv5s also uses binary cross entropy loss (BCE Loss) as the confidence loss function. Confidence loss ensures that the model can accurately determine whether the anchor box contains the target by minimizing the difference between the predicted value and the true value. The formula is as follows:

$$L_{\text{obj}} = -\sum_{i=1}^{S^2}\sum_{j=1}^{B}\left[\hat{o}_j \log(o_j) + (1-\hat{o}_j)\log(1-o_j)\right] \quad (3)$$

In this formulas, $\hat{O}_j$ represents the confidence of the target existence in the true label, and O_j represents the confidence of the target predicted by the model.

Finally, the total loss function combines the three losses by weighted summation:

$$L_{\text{loss}} = \lambda_{\text{coord}} L_{\text{box}} + \lambda_{\text{obj}} L_{\text{obj}} + \lambda_{\text{class}} L_{\text{class}} \quad (4)$$

In this formulas, λ_{coord}, λ_{obj} and λ_{class} are the weight coefficients of localization loss, confidence loss and classification loss respectively.

In this paper, the three weights are set to $\lambda_{\text{coord}} = 1$, $\lambda_{\text{obj}} = 1$, $\lambda_{\text{class}} = 1$ to balance the contribution of each part of the loss to the model training.

Model evaluation metrics

We employed three indicators to assess the performance of the detection model, including precision, recall mean average precision (mAP), and F1-score, which serve as the evaluation metrics for the algorithm.

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \times 100\% \quad (7)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

In these formulas, TP represents the total number of correctly predicted positive samples; FP denotes the total number of incorrectly predicted positive samples; FN indicates the total number of incorrectly predicted negative samples; AP_i refers to the detection accuracy for the i-th class; and N signifies the total number of classes.

Model training and result analysis

We employed the Stochastic Gradient Descent (SGD) algorithm for model training in this paper. The initial learning rate was set to 0.01, and the initial momentum was established at 0.937. The batch size was configured to 16, and the training process spanned 300 epochs. The model parameters with the highest accuracy, as well as the final model parameters, were trained and saved. This procedure yielded the precision, recall, and average precision curves for the trained model. The training loss of the model is illustrated in Fig. 7.

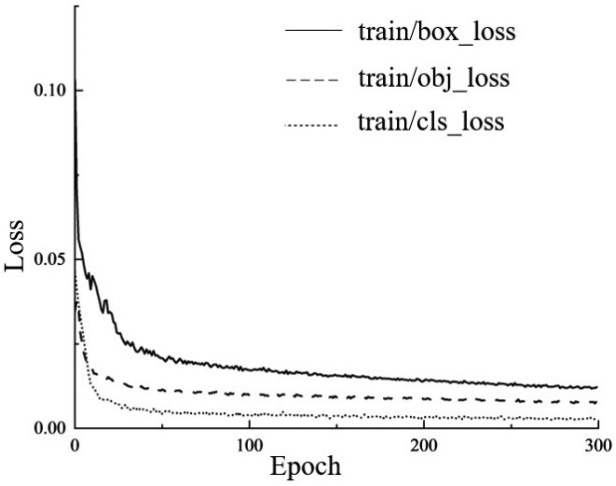

Figure 7. Training loss curve.

As illustrated in Fig. 7, the training loss of the model decreases rapidly during the initial 50 epochs. After 300 epochs of training, the loss curve of the model reaches a plateau, indicating effective convergence of the model.

As illustrated in Fig. 8, during the initial 50 epochs, the precision, recall, and average precision curves of the model increase sharply. Once the epochs reach 300, these curves generally reach a plateau. The final model attains a precision of 99.76%, a recall of 99.59%, and an average detection precision of 99.5%, indicating that the fish body direction detection model not only is effective but also exhibits a high recognition success rate.

As shown in Fig. 9, the normalized confusion matrix is strictly diagonal, indicating that the model has better recognition performance.

Comparison of recognition results from different object detection algorithms

To objectively evaluate the performance of the selected model, we compared the detection results on the test set using the YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x network models. The evaluation metrics employed included precision, recall, F1-score, average precision, average detection speed, and model size. The comparison of performance of various network models is presented in Table 5.

Table 5. Comparison of detection results of different target detection algorithms.

