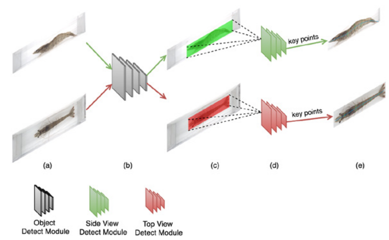
Study on deep learning-based detection method of key points of unconstrained Pacific white shrimp in water



Estudio de un método de detección basado en deep learning de puntos clave de mediciones no restringidas del camarón blanco del Pacífico en el agua

Xiujun Zhang[1], **Su Fang**[2], **Zechao Jin**[3], **Sheng Luan**[4,*]

[1] School of Applied Engineering, Zhejiang Business College – 310053 Hangzhou (China)
[2] Xiaomi Communications Co. Ltd - 100085 Beijing (China)
[3] School of Information and Electrical Engineering. Hangzhou City University – 310015 Hangzhou (China)
[4] State Key Laboratory of Mariculture Biobreeding and Sustainable Goods, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao – 266071 Shandong (China).

DOI: https://doi.org/10.52152/D11350 | Received: 03/oct/2024 • Reviewing: 03/oct/2024 • Accepted: 03/dec/2024

To cite this article: ZHANG, Xiujun; FANG, Su; JIN, Zechao; LUAN, Zechao. STUDY ON DEEP LEARNING-BASED DETECTION METHOD OF KEY POINTS OF UNCONSTRAINED PACIFIC WHITE SHRIMP IN WATER. *DYNA* January February 2025. Vol. 100, n.1, pp. 83-89 | DOI: https://dx.doi.org/10.6036/D11350

FUNDING

This research was funded by the Zhejiang Public Welfare Technology Application Research Project of China (LGN21C190002), China Agriculture Research System of MOF and MARA (No. CARS-48), Central Public-interest Scientific Institution Basal Research Fund, CAFS (No. 2020TD26).

RESUMEN

- En la industria acuícola, las plántulas acuáticas se conocen comúnmente como «chips». Los datos fenotípicos de las plántulas de camarón, que pueden reflejar su crecimiento, son importantes índices de referencia para la cría. En el cultivo tradicional del camarón, los puntos clave del cuerpo del camarón se determinan principalmente por medios artificiales, y los parámetros se miden manualmente para obtener datos fenotípicos relacionados con la cría. Sin embargo, este método manual no sólo requiere mucho tiempo y trabajo, sino que también puede dar lugar a errores humanos. Además, las gambas son muy sensibles a la manipulación, lo que puede causarles daños físicos, propagar enfermedades y contaminar el agua durante las mediciones manuales. Para mejorar la velocidad y la precisión de la recopilación de datos fenotípicos de camarones, este estudio propuso un enfoque novedoso: una red basada en el aprendizaje profundo para la detección automática de puntos clave de camarones. El método propuesto minimizó el contacto físico, evitó posibles daños y permitió la recopilación de datos fenotípicos más completos. Con el camarón blanco del Pacífico como objeto de estudio, se aplicó el aprendizaje profundo para detectar los puntos clave del cuerpo del camarón. La red de detección de puntos clave pudo detectar 23 puntos clave de la vista superior y 10 puntos clave de la vista lateral, lo que podría proporcionar datos de apoyo para mediciones posteriores de parámetros fenotípicos y para modelar el crecimiento del cuerpo del camarón. Los resultados muestran que, en comparación con los métodos tradicionales, el enfoque propuesto tiene un rendimiento avanzado, velocidad y robustez en términos de precisión y eficiencia. La precisión de los puntos clave de la vista superior alcanza el 97,57%, con un tiempo medio de prueba de 137,3 ms, y la verificación de la vista lateral establece una precisión de detección de puntos clave del 98,61%, con un tiempo medio de prueba de 49,3 ms. Este estudio realiza la medición sin restricciones de camarones con agua, que puede obtener con precisión y rapidez los puntos clave de los camarones y satisfacer las necesidades de medición de múltiples datos fenotípicos.

- **Palabras clave:** datos fenotípicos, aprendizaje profundo, camarón blanco del Pacífico, detección de puntos clave, medición sin restricciones.

