Proceedings of 2022 IEEE
International Conference on Mechatronics and Automation
August 7 - 10, Guilin, China

Study on Real-time Recognition of Underwater Live Shrimp by the Spherical Amphibious Robot Based on Deep Learning

Shaolong Wang[1], Jian Guo[1,2*]
[1]Tianjin Key Laboratory for Control Theory& Applications In Complicated systems and Intelligent Robot Laboratory
Tianjin University of Technology
BinshuiXidao Extension 391,Tianjin,300384,China
Fuqiang6369@hotmail.com; 171029492@qq.com

Shuxiang Guo[1,2,3*]
[2] Shenzhen Institute of Advanced Biomedical Robot Co., Ltd.
No.12, Ganli Sixth Road, Jihua Street, Longgang District, Shenzhen,518100,China
*corresponding author : jianguo@tjut.edu.cn

Qiang Fu[1]
[3]Key Laboratory of Convergence Medical Engineering System and Healthcare Technology, The Ministry of Industry and Information Technology, School of Life Science Beijing Institute of Technology
No.5, Zhongguancun South Street, Beijing,100081, China
*corresponding author : guoshuxiang@hotmail.com

Jigang Xu[4*]
[4]Unit68709
Qinghai Haidong,810700, China
*corresponding author : xujigang216@163.com

Abstract **–In this paper, spherical robots are used for the detection and identification of lobsters in aquaculture. Lobster farmers are often faced with tasks such as observation, feeding, and fishing, which are all done manually, with low efficiency and high operating costs. Therefore, this paper proposes a real-time underwater lobster detector based on Generative Adversarial Networks and Convolutional Neural Networks, implemented by a spherical amphibious robot. Firstly, the underwater lobster image dataset is established, and the improved GAN algorithm and data increment method are used for data enhancement preprocessing. Secondly, the single-shot multi-frame detector (SSD) is improved as follows, using the lightweight network MobileNetV2 as the backbone of the SSD network; in the network prediction layer, using depthwise separable convolution instead of standard convolution to accelerate inference; compressing the fully connected layer The parameters construct a lightweight model. Finally, the model is trained on the underwater lobster dataset and deployed on a spherical amphibious robot, and the changes in the loss function value during training before and after image enhancement and algorithm improvement are plotted. Two sets of experimental test results show that the model optimizes the target recognition accuracy of underwater lobsters, and the recognition accuracy reaches 90.32%. The reduced model size facilitates model deployment and is only 24MB in size. The model has good stability and high recognition accuracy in identifying lobsters in complex situations.**

Index Terms –Intelligent aquaculture, Live shrimp detection, Shrimp dataset, Underwater detection, Mobilnet-V2 SSD.

I. INTRODUCTION

In recent years, with the application of artificial intelligence in the fields of image, machine vision and speech recognition becoming more and more mature, deep learning has become one of the research hotspots and mainstream development directions in this field. Procambarus clarkii (crayfish) is an important aquaculture species. Due to its easy cultivation, significant yield and sufficient nutritional value, the breeding range in eastern Asia has expanded rapidly. However, the current aquaculture methods of shrimp seriously rely on the inefficient artificial farming method, and observe the growth status of shrimp through the fishing grab method. This invasive observation method will disrupt the growth law of shrimp, and it is unable to observe the number of shrimp in a specific place and feed accurately, resulting in inconsistent growth status of shrimp and excessive waste of feed. Therefore, it is very important to improve the intelligent identification and detection of shrimp aquaculture. [1-4]