Model	Precision/%	Recall/%	F1-Score/%	mAP/%	Average detection speed/s	Model size/MB
YOLOv5s	99.76	99.56	99.66	99.5	0.029	14.1
YOLOv5m	99.9	99.7	99.80	99.5	0.060	42.2
YOLOv5l	99.8	99.5	99.65	99.5	0.070	92.9
YOLOv5x	99.9	99.7	99.80	99.5	0.078	173.1

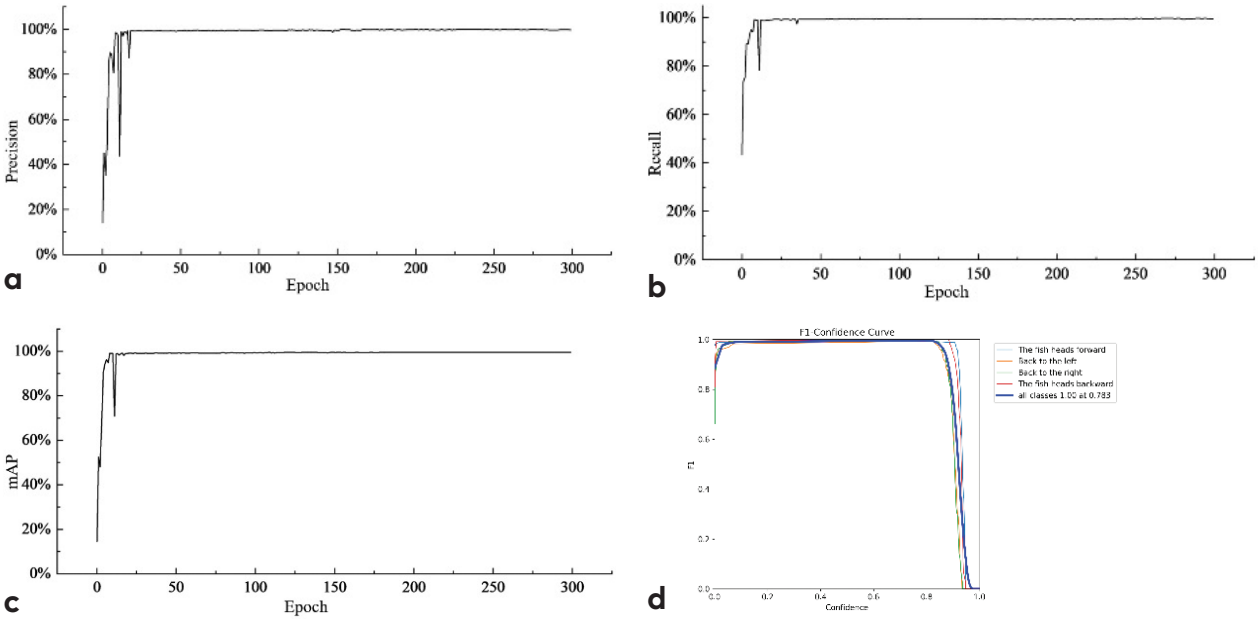

Figure 8. Evaluation indicators for the model: **a.** Precision; **b.** Recall; **c.** mAP; **d.** F1-score.

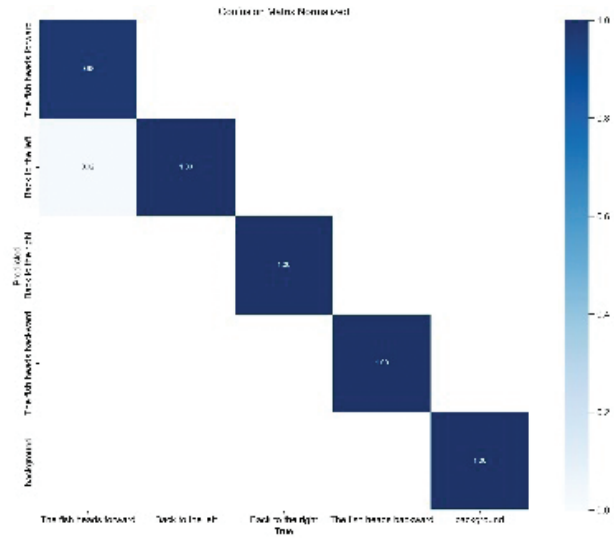

Figure 9. Confusion matrix of the model.