ABSTRACT

In the aquaculture industry, aquatic seedlings are commonly known as "chips." The phenotypic data of shrimp seedlings, which can reflect their growth, are important reference indexes for breeding purposes. In traditional shrimp culture, key points on the shrimp body are mainly determined through artificial means, and parameters are manually measured to obtain breeding-related phenotypic data. However, this manual approach is not only time-consuming and laborious but can also lead to human errors. Moreover, shrimp are highly sensitive to handling, which can easily cause physical harm, spread diseases, and lead to water contamination during manual measurements. To improve the speed and accuracy of shrimp phenotypic data collection, this study proposed a novel approach: a deep learning-based network for the automatic detection of shrimp key points. The proposed method minimized physical contact, prevented potential damage, and enabled the collection of more comprehensive phenotypic data. With Pacific white shrimp as the study object, deep learning was applied to detect key points of the shrimp body. The key point detection network could detect 23 key points of the top view and 10 key points of the side view, which could provide data support for subsequent measurements of phenotypic parameters and for modeling shrimp body growth. Results show that, compared with traditional methods, the proposed approach has advanced performance, speed, and robustness in terms of accuracy and efficiency. The accuracy of key points of the top view reaches 97.57%, with an average test time of 137.3 ms, and the side view verification set a key point detection accuracy of 98.61%, with an average test time of 49.3 ms. This study realizes the unconstrained measurement of shrimp with water, which can accurately and quickly obtain the key points of shrimp and meet the needs of multiple phenotypic data measurement.

Keywords: phenotypic data, deep learning, Pacific white shrimp, key point detection, unconstrained measurement.

1. INTRODUCTION

Phenotypic data of aquatic animals are the main observation data during the breeding of seedlings. However, the long-term manual measurement of phenotypic data cannot meet the speed and precise measurements required in the breeding process. Researchers have been exploring the use of traditional image processing algorithms to complete the measurement of phenotypic data through object detection and segmentation, but strict constraints were required [1-2]. To solve the generalization performance of the algorithms, deep learning algorithm was used to complete the phenotypic measurement of aquatic animals [3-4]. However, few phenotypic data indicators were obtained, which cannot meet the requirements of breeding data quantity, and the measured phenotypic data quantity was insufficient.

The accuracy of key point detection is important for measuring the phenotypes of aquatic animals. Most existing deep learning methods for key point detection utilize thermal map and regression techniques. While these methods significantly improve accuracy when high-quality images are available, they require reliable image quality as a prerequisite. Aquatic animals, however, often exhibit large deformations and high swimming speeds, making it challenging to capture high-quality images in real-time. Factors such as impurities and varying light conditions in the water can further hinder the acquisition of suitable image samples. In addition, rapid and accurate measurement in batches is needed to breed and raise seedlings, yet there is currently no relevant literature on this topic [5-6].

Therefore, with Pacific white shrimp as the study object, this study established an unconstrained aquatic animal image acquisition device that can obtain both the top and side views of white shrimp. Considering the large number of key points of the top view, an improved hourglass network was designed to improve the accuracy. There were few key points of the side view, so a lightweight design was adopted to improve the speed of key point identification. Finally, a unified deep learning model for key point detection was established, and key point identification of the top view and a test graph were completed, which ensured the accuracy and speed of identification.

2. STATE OF THE ART

The development of computer vision technology has enabled researchers to explore new methods for measuring the phenotypic data of aquatic animals. Monkman [7] proposed the use of R-CNN to detect the predicted position and the region of the color plate. By taking the length of the maximum detection frame as the predicted width of fish, the measurement method was limited by the detection accuracy of the boundary frame. Tseng [8] used a convolutional neural network to detect the position of fish head, fish tail, and color plate. By carefully designing each image processing unit module, the actual length of the fish can be accurately calculated. The algorithm has high complexity due to the mutual influence between various image modules and the large number of parameters in each module, and the locations of the fish head and tail points were easily affected by environmental interference. Although the measurement of underwater fish body size was not realized, the difficulty of realization was discussed. Lai[9] adopted YOLO4 to complete target detection for shrimp body images obtained underwater and used the image segmentation algorithm to obtain the length and width of shrimp body. Setiawan [10] adopted traditional image processing algorithms to obtain shrimp body morphology indicators through edge detection and region of interest processing, as well as constantly optimized the parameters of the algorithm according to image quality. Bao [11] applied a two-stage image segmentation algorithm to obtain the length of shrimp. However, measurements out of water and segmentation algorithm were not carried out in real-time, so it was difficult to apply to the breeding process requiring rapid measurement. Saleh [12] used the key point detection method to measure morphological and phenotypic data of shrimp body, but the method was mainly used for weight prediction and disregarded the accuracy of phenotypic data such as body length. Meanwhile, there was no measurement data of the top view of shrimp body. The above studies used target detection and image segmentation technology to obtain the outline of images and, subsequently, the length and width of aquatic animals. However, such methods could only measure a small number of phenotypic data indicators of aquatic animals.