Many research institutions and scholars have invested a lot of research in the research of underwater intelligent vision in aquaculture. Arthur F.A. Fernandes used deep learning image segmentation techniques to extract body dimensions of Nile tilapia from predicted maps of body weight and ketone body features. Through image analysis, Zhang proposed the fish shoal quality estimation method of principal component analysis calibration factor and neural network, and enhanced underwater vision by using pre operations such as image segmentation and image enhancement, which solved the problem of inconsistent distance between image acquisition equipment. Zhou. Near infrared sensor and fuzzy neural network are combined to build a machine vision system that can intelligently control fish feeding, which can complete accurate feeding in aquaculture, save aquaculture costs and reduce aquaculture water pollution. Chao through computer vision image processing technology, image texture and support vector machine methods are used to classify and screen images of interest in aquaculture, and achieve high classification accuracy.[5-11]

Due to the special nature of underwater target data set, the data set that can be obtained under public conditions is limited, and the image imaging effect of underwater images is poor. Also, because crayfish is good at hiding underwater and the target environment is different, only one small data set can be obtained. In order to solve this problem, image enhancement is needed to expand the data set . MD jahidul Islam realizes the real-time underwater image enhancement model, which is applied to optimize the conditions of global content, color, local texture and style of the image to generate the countermeasure network, and improves the performance of underwater target detection, human posture evaluation and significance prediction .[12-15]

This paper proposes a real-time target algorithm for underwater live shrimp of spherical amphibious robot based on deep learning. (1) An image preprocessing method based on improved generation countermeasure network (GAN) is used to improve the quality of underwater visual scene. (2) Lightweight MobileNet V2 is selected as the backbone of single-stage multi frame predictor (SSD) to realize real-time and stable target detector. (3) In order to make up for the deficiency of SSD learning the same features without network layer, a feature pyramid network (FPN) is used in the algorithm network structure to improve the accuracy of multi-scale target detection. Realize the real-time recognition and detection of underwater moving targets of spherical amphibious robot.

II. Method and Verification

With the rapid development of machine vision technology, real-time detection and observation of intelligent aquaculture becomes possible. It is worth noting that many excellent scholars and institutions not only focus on the underwater identification of shrimp as the premise of fixed feeding, but also focus on the research of feeding behavior, activity trajectory, population number, Stationary rearing and reproduction of shrimp. In this paper, we focus on real-time detection of underwater shrimp. [16-17]

Fig.1 is a schematic diagram of a spherical amphibious robot. In this study, Procambarus clarkii (crayfish) was selected as the target detection object. The crayfish recognition algorithm studied in this paper is deployed on the spherical amphibious robot to complete target recognition. Since it is impossible to observe aquatic animals in real time underwater, it is difficult to obtain effective breeding data. Therefore, the spherical robot is required to complete underwater target shooting and monitoring, which is composed of upper and lower parts.

Fig 2 is the schematic diagram of the system. animals in real time underwater, it is difficult to obtain effective

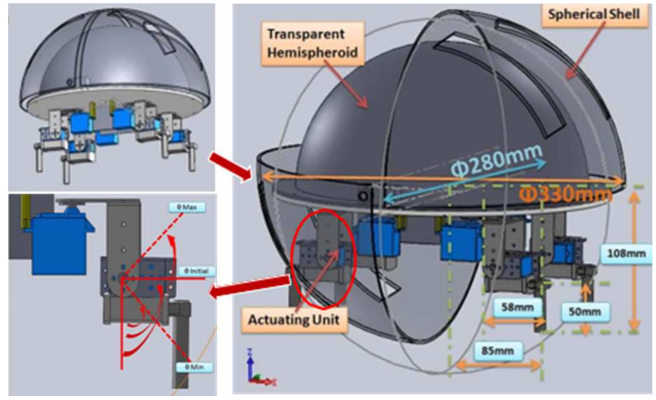

Fig.1 Spherical amphibious robot model

breeding data. Therefore, the spherical robot is required to complete underwater target shooting and monitoring, which is composed of upper and lower parts. Fig 3 shows the internal hardware introduction and overall physical map of the spherical amphibious robot. [18]