A trade-off curve is a chart used to show the balance between two competing metrics in different models, such as detection speed and accuracy. For object detection tasks, the commonly used evaluation metrics are detection speed (frames per second, FPS) and accuracy (mean Average Precision, mAP). The trade-off of model is shown in Fig. 10.

As shown in Table 5, the precision, recall, and average detection precision of various models show no significant differences. However, the YOLOv5s algorithm employed in this study demonstrates distinct advantages in terms of model size and detection speed relative to other models. The model size is only 14.1 MB, and the average detection speed is 0.029 second. These results indicate that the chosen algorithm offers substantial benefits in both size and speed while maintaining recognition accuracy, making it particularly appropriate for the real-time detection of fish body head-tail and dorsal-ventral directions.

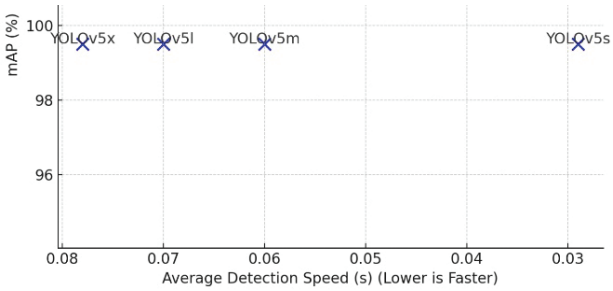

Figure 10. Trade-off between detection speed and accuracy.

Results

Silver carp sample and test equipment

The fish samples used in this study consisted of 100 live silver carp purchased from the market. The samples were fresh and intact in appearance. After acquisition, the samples were placed in a tank for temporary holding. Prior to testing, most of the water in the tank was drained to induce natural death of the fish due to hypoxia while ensuring that the surface of the fish remained moist.

The employed experimental equipment is a self-developed directionally arranged conveyance device for silver carp, which was developed using machine vision processing and based on the a forementioned structure and principles, as illustrated in Fig. 11.

Evaluation indices

We used the success rates of head-tail and dorsal-ventral directions as the evaluation indices. The success rate of head-tail direction is defined as the ratio of the number of fish samples that successfully achieve head-tail direction