The phenotypic data of aquatic animals are mainly measured according to the key points on their bodies. Thus, the accuracy of key point detection affects the accuracy of the final phenotypic data measurement. DeepPose proposed by Toshev [13] introduced deep learning into human key point detection for the first time. The cascade network composed of deep neural network regressors was constructed to detect key point coordinates of the human body and obtain accurate human posture. The model can learn input spatial features but cannot model the dependency relationship explicitly in output space. To further improve the recognition accuracy of key points, Park [14] used depth information to improve the confidence of key point location and solve the problem of key point recognition with an obscured human body. Moccia [15] improved the accuracy of pose detection through spatiotemporal characteristics. The error value of the key point position of the algorithm does not affect the pose estimation, but it cannot meet the application in the measurement field. On the one hand, the detection algorithm based on key points pursues the accuracy of recognition. On the other hand, it constantly explores the lightweight quality of the model. Wang [16] used an improved Kalman filter to compensate for errors in fast motion. The aim was to enhance the detection accuracy of gestures and avoid gesture loss in complex backgrounds, thereby improving the recognition rate of key points of the hand and reducing the number of parameters in the YOLO3 ratio. However, this method is still a two-stage operation, which is not conducive to improving subsequent recognition speed. Groos [17] proposed Efficient Pose, a new convolutional neural network architecture that can quickly generate key point heat maps by limiting memory usage and computational cost to support the deployment of key point detection programs on edge devices. Xu [18] proposed the ViTPose model of key point detection in the latest transformer framework. Although the structure has not been refined, it can obtain relatively good detection accuracy on various data sets and has relatively good generalization performance. ViTPose allows the flexible adjustment of attention mechanism types, input resolution, pre-training, and finetune (fine-tuning) strategies, empirically demonstrating that knowledge of large models can be transferred to small models by setting simple knowledge tokens. Considering the running speed of the algorithm, Huang [19] embedded a coordinate transformer in the target detection model to realize the lightweight model of multi-person attitude detection. However, missing detection still occurs.

The above target detection algorithms are mostly used in the field of pose detection. Although Yao [20] used the key-point detection algorithm for estimating the body size of pigs, the centimeter-level error produced cannot meet the accuracy requirements for small-sized aquatic animals. The deviation of key point detection has minimal impact on pose estimation, but it introduces relatively large measurement errors. In order to improve the accuracy of key point detection, Newell [21] first proposed the Hourglass net network structure, which achieved accurate detection of human skeletal key points. Zhou [22] proposed a Graphical Model based Structure Context Enhancement Network (GM-SCENet) based on an hourglass network. The proposed network enhances multi-scale information and improves the robustness of key point location through a Structure Context Mixer (SCM) and a Cascade Multi-Level Supervision (CMLS) module. The deep network model based on the hourglass architecture has a concise and easily extensible structure, showing accurate and efficient performance. To improve the feature extraction capabilities of deep learning networks and create a lightweight key point recognition model, Sun[23] proposed a high resolution network (HRNet) that integrates multi-scale features, reducing the complexity of the network, and Zhang[24] proposed a high-resolution feature learning network that can detect key points faster.

The above studies focused on key point detection method based on deep learning. In deep learning, it is difficult for the coordinate regression method to accurately extract the dependency information between key points and structures due to its simple models and poor ability to learn the dependency relationship between structures, resulting in low algorithm accuracy. The heat map contains the probability information of key points in space, and it is more suitable for the whole key point detection. The key point algorithm based on heat maps generates heat maps by calculating