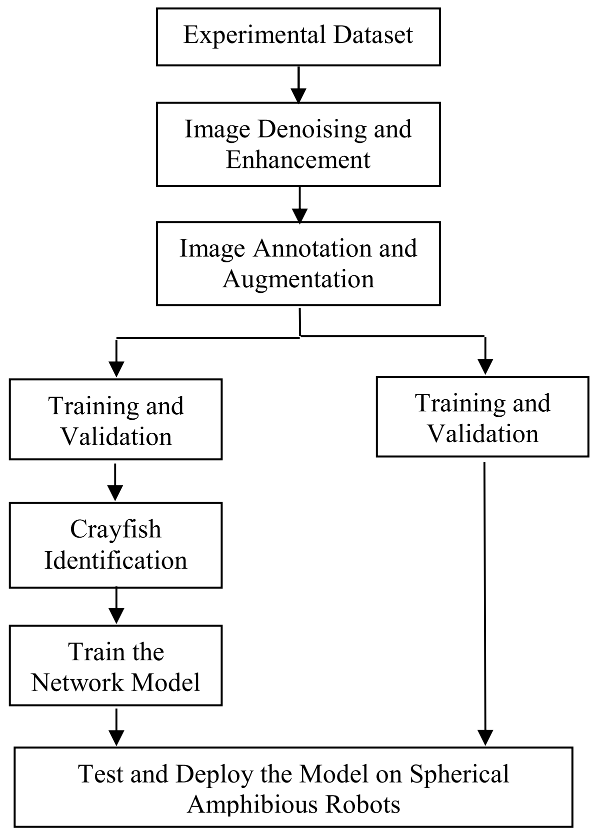

Fig.2 Overall framework of algorithm

The third part mainly introduces the experimental data set, the preprocessing methods of image denoising and enhancement, and the target recognition algorithm based on MobileNet SSD. The fourth part mainly introduces the image enhancement effect, the generalization ability of various shrimp target detection, the detection effect of this algorithm and the comparison with other algorithms. The experimental results in the fifth part show that the optimized target detection network effectively improves the recognition accuracy. [19]

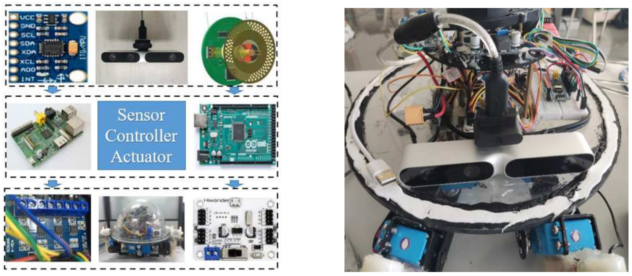

Fig.3 The internal hardware introduction and overall physical map of the spherical amphibious robot

III. Target Recognition Algorithm Based on MobileNet-SSD

We use the fusion method of MobileNet V2 algorithm and SSD algorithm to identify underwater live shrimp in real time. On this basis, we use Funie-Gan to preprocess the image enhancement of the experimental data set, so as to input the processed data set into the MobileNet SSD network for training and recognition, and finally obtain the parameter distribution of the fitted underwater shrimp data set image. Then the training model is deployed on the spherical amphibious robot to practice shrimp recognition.

A. Experimental Data

The experimental data in this paper is composed of ImageNet dataset and self-made dataset, and the storage format is VOC 2007. The self-made dataset consists of spherical robot camera and crayfish images downloaded from Baidu and Google. Using image rotation, mirroring, cropping, brightness/contrast transformation and other dataset enhancement methods, 8052 photos are finally obtained. Parts are shown in Fig 4.