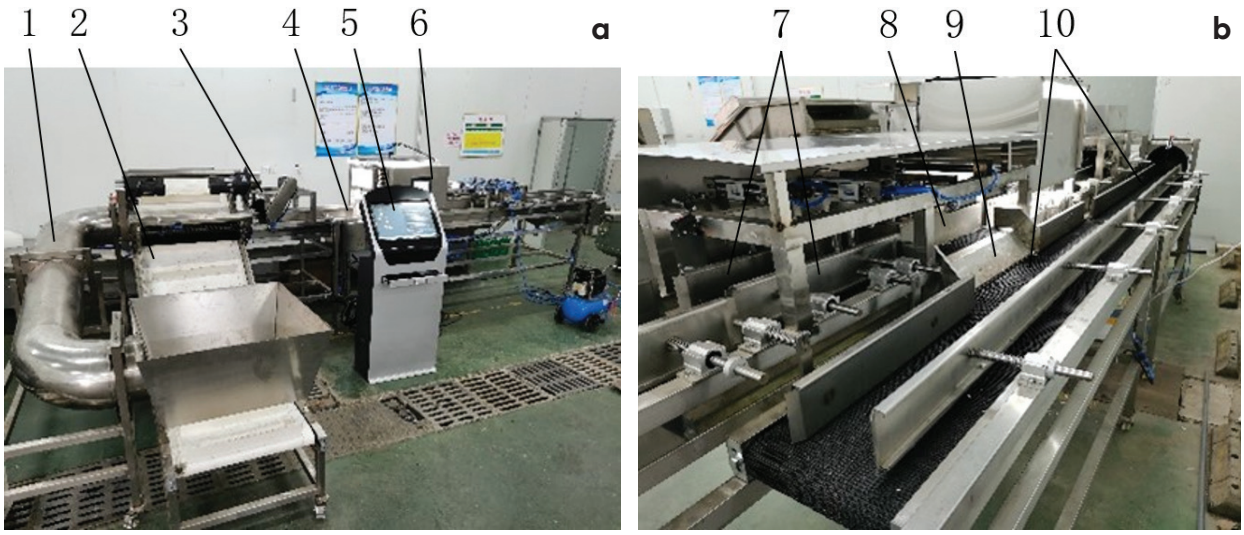

Figure 11. The fabricated directionally arranged conveyance device for silver carp. **a.** Overall device diagram; **b.** Overall device diagram. 1. Fish slipway; 2. Fish body lifting system; 3. Fish body separation system; 4. Fish body direction identification and assignment system; 5. Control system; 6. Machine vision acquisition system; 7. Dorsal-ventral actuator; 8. Head-tail actuator; 9. Sliding apparatus; 10. Fish body return system.

after passing through the device to the total number of samples. The formula is as follows:

$$W_{ht} = \frac{W_{up}}{W_{all}} \times 100\% \quad (9)$$

In the formula, Wht represents the success rate of head-tail direction, which is expressed as a percentage; Wup denotes the number of fish samples with successful head-tail direction in each trial; while Wall indicates the total number of fish samples in each trial.

The success rate of dorsal-ventral direction is defined as the ratio of the number of fish samples that successfully achieve dorsal-ventral direction after passing through the device to the total number of samples. The calculation formula is as follows:

$$N_{ad} = \frac{N_{up}}{N_{all}} \times 100\% \quad (10)$$

In the formula, Nad represents the success rate of dorsal-ventral direction, which is expressed as a percentage; Nup denotes the number of fish samples that successfully achieve dorsal-ventral direction in each trial; while Nall refers to the total number of samples in each trial.

Optimization of the operation parameters

The recognition accuracy of the YOLOv5s detection model directly determines the success rate of detecting the head-tail and dorsal-ventral directions of fish samples. The machine vision detection system captures images of individual fish bodies, which is crucial for the accurate detection by the YOLOv5s model. This necessitates the effective separation and convey of single fish bodies. The conveying speed of the fish in the lifting system, separation system, and direction identification and assignment system of fish bodies significantly influence the effectiveness of separating individual fish bodies. Therefore, we evaluated the success rates of head-tail and dorsal-ventral directions as key performance indicators based on the conveying speed of the abovementioned systems to optimize the operation parameters for the device (Howard et al. 2017; Liu et al. 2018; Tan et al. 2020). The factor-level coding table for the test is presented in Table 6.

Table 6. Table of test factors for optimizing the operation parameters of the device.

Levels	Lifting device conveying speed/ (ms^{-1})	Separation device conveying speed/ (ms^{-1})	Orientation device conveying speed/ (ms^{-1})
1	0.05	0.40	0.50
2	0.10	0.45	0.55
3	0.15	0.50	0.60

Ninety silver carp samples were randomly selected, with thirty samples being allocated to each factor combination test. The average value was calculated and repeated three times. The test results are presented in Table 7.

As shown in Table 7, the primary and secondary factors influencing the success rate of fish head-tail and dorsal-ventral directions are the conveying speed of the fish in the lifting system, separation system, and direction identification and assignment system of fish bodies. The optimal parameter combination is $A_1B_2C_3$ for head-tail direction, and $A_2B_2C_3$ for dorsal-ventral direction. Since our major aim is to ensure a high success rate of both head-tail and dorsal-ventral directions of fish bodies, a comprehensive trade-off approach is applied to obtain the optimal combination. The optimum success rate for fish body direction can be achieved with the head-tail and dorsal- ventral direction success rate could reach 97.8% and 95.5% respectively, when the conveying speed of the

Table 7. Test results for optimizing the operation parameters of the device.

	Test No.	*A*	*B*	*C*	Empty column	Head-to-tail orientation success rate/%	Dorsal-ventral orientation success rate/%
	1	1	1	1	1	84.4	78.9
	2	1	2	2	2	86.7	86.7
	3	1	3	3	3	93.3	90
	4	2	1	2	3	80	83.3
	5	2	2	3	1	97.8	95.5
	6	2	3	1	2	83.3	83.3
	7	3	1	3	2	76.7	76.7
	8	3	2	1	3	73.3	73.3
	9	3	3	2	1	73.3	76.7
Head-tail orientation success rate	k_1	88.1	80.4	80.3			
	k_2	87	85.9	80			
	k_3	74.4	83.3	89.3			
	R	13.7	5.5	9.3			
Dorsal-ventral orientation success rate	k_1	85.2	79.7	78.5			
	k_2	87.4	85.2	82.2			
	k_3	75.4	83.3	87.4			
	R	12	5.5	8.9			

Note: *A*, *B* and *C* are the conveying speed of the fish in the lifting system, separation system, and direction identification and assignment system, respectively.

lifting system, separation system, and direction identification and assignment system being set at 0.05, 0.45, and 0.60 m/s, respectively.