Serial number	Point name	Position point specification
Key points of the top view		
1	top-head	Edge of the shrimp head
2–3	eye-up (down)	Marginal point of the upper (lower) eye socket
4–5	bw-up (down)	Marginal point on the upper (lower) edge of the widest part of the shrimp head
6	mid	Midpoint of the first shrimp curve
7–20	bn-up (down)	Upper (lower) edge point where the n node intersects with the next node
21	tail-up	Upper edge point of the shrimp tail
22	top-tail	Marginal point of the central axis of the shrimp body away from the head
23	tail-down	Lower edge point of the shrimp tail
Key points of the side view		
1	side-head	Edge of the shrimp head
2–3	b1-up (down)	Upper (lower) edge where the head and the body intersect
4–5	b3-up (down)	Upper (lower) edge point where the third node of the shrimp intersects with the next node
6–7	b5-up (down)	Upper (lower) edge point where the fifth node of the shrimp intersects with the next node
8–9	b7-up (down)	Upper (lower) edge point where the seventh node of the shrimp intersects with the next node
10	side-tail	Edge point of the shrimp tail

Table 1. Description of key points of the top view and side view of Pacific white shrimp.

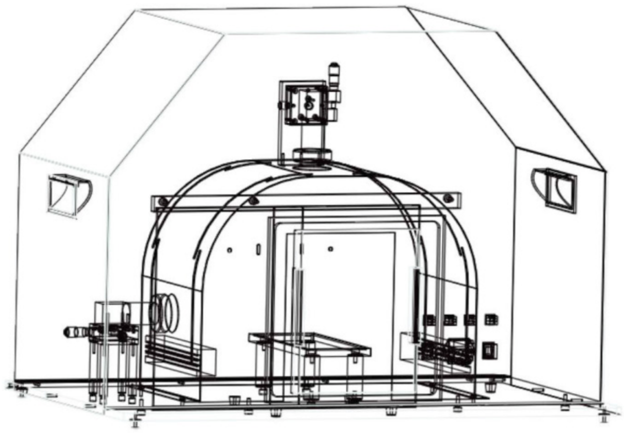


Fig. 1. Hardware structure of the device.

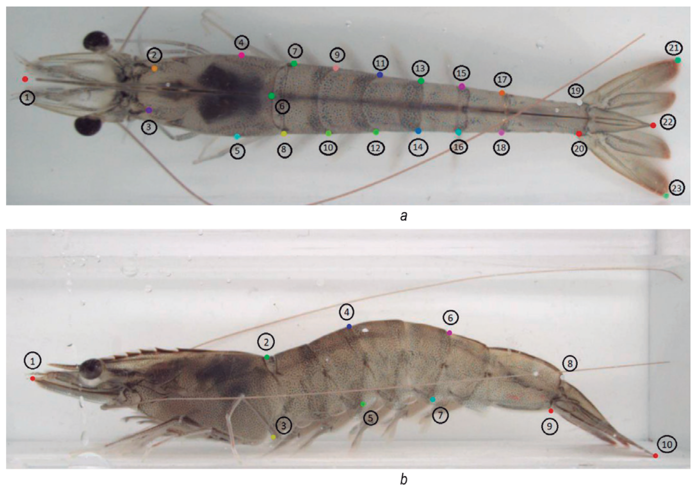


Fig. 2. Shrimp key points: (a) Key points of the top view (b) Key points of the side view.

the probability that each pixel in the image is critical, rather than directly regressing the coordinate information of the key points. This enables the model to provide the information of key points and pixels around the key points, thus improving the accuracy and generalization ability of key point detection. The studies discuss the accuracy of key point identification in input high-quality image samples, which cannot be applied to the rapid and accurate measurement of phenotypic data in batches during the breeding process of Pacific white shrimp. During population breeding, the phenotypic data of a large number of shrimp seedlings must be collected quickly and accurately. This entails not only the algorithm to meet the requirements but also the equipment to reliably collect unconstrained shrimp body image samples in water. In this study, an image acquisition device was established that can quickly obtain the top and side views of Pacific white shrimp. A deep learning network combined with hourglass network and HRNet structure features was designed, which can detect the key points of the top and side views of Pacific white shrimp in quasi-real time, meeting the requirements of accuracy and speed. The identification of key points of the top view can be improved by inserting a relay module in the hourglass network. The side view can optimize the HRNet network structure to improve the identification speed.