Fig.4 partial experimental data set

B. Image Denoising and Enhancement

Combining the model of deep convolution neural network (CNN) and generative adversary network (GAN), we can learn from paired and unpaired data and improve the image perception performance. Fig 5 shows the denoising and enhancement results of some images of the experimental data set in this paper. In this paper, the fast real-time underwater image enhancement model based on convolution condition (FUnIE GAN) proposed by Islam MJ is used as the underwater image algorithm. Improve the perception of underwater image contrast target through the global content, color, local texture and style information of the image. The conditional Gan loss function is shown in equation 1. Quantitative perception of image quality is divided into three parts: (1) local texture and style. (2) Global similarity loss function L1 (g). (3) Image content loss function Lcon (g). [20]

$$L_{cGAN}(G,D) = E_{X,Y}[\log D(Y)] + E_{X,Y}[\log(1-D(X,G(X,Z)))] \tag{1}$$

$$L_1(G) = E_{X,Y,Z}[\| Y - G(X,Z) \|_1] \tag{2}$$

$$L_{con}(G) = E_{X,Y,Z}[\| \Phi(Y) - \Phi(G(X,Z)) \|_2] \tag{3}$$

Its essence is the confrontation between generator g and discriminator D. In the process of training, generator g is always minimized, on the contrary, discriminator D is always maximized to form confrontation. The experimental data in this paper are paired images, so the paired training loss function is shown in equation (4).

$$G^* = \arg\min_G \max_D L_{cGAN}(G,D) + \lambda_c L_{con}(G) \tag{4}$$

Fig.5 image denoising and enhancement

C. SSD Detector

Single stage multi frame detector (SSD) is a detection algorithm for many categories. The detector combines the two main advantages of the anchor box mechanism of fast RCNN algorithm and the regression idea of Yolo algorithm, makes multi-scale prediction on multiple feature maps with different resolutions, and uses a small convolution filter to map on the feature map to help the detector identify small targets, and has high accuracy and rapidity of target detection. To sum up, SSD is selected as the model backbone network to identify underwater lobster. Fig 6 shows the SSD network model structure. It can be seen from the figure that SSD is based on VGG16 template model. The network is divided into six parts, each part completes candidate box generation, feature extraction and multi-scale information fusion respectively, and finally completes the target detection task. [21]

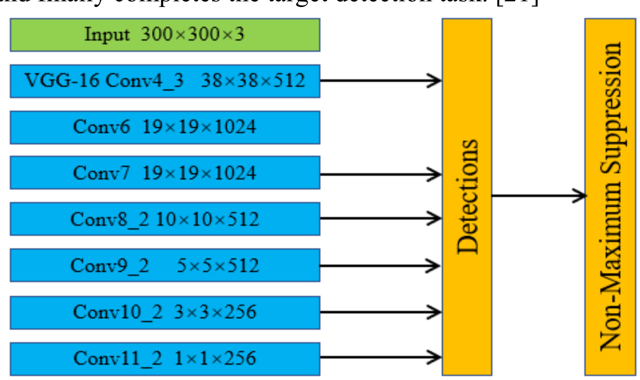

Fig.6 The network architecture of the SSD object detector.

D. MobileNet-v2 Network

Deep convolution neural network has the disadvantages of complex structure, too many parameters and too long calculation time, which makes it difficult to be applied in mobile terminal. Therefore, if you want to apply the depth network model to real scenes, such as embedded mobile robots, you need to choose a small and fast depth network model. MobileNet introduces two super parameters to reduce the amount of parameters and calculation. The basic model unit is depthwise separable revolution, which essentially separates the convolution calculation operation into two smaller convolution operations: depth convolution and point-by-point convolution. Firstly, the depth separation convolution is used to convolute multiple different input channels respectively, and then the point-to-point convolution is used to combine the outputs of each channel to complete the mapping of the input image feature map. Compared with standard convolution, deep separable convolution has less computation and model parameters.