Verification of the optimal operation parameters

According to the optimization results for the best operation parameters of the device, 100 silver carp samples were selected for validation of the optimal operation parameters. The conveying speed of the lifting system, separation system, and direction identification and assignment system of fish bodies was set at 0.05, 0.45, and 0.6 m/s, respectively. The average of five tests was calculated, and the results are presented in Table 8.

Table 8. Validation test results.

Test No.	Head-tail orientation success rate/%	Dorsal-ventral orientation success rate/%
1	97.0	96.0
2	98.0	98.0
3	96.0	95.0
4	98.0	95.0
5	97.0	94.0
Average	97.2	95.6

The experimental results indicated that the average success rate is 97.2% and 95.6% for head-tail and dorsal-ventral direction under the optimal parameter combination, respectively. Additionally, the direction efficiency reached 15 fish per minute, which is closely in line with the optimization results, thereby confirming the reliability of the parameter optimization.

Discussion

Study limitations and future work

This study has several limitations that should be acknowledged. First, the image dataset was constructed using a limited sample size and only one fish species (silver carp), which may constrain the generalizability of the detection model to other species with different morphological characteristics. Second, all experiments were conducted under controlled laboratory conditions; as such, the robustness and reliability of the system in real-world industrial environments have yet to be validated. Third, the current system is limited by its hardware configuration, particularly in terms of sorting throughput and long-term operational stability. Future efforts will focus on optimizing the device structure and upgrading key components, including the actuators and sensors, to support high-speed processing and prolonged industrial use. Future efforts will focus on optimizing the device structure and upgrading key components, including the actuators and sensors, to support high-speed processing and prolonged industrial use.

In practice, some orientation failures were observed. These can be attributed to slight discrepancies in the conveyor speeds among different subsystems, which occasionally result in insufficient spacing between adjacent fish. Consequently, the industrial camera may capture images containing more than one fish, leading the YOLOv5s detection model to generate incorrect predictions and trigger inappropriate actuator responses. Regarding this issue, we will explore advanced sensing configurations—such as stereo or multi-camera arrays and affordable depth sensors—to address occlusion and overlapping fish in high-throughput settings.

Moreover, due to the emphasis on end-to-end system performance and constraints in time and computational resources, this study did not include ablation experiments to isolate and evaluate the impact of individual network components or data augmentation strategies. Nor was cross-validation performed, as the available dataset size was limited. Instead, we adopted a fixed training–testing split to assess the detection performance. In future work, we plan to conduct systematic ablation studies and apply k-fold cross-validation to obtain a more comprehensive evaluation of the model's robustness and generalization ability.

Finally, although the current system employs an Arduino development board, which offers cost-effective integration and supports a wide range of sensors, it may not meet the high-precision and long-term reliability requirements of industrial settings. In future work, we intend to develop a custom microcontroller-based system to enhance performance and adaptability for commercial-scale deployment.

Commercial application prospects

The system's modular architecture—featuring standard industrial cameras, pneumatic actuators, and a compact YOLOv5s inference engine—enables straightforward integration into existing fish-processing lines. In single-lane tests, we have demonstrated a throughput of 15 fish per minute; by deploying parallel lanes, processors can multiply this capacity without necessitating fundamental redesigns. Moreover, our preliminary cost–benefit analysis suggests that, in a mid-size facility, the combined capital and operating expenses can be fully recovered within 12–18 months via reductions in labor costs and product rework. Leveraging off-the-shelf components not only lowers the initial investment but also simplifies maintenance and spare-parts management, further enhancing the solution's commercial viability.

Conclusion

Based on the detection of morphological and physical characteristics of silver carp and in conjunction with the directionally arranged conveyance device, this study designs a method for the directional arrangement and conveyance of silver carp bodies using machine vision. By analyzing and experimentally verifying the performance of the method, the main conclusions are as follows.