The remainder of this study is structured as follows. Section 3 describes the image acquisition device for the Pacific white shrimp and constructs a deep learning network model that can detect the key points of the top and side views. Section 4 completes the detection of the top view and key points based on the model, rea-

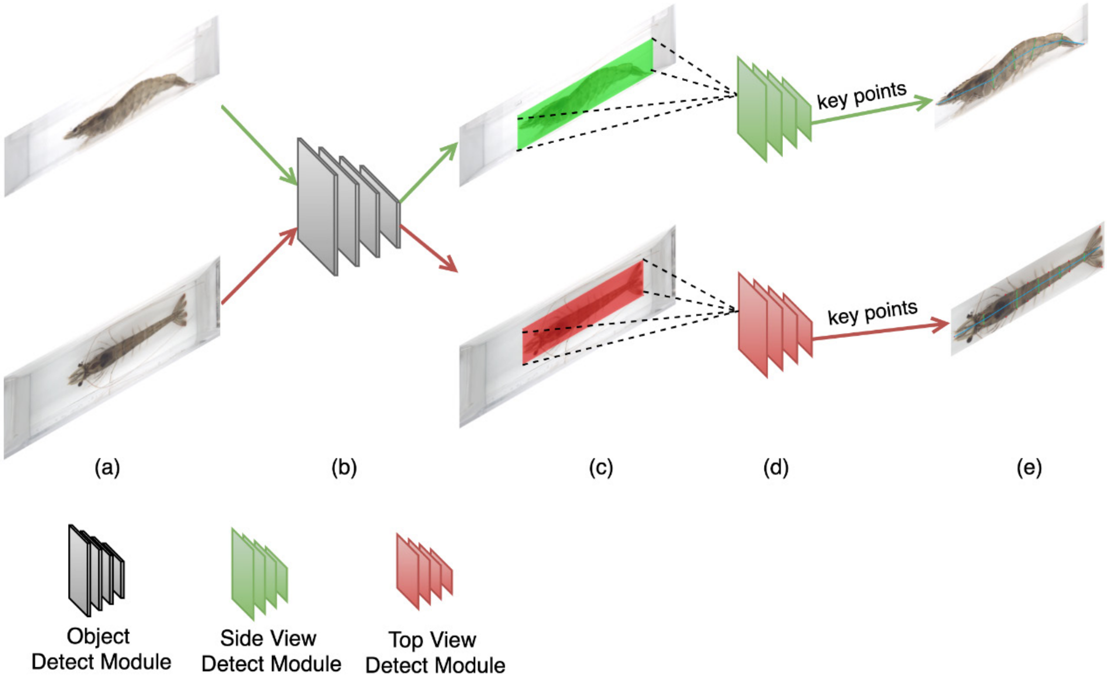


Fig. 3. Overall network framework: (a) image reading, (b) target detection module, (c) accurate target positioning after target detection, (d) key point detection module, and (e) final output results, including key points of shrimp and shrimp motion posture.

lizes the index evaluation of different models. Finally, Section 5 summarizes this study and draws relevant conclusions.

3. MATERIAL & METHODS

3.1. HARDWARE DEVICE STRUCTURE

The hardware used in this study includes four strip light sources on the inside surface of the device, a slide track inside the device, and two 4 MP cameras positioned in the overhead and side-view directions. A shrimp box was used, with one side of the box transparent and the other side white. When detecting key points, 3/4 of the shrimp box was filled with water, and the shrimp box was placed on the slide and pushed to collect the top view and side view of the shrimp. The device structure is shown in Fig. 1. Note that when the shrimp box is pushed into the device, if the unconstrained shrimp do not swim quickly in water, the camera can be turned on to take both top and side views of the shrimp, providing image data required for the algorithm model.

3.2. IMAGE ACQUISITION AND DATA PREPARATION

Shrimp do not always maintain the same posture and angle in water. Therefore, to obtain a more robust model and adapt it to the different postures and body types of Pacific white shrimp, this study collected 3,000 top and side views of shrimp with different body types and postures using the hardware in the previous section. Next, 2,500 of the views were used to train highly accurate and efficient key-point detection models, and the remaining views were used to evaluate the performance of the trained models.

Meanwhile, to solve the difficulties and problems existing in the measurement of phenotypic parameters or pose estimation by shrimp breeders, this study defined 23 key points of the top view and 10 of the side view based on the needs of the phenotypic parameters and pose estimation, as shown in Fig. 2. The label description for each point is shown in Table 1.