E. Improved MobileNet-SSD Algorithm

We take SSD as the main backbone network of MobileNet SSD algorithm, replace the front-end network of the original SSD network with the MobileNet network framework to realize target detection, and adopt the deep separable convolution of MobileNet from conv1 to conv13. The combination of MobileNet algorithm and SSD algorithm can greatly increase the recognition speed of the algorithm and reduce the size of the model. However, it is still unable to achieve the ideal real-time detection goal. Therefore, we propose to add BN layer and convolution layer to MobileNet SSD algorithm, which can avoid the problems of gradient explosion and gradient disappearance, so as to improve the performance of the model. Fig 7 shows the network structure of MobileNet SSD algorithm.[22]

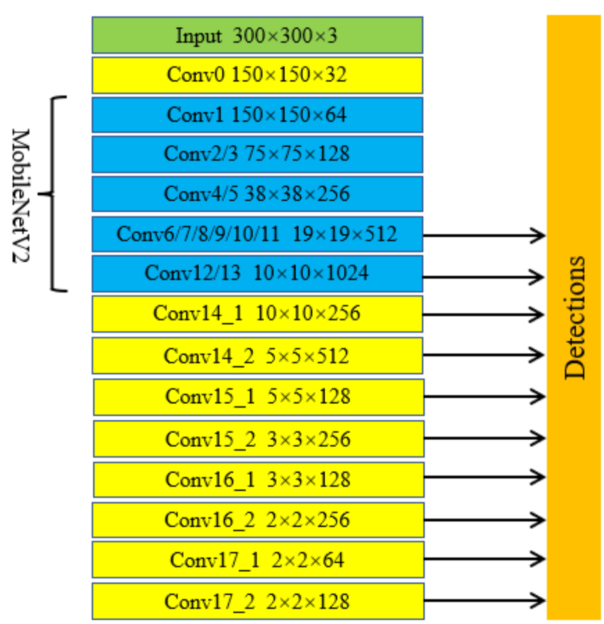

Fig.7 MobileNet-SSD network structure.

The blue convolutional layer in the network framework is the depthwise separable convolution of MobileNet from conv1 to conv13. MobileNet-SSD removes the last global average pooling, fully connected layer and Softmax layer of the MobileNet network, and retains the SSD post-network structure.

IV. EXPERIMENTAL TEST AND RESULT ANALYSIS

A. Experimental Configuration

Our experimental dataset is divided into three parts in a ratio of 7:2:1: training set, test set and validation set. The GPU used for network model training uses Tesla V100 16GB, the controller of the spherical robot uses Raspberry Pi 3 Model B V1.2, and the shooting uses Raspberry Pi camera Rev 1.3. The image size of the device input to the network is 300×300. The initial learning rate is 0.001 and is set to 0.5 momentum every 1/10 decrease every 3 epochs during training.

In this experiment, we carried out two groups of experiments on underwater shrimp recognition with SSD and MobileNet-SSD algorithms. (1) Verify the underwater image perception quality enhancement experiment. The original SSD is used as the recognition algorithm to verify the recognition accuracy, model size and target inference time (MS) before and after data enhancement. (2) The efficiency of underwater shrimp recognition is compared using MobileNet-SSD and SSD algorithm.

In experiment (1), we use SSD algorithm to evaluate the target recognition accuracy, model size and reasoning speed in the same GPU before and after data set enhancement. In experiment (2), on the premise of data set enhancement, in order to prove the performance of the improved algorithm, verify the recognition accuracy of MobileNet SSD and SSD algorithm, FPS, calculation time on the same GPU and CPU and model size.

B. Experimental Results and Analysis

Table I shows the recognition accuracy, model size, and reasoning speed at the same GPU before and after data set enhancement. Then it is verified that the data set image enhancement can greatly improve the recognition accuracy and recognition speed of the model. The recognition accuracy of the data set image enhanced is 4.33% higher than that of the non enhanced image, and the average speed of recognizing the target in the same image is increased by 15344ms, but it does not contribute to reducing the size of the model.