1) A detection model for the direction of silver carp bodies was constructed based on YOLOv5s. In detecting and recognizing the direction of silver carp bodies, the model could achieve an accuracy of 99.76%, a recall of 99.59%, an average precision of 99.5%, and a F1-score of 99.66%.
2) Optimization of the conveying speeds of various systems in terms of direction success rate of fish bodies indicated that when the lifting system, separation system, and direction identification and assignment system of fish bodies were set at a speed of 0.05, 0.45, and 0.6 m/s, respectively, the device showed the optimal performance.
3) The verification test results demonstrate that under the optimal parameter combination, the average success rate for head-tail and dorsal-ventral direction of fish bodies is 97.2% and 95.6%, respectively, with a direction efficiency reaching 15 fish per minute.

Author contributions

Conceptualization: PW, YH. Data curation: JH. Formal analysis: ZL. Supervision: PW. Validation: XZ, PW, YH, SG, ZL. Visualization: PW. Writing – original draft: ZL. Writing – review and editing: XZ, SG, JH, ZL. .

Acknowledgements

This research was funded by the Key Laboratory of Modern Agricultural Intelligent Equipment in South China, Ministry of Agriculture and Rural Affairs, P.R.China, grant number 2023BBA001, and by the Key Research and Development Program of Hubei Province of China, grant number 2023BBB115.

References

Ai Z, Tan H, Wan P (2022) Design and experiment of a mackerel orientation device based on machine vision. Journal of Huazhong Agricultural University 41(2): 208–216. https://doi.org/10.13300/j.cnki.hnlkxb.2022.02.025

Aziz RM, Desai NP, Baluch F (2023) Computer vision model with novel cuckoo search based deep learning approach for classification of fish image. Multimedia Tools and Applications 82: 3677–3696. https://doi.org/10.1007/s11042-022-13437-3

Borderias AJ, Sanchez-Alonso I (2011) First processing steps and the quality of wild and farmed fish. Journal of Food Science 76(1): 1–5. https://doi.org/10.1111/j.1750-3841.2010.01900.x

Bureau of Fisheries and Fisheries Administration Ministry of Agriculture (2023) China Fisheries Statistical Yearbook 2023. China Agriculture Press, Beijing. https://www.stats.gov.cn/sj/ndsj/2023/indexeh

Dan H, Zhang C, Li C, Zhao H, Tian Y (2021) Design and test of directional conditioning feeding system for fish processing and production. Journal of Chinese Agricultural Mechanization 42(3): 91–97. https://doi.org/10.13733/j.jcam.issn.2095-5553.2021.03.013

Duan J, Wang Z, Zou X, Yuan H, Huang G, Yang Z (2022) Recognition of bananas to locate bottom fruit axis using improved YOLOv5. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE) 38(19): 122–130. https://doi.org/10.11975/j.issn.1002-6819.2022.19.014

Fang W, Wang L, Ren P (2020) Tinier-YOLO: A real-time object detection method for constrained environments. IEEE Access 8: 1935–1944. https://doi.org/10.1109/ACCESS.2019.2961959

Gao X (2011) Research on the device of setting the freshwater fishes' heads and backs in the desired direction. PhD Thesis. Huazhong Agricultural University, Hubei, China.

Gulzar Y (2024) Enhancing soybean classification with modified inception model: A transfer learning approach. Emirates Journal of Food and Agriculture 36: 1–9. https://doi.org/10.3897/ejfa.2024.122928

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June. https://doi.org/10.1109/CVPR.2016.90

Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv. https://doi.org/10.48550/arXiv.1704.04861

Kuswantori A, Suroso DJ (2024) Detection and classification using YOLOv8 for automated sorting systems. Indonesian Journal of

Artificial Intelligence and Data Mining (IJAIDM) 7(2): 447–159. https://doi.org/10.24014/ijaidm.v7i2.30967

Li L, Zong L, Wang J, et al. (2010) Research status and development trend of massive freshwater fish pre-treatment processing technology and equipment. Fisheries Modernization 37(5): 43–46. https://doi.org/10.3969/j.issn.1007-9580.2010.05.011