3.3. ALGORITHM MODEL STRUCTURE

In this study, a new shrimp detection algorithm was designed based on the characteristics of top and side views of Pacific white shrimp in image acquisition hardware, considering the balance bet-

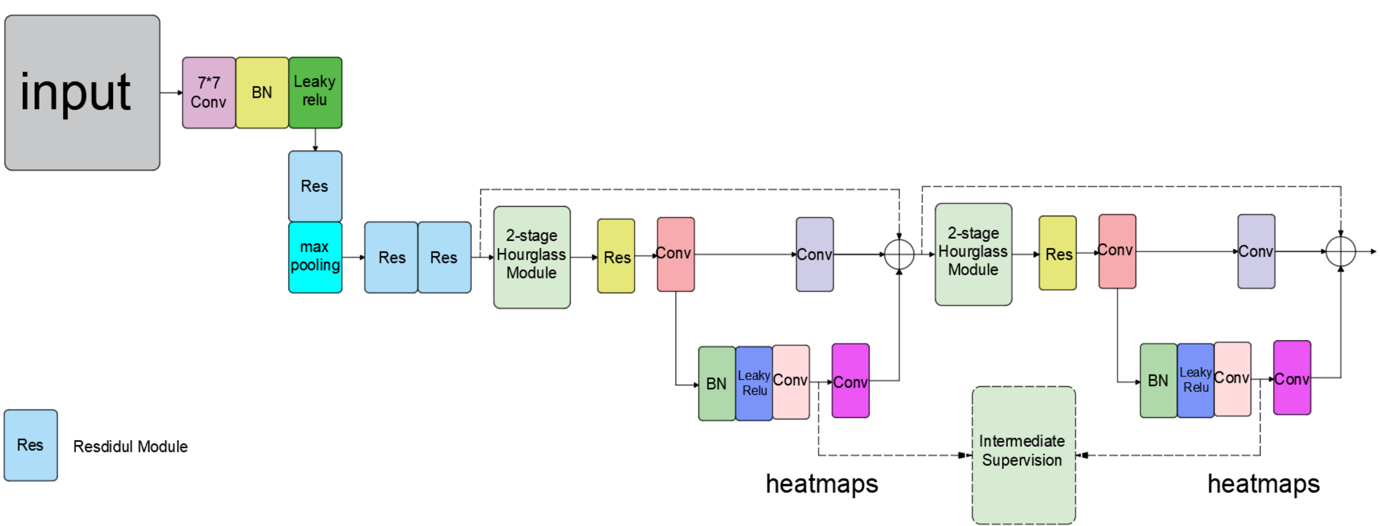


Fig. 4. Overall network architecture of detection module in top view.

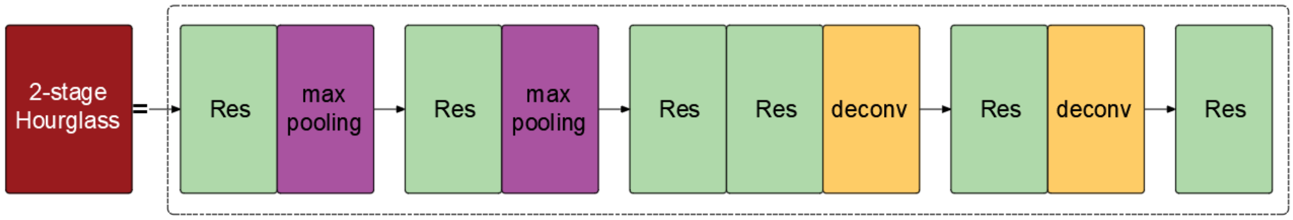


Fig. 5. Improved hourglass module.

ween operation accuracy and operation efficiency. The overall model structure (Fig. 3) is divided into a target detection sub-module, a top view detection sub-module, and a side view detection sub-module.

3.4. TOP VIEW DETECTION MODULE

As shown in Fig. 4, the top view detection module consists of two cascaded 2-stage hourglass modules and relay supervision modules. In the hourglass module [21], the input image is first downsampled through a max pooling layer to reduce its area to 1/4 of the original image, and then upsampled using bilinear interpolation to restore its original size. At the same time, the hourglass module preserves the original information through residual addition, allowing the hourglass module to learn features at different scales and preserve the original features as much as possible. However, using bilinear interpolation for upsampling simply restores the image to its original size, while losing some of the data lost during the downsampling process. Therefore, this paper uses deconvolution for upsampling to reduce the information lost during the upsampling process. Fig. 5 illustrates that 2-stage improved hourglass.