TABLE I
RECOGNITION EFFICIENCY BEFORE AND AFTER DATA SET ENHANCEMENT

	SSD300	
No image enhancement	✔	
FUnIE GAN		✔
Data test set mAP	61.42	65.75

Fig 8 shows the demonstration of recognition effect before data enhancement based on SSD algorithm, after data enhancement based on SSD algorithm and after data enhancement using MobileNet-SSD in the verification set. In the identification demonstration of the verification set part, we can intuitively find the data set, enhance and improve the SSD algorithm, and greatly improve the identification effect of MobileNet SSD algorithm.

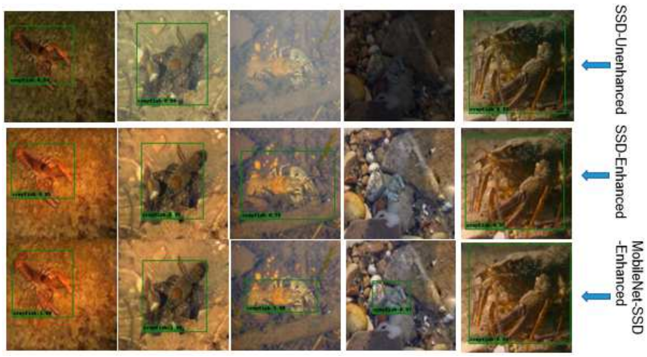

Fig.8 Comparison of data enhanced shrimp identification

Fig 9 shows the changes of Loss function in the training process of each model on self-made shrimp data set, where the vertical axis represents Loss value and the horizontal axis represents iteration times. As can be seen from the figure, the Loss value of each model tends to be stable with the increase of iteration times. Obviously, when the iterations of VGG16-SSD-Unenhanced and VGG16-SSD-Enhanced models reach 50000 times, It can be seen that the mean value of Loss function of VGG16-SSD-Enhanced network model after image enhancement is lower than that of VGG16-SSD-Unenhanced network model. After that, both Loss function values tend to irregular oscillation and obvious fluctuation. Therefore, the experiment shows that the data set image enhancement can better train the depth model.

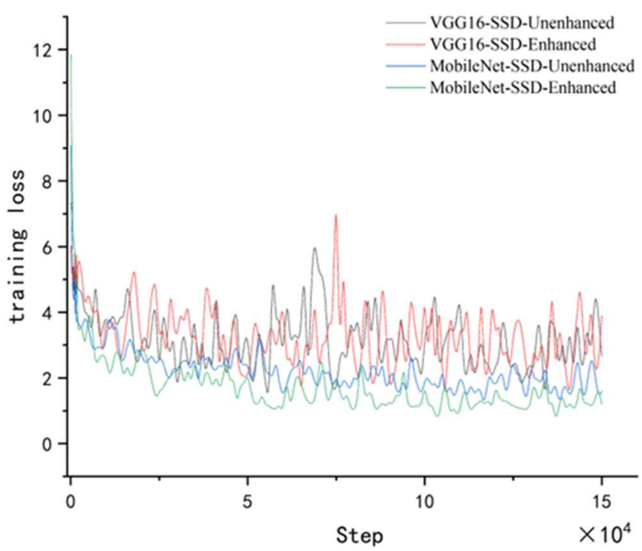

Fig.9 Comparison of data enhanced shrimp identification.

After verifying the image enhancement effect of shrimp data sets, in order to further verify the difference between each network model, we used Mobilenet-SSD to conduct iterative training on shrimp data sets without image enhancement, and compared the Loss function value with VGG16-SSD-Unenhanced. It can be seen that the convergence rate of Loss function value of Mobilenet-SSD is faster and more stable than that of VGG16-SSD-Unenhanced, and the final value tends to be about 1.5. This experiment shows that MobileNet-SSD model can train depth model well.

It is obvious that the convergence rate of Loss value is faster and more stable than that of Mobilenet-SSD-Unenhanced when using the mobilenet-SSD model to conduct in-depth model training on shrimp data set. From the beginning to the end of the model training iteration, the Mobilenet-SSD-Enhanced model maintained better convergence speed and stability, and had a lower and stable Loss function value compared with the VGG16-SSD-Enhance model. It shows that the image enhancement of shrimp data set and Mobilenet-SSD model can better train depth model.