Liu S, Qi L, Qin H, Shi J, Jia J (2018) Path aggregation network for instance segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1803.01534. https://doi.org/10.1109/CVPR.2018.00913

Liu W (2013) Equipment Manufacture and Orientation Mechanism Study on Freshwater Fishes' Heads and Backs. PhD Thesis. Huazhong Agricultural University, Hubei, China. https://doi.org/10.7666/d.Y2394446

Lv S, Lu S, Li Z, Hong T, Xue Y, Wu B (2019) Orange recognition method using improved YOLOv3-LITE lightweight neural network. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE) 35(17): 205–214. https://doi.org/10.11975/j.issn.1002-6819.2019.17.025

Ouyang J, Shen J, Zheng X, Huang Y, Xu H (2017) Research and application status and development tendency of aquatic products processing equipment. Fisheries Modernization 44(5): 6.

Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556. https://doi.org/10.48550/arXiv.1409.1556

Sun Y, Yuan B, Zhao T, Sun J, Fang Z, Zhao D (2023) Detecting river crab and bait using improved YOLOv5s. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE) 39(22): 178–187. https://doi.org/10.11975/j.issn.1002-6819.202308197

Tan M, Pang R, Le Q (2020) Efficient Det: Scalable and efficient object detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/CVPR42600.2020.01079

Tang Y, Zhou H, Wang H, Zhang Y (2023) Fruit detection and positioning technology for a *Camellia oleifera* C. Abel orchard based on improved YOLOv4-tiny model and binocular stereo vision. Expert Systems with Applications 211: 118573. https://doi.org/10.1016/j.eswa.2022.118573

Wan P, Huang J, Tan H, Zhu M, Zhao J (2020) Design and test of orientation device for freshwater fishes' abdomen and back. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE) 36(6): 28–35. https://doi.org/10.11975/j.issn.1002-6819.2020.06.004

Wan P, Guo S, Yang J, Zhao J, Tan H, Zhu M (2021) Horizontal reciprocating vibration method for head-to-tail directional transportation of fresh water fish. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE) 37(4): 40–48.

Wan P, Huang Y, Wang R, Li M, Xiao C, Wu W (2023) Design and experiments of the machine vision-based body orientation arrangement convey device for the *Scomber japonicus*. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE) 39(14): 271–282. https://doi.org/10.11975/j.issn.1002-6819.202304120

Wang Z, Liu H, Zhang G, Yang X, Wen L, Zhao W (2023) Diseased fish detection in the underwater environment using an improved yolov5 network for intensive aquaculture. Fishes 8: 169. https://doi.org/10.3390/fishes8030169

Wei L, Li L, Ma R, Bai S, Zhang A, Niu K (2023) Research progress on key technologies of primary fish processing equipment. Modern Agricultural Equipment 44(1): 2–7. https://doi.org/10.3969/j.issn.1673-2154.2023.01.001

Xiang Y (2016) Vibration orientation laws and experimental research of head and tail of freshwater fish. PhD Thesis. Huazhong Agricultural University, Hubei, China. https://doi.org/10.7666/d.Y3052716

Yang J, Wan P, Huang Y, Zhu M, Tan H, Wu W, Ding A (2023) Design and test of silver carp head processing device based on machine vision. Journal of Huazhong Agricultural University 42(1): 178–187. https://10.13300/j.cnki.hnlkxb.2023.01.022

Ye X, Liu Y, Zhang D, Hu X, He Z, Chen Y (2023) Rapid and accurate crayfish sorting by size and maturity based on improved YOLOv5. Applied Sciences 13(15): 8619. https://doi.org/10.3390/app13158619

Zhang Z, Qu Y, Wang T, Rao Y, Jiang D, Li S, Wang Y (2024) An improved YOLOv8n used for fish detection in natural water environments. Animals 14(14): 2022. https://doi.org/10.3390/ani14142022

Zhou T, Xu H (2000) Investigation on the status of Nordic fisheries equipment. Fisheries Modernization (6): 3–6. https://10.3969/j.issn.1007-9580.2000.06.001

Zion B, Shklyar A, Karplus I (2000) *In-vivo* fish sorting by computer vision. Aquacultural Engineering: An International Journal 22(3): 165–179. https://doi.org/10.1016/S0144-8609(99)00037-0