In addition, in the network model shown in Figure 4, Leaky ReLU is selected as the activation function. Meanwhile, to stack multiple hourglass modules in the entire network architecture and prevent gradient disappearance caused by network depth, the network can constantly repeat the bottom-up and top-down feature extraction to improve network reliability. One characteristic of Hourglass networks is Intermediate Supervision. It adds a supervised part in each stage of the network, so that the output heatmap of each order calculates the loss using the ground truth and then calculates the gradient for backpropagation, that is, the training loss after each hourglass module was calculated to help the training of the hourglass module in the later stage.

3.5. SIDE VIEW DETECTION MODULE

Different from the top view detection module, the structure of key points of the side view is simple, and the number of key points is greatly reduced compared with those of the top view. To further accelerate the overall efficiency of the algorithm, a lighter network structure was used in the side view detection module [23], as shown in Fig. 6(see section: supplementary material). The side view detection module replaces the intermediate supervision and hourglass modules in the top view detection module by connecting the subnet from high resolution to low resolution in parallel. While greatly reducing the model parameters, the multi-scale repeated fusion was carried out by repeatedly exchanging information on the parallel multi-resolution subnetwork to maintain the precision of key point detection. The paper did not make any structural modifications to the model proposed in reference [23], but only trained the network model based on the number of key points in the shrimp side view.

4. RESULTS

4.1. KEY POINT DETECTION MODEL TRAINING AND PREDICTION

To train a robust and high-precision key point detection model, the 2,500 training sets prepared in Section 3.2 were used for

Model \ CK threshold	top view						side view					
	0.04	0.06	0.08	0.1	0.12	0.14	0.04	0.06	0.08	0.1	0.12	0.14
MobileNetV2	0.116	0.229	0.321	0.407	0.475	0.536	0.291	0.377	0.461	0.532	0.595	0.652
ResNet50	0.372	0.586	0.725	0.823	0.883	0.925	0.593	0.64	0.729	0.779	0.811	0.832
ResNet101	0.233	0.4	0.526	0.631	0.705	0.764	0.497	0.536	0.636	0.702	0.749	0.786
ShuffleNetV2	0.085	0.174	0.257	0.339	0.408	0.473	0.318	0.429	0.526	0.601	0.661	0.616
Hourglass Net *	0.715	0.893	0.955	0.978	0.986	0.991	0.669	0.821	0.895	0.928	0.944	0.954
HRNet	0.534	0.747	0.856	0.916	0.947	0.966	0.62	0.771	0.857	0.899	0.924	0.943

*Table 2. PCK comparison of different models in top view and side view: Hourglass Net * in the table was proposed by this paper.*

Model	Average training time s/epoch	Average prediction time ms/per	Model parameter MB (Megabyte)
MobileNetV2	10.0	28.1	33.1
ResNet50	16.5	35.5	272.3
ResNet101	35.2	64.2	424.5
ShuffleNetV2	10.5	31.2	24.7
Hourglass Net*	55.6	137.3	199.2
HRNet	18.6	49.3	32.0

*Table 3. Comparison of efficiency of different models: Hourglass Net * in the table was proposed by this paper.*

training. The resolution of the input image was 512×512, and a 128×128 probability heat map was generated for the corresponding truth value coordinates. A probabilistic heat map was used to represent the coordinates of key points, and the probability value of each pixel on the probabilistic heat map generated was calculated using the multivariate Gaussian model. In particular, the closer the key point was, the closer the corresponding probability value was to 1 (Fig. 7(see section: supplementary material)).

During the training and reasoning process, the training device configured with a 3.70 GHz Intel i7-8700K CPU and an NVIDIA RTX2080 SUPER graphics card was used. During the overall training, the accuracy of the training set and test sets changed, as shown in Fig. 8(see section: supplementary material). The prediction results based on the trained model are shown in Fig. 9(see section: supplementary material). The training speed of the top view detection model is slower than that of the side view, while the accuracy of the final verification set is slightly lower than that of the side view due to the large number of keys and structural changes in the top view. Finally, the accuracy of key point detection in the top view and side view verification sets reached 97.57% and 98.61%, respectively, after training.

4.2. PRECISION COMPARISON EXPERIMENT OF THE KEY POINT DETECTION MODEL

To compare the accuracy of different key point detection networks in shrimp key point detection, six of the latest key point detection networks were selected for comparison: MobileNetV2[25], ResNet50[26], ResNet101, ShuffleNetV2[27], Hourglass Net, and HRNet. The percentage of correct key-points (PCK), referring to the proportion of normalized distance between detected key points and their corresponding true values less than the set threshold, was adopted as the indicator to evaluate the accuracy performance of the model. Its normalized threshold is from 0 to 1. In the experiment, each of the six models was tested on 500 images, and PCK was calculated using different thresholds as an evaluation indicator. The results are shown in Table 2 ~~and 3~~.