In the shrimp dataset, we tested each model after training on the test set, and got the test results in Table 2. We mainly test the recognition accuracy of the test set, the inference time of the model deployed on the spherical robot, and the size of the model.

Table II shows the comparison of MobileNet SSD and SSD algorithms on the basis of data augmentation in terms of recognition accuracy, inference time, and model size. The inference time is the average time, in seconds, from grabbing the image to the spherical robot inferring the result. Obviously, the recognition accuracy of the MobileNet SSD algorithm on the shrimp dataset is 24.57% higher than that of the SSD algorithm. The number of images that can be processed per second is significantly increased, and the model size is reduced by 157161 KB, which can effectively reduce the running memory of embedded mobile terminals. The recognition speed on the spherical robot is also significantly improved, and the recognition of the same target in the same image is also significantly improved.

TABLE II
COMPARISON OF RECOGNITION EFFICIENCY BETWEEN MOBILENET-SSD AND SSD ALGORITHMS

Method	Input	Precision	Inference	Size
VGG16-SSD	300×300	65.75	269.508	181210
MbileNet-SSD	300×300	90.32	19.568	24049

VI. CONCLUSIONS AND FUTURE WORK

According to the recognition requirements of underwater shrimp detector, an underwater shrimp recognition and detection method based on generative countermeasure network and convolutional neural network is proposed in this paper. Make underwater shrimp data set. Underwater video real-time recognition shrimp detector. Firstly, the underwater shrimp image data set is established, and the image of the data set is enhanced by FUnIE GAN algorithm, and the number of images is expanded. Secondly, the single-stage multi frame detector (SSD) is used as the main framework and MobileNetv2 is selected as the backbone of the algorithm, so as to replace the standard convolution with deep separable convolution and improve the detection accuracy of multi-scale shrimp. The experimental results show that the model improves the performance of underwater shrimp recognition, and the recognition accuracy of the data set reaches 90.32%. It has high recognition accuracy and stability, and can implement and stably complete the task of underwater shrimp recognition.

ACKNOWLEDGMENTS

This research is supported by National Natural Science Foundation of China (61703305), Key Research Program of the Natural Science Foundation of Tianjin (18JCZDJC38500) and Innovative Cooperation Project of Tianjin Scientific and Technological (18PTZWHZ00090).

REFERENCE

[1] Yang X, Zhang S, Liu J , Gao Q , Dong S , Zhou C, “Deep learning for smart fish farming: applications, opportunities and challenges”, *Reviews in Aquaculture*, No.13,2021.

[2] Antonucci F, Costa C, “Precision aquaculture: A short review on engineering innovations”, *Aquaculture International*, No：28, pp:41–57,2020.

[3] Holdich, David M, “A review of astaciculture: freshwater crayfish farming”, *Aquatic Living Resources*, Vol.6,No.4, pp: 307-317,1993.

[4] X Yang, Zhang S, Liu J , Gao Q , Dong S , Zhou C, “Deep learning for smart fish farming: applications, opportunities and challenges”, *Reviews in Aquaculture*, pp: 14-19,2020.

[5] Jesse E, Dylan K, Janine B , Scott M , Dan Z, “Advancements towards selective barrier passage by automatic species identification: applications of deep convolutional neural networks on images of dewatered fish”, *ICES Journal of Marine Science*, Vol.12,No.8,pp:7-8,2020.

[6] Fernandes A , Turra E. M Alvarenga, R Passafaro, T. L , Rosa G , “Deep learning image segmentation for extraction of fish body measurements and prediction of body weight and carcass traits in nile tilapia”, *Computers and Electronics in Agriculture*, No.170 pp:105274-,2020.

[7] Spanig S , Emberger-Klein A, Sowa J.P, Can bay A, Men rad K., Heider D, “The virtual doctor: An interactive clinical-decision-support system based on deep learning for non-invasive prediction of diabetes”, *Artificial Intelligence in Medicine*, Vol.100,pp: 101706, 2019.

[8] Lu Z , Jw C , Qda B, “Estimation for fish mass using image analysis and neural network”, *Computers and Electronics in Agriculture*, Vol.173,2020.

[9] Chao Zhou, Kai Lin, Daming Xu, Lan Chen, Qiang Guo, Chuan heng Sun, Xinting Yang, “Near infrared computer vision and neuro-fuzzy model-based feeding decision system for fish in aquacul ture, Computers and Electronics in Agriculture”, *Computers & Electronics in Agriculture*, Vol.146,pp:114-124,2018.

[10] Zhou, Chao Sun, Chuan heng Lin, Kai Xu, Daming Guo, Qiang Chen, Lan Yang, Xinting, “Handling Water Reflections for Computer Vision in Aquaculture”, *Transactions of the ASABE*, Vol.10, No.5, pp: 469-479,2018.

[11] Shuoxin Gu, Shuxiang Guo, Liang Zheng, “A highly stable and efficient spherical underwater robot with hybrid propulsion devices”, Autonomous Robots, Vol.44, No.5, pp.759-771, DOI: 10.1007/s10514-019-09895-8, 2020.

[12] Islam M J , Xia Y , Sattar J, “Fast Underwater Image Enhancement for Improved Visual Perception”, *IEEE Robotics and Automation Letters*, Vol.5, No.2,pp: 3227-3234, 2020.

[13] Karras T, Laine S, Aittala M, Hellsten J, Lehtinen J, Aila T, “Analyzing and Improving the Image Quality of StyleGAN”, *IEEEXplore*, 2019.

[14] Xihuan Hou, Shuxiang Guo, Liwei Shi, Huiming Xing, He Yin, Zan Li, Mugen Zhou, Debin Xia, “Improved Model Predictive-Based Underwater Trajectory Tracking Control for the Biomimetic Spherical Robot under Constraints”, Applied Science, Vol.10, No.22, 2020.

[15] Ruochen An, Shuxiang Guo, Yuanhua Yu, Chunying Li, Tendeng Awa, “Multiple Bio-Inspired Father-Son Underwater Robot for Underwater Target Object Acquisition and Identification”, Micromachines 2022, Vol.13, No.1,2021.

[16] Shuxiang Guo, Sheng Cao, Jian Guo, “Study on Decentralization of Spherical Amphibious Multi-Robot Control System Based on Smart Contract and Blockchain”, Journal of Bionic Engineering, 2021.

[17] Jian Guo, Chunying Li, Shuxiang Guo, “Path Optimization Method for the Spherical Underwater Robot in Unknown Environment”, Journal of Bionic Engineering, Vol.17, No.5, pp.944-958, 2020.

[18] Ruochen An, Shuxiang Guo, Chunying Li, Tendeng Awa, “Underwater Motion Control of a Bio-inspired Father-son Robot Based on Hydrodynamics Analysis”, International Journal of Mechatronics and Automation, in press, 2021.

[19] Islam M J , Xia Y , Sattar J, “Fast Underwater Image Enhancement for Improved Visual Perception”,*IEEE Robotics and Automation Letters*, pp: 3227-3234, 2020.

[20] Liu W, Anguelov D , Erhan D , Szegedy C , Reed S , Fu C. Y, “SSD: Single Shot MultiBox Detector”, *Springer, Cham,2016.*

[21] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. C, “MobileNetv2: inverted residuals and linear bottlenecks” , *IEEE* , 2018.

[22] Cob-Parro AC, Losada-Gutiérrez C, Marrón-Romera M, Gardel-Vicente A, Bravo-Muñoz I. “Smart Video Surveillance System Based on Edge Computing. Sensors”, Vol.21, No.9, pp: 2958,2021.