Hourglass Net is found to be considerably more accurate than other networks in the detection and comparison of key points of the top view. In the comparison of key point detection of the side view, because the structure of key points here is simpler than that of the top view, the performance gap between models is smaller than that of the top view. Both Hourglass Net and HRNet have good performance in terms of accuracy.

4.3. EFFICIENCY COMPARISON EXPERIMENT OF KEY POINT DETECTION MODELS

To further compare the operating efficiency of different key point detection networks in shrimp key point detection, the six key point detection networks in the previous section were selected for comparison. Model size, training time per batch, and average time of model prediction were taken as the indexes to evaluate model efficiency and performance. In the experiment, the six models were each tested on 1,000 images, and the results are shown in Table 3.

According to the above table, although Hourglass Net achieves the best key point detection accuracy according to the experimental results in Section 4.2, its efficiency is far behind that of other networks. HRNet has high efficiency while maintaining excellent performance in accuracy. Therefore, based on the above experimental results and the actual application scenario of shrimp key point detection, Hourglass Net was selected in this study as the top view key point detection network and HRNet as the side view key point detection network.

5. DISCUSSION

In light of the pain points of traditional methods for measuring shrimp phenotypic data and the shortcomings of existing computer vision approaches, this study took Pacific white shrimp as the study object and applied a deep learning algorithm to the field of shrimp key point detection for the first time. A novel shrimp key point detection method based on deep learning was proposed, which detected a total of 33 key points from both top and side views simultaneously. The following conclusions could be drawn:

(1) High detection accuracy was achieved, with the key point detection accuracy for the top and side views reaching 97.57% and 98.61%, respectively.

(2) Compared to other advanced algorithms, the proposed method balances detection accuracy and efficiency. It took an average of 49.3 ms to detect a key point of the side view and an average of 137 ms to detect a key point of the top view, saving substantial manpower and material resources. Moreover, the measured phenotypic data were comprehensive.

(3) The method enabled the non-constrained measurement of shrimp in water. Compared to measurements taken out of water, the shrimp were completely in their natural state, which reduced their stress response and improved the level of aquaculture outcomes.

Although this study focused on Pacific white shrimp, the proposed key point detection algorithm is applicable to the key point detection and phenotypic data measurement of other shrimp species.

REFERENCES

[1] Hsieh, C. L., Chang, H. Y., Chen, F. H., et al., "A simple and effective digital imaging approach for tuna fish length measurement compatible with fishing operations". Computers and Electronics in Agriculture, 2011. Vol.75-1. p.44–51.DOI: https://doi.org/10.1016/j.compag.2010.09.009

[2] Balaban, M. O., Sengör, G. F. U., Soriano, M. G., et al., "Using image analysis to predict the weight of Alaskan salmon of different species". Journal of Food Science, 2010. Vol.75-3. p.E157–62. DOI: https://doi.org/10.1111/j.1750-3841.2010.01522.x

[3] Deng, Y. X., Tan, H. Q., Tong, M. H., Zhou, D.H., Li, Y.X., Zhu, M., "An automatic recognition method for fish species and length using an underwater stereo vision system". Fishes, 2022, Vol.7-6. p.326. DOI: https://doi.org/10.3390/fishes7060326

[4] Zhou, M.G., Shen, P.F., Zhu, H., Shen, Y., In-water fish body-length measurement system based on stereo vision. Sensors, 2023. Vol.23-14. p.6325. DOI: https://doi.org/10.3390/s23146325

[5] Yang,L., Zhong, J.H., Zhang, Y., et al., "An Improving Faster-RCNN With Multi-Attention ResNet for Small Target Detection in Intelligent Autonomous Transport With 6G". IEEE Transactions on Intelligent Transportation Systems, 2023.Vol. 24-7. p. 7717-7725. DOI:https://doi.org/10.1109/TITS.2022.3193909

[6] Li, G., Yang, L., Bai, S. et al. BIKAGCN: Knowledge-Aware Recommendations Under Bi-layer Graph Convolutional Networks. Neural Processing Letter,2024. Vol. 56-20.p.1-21. DOI:https://doi.org/10.1007/s11063-024-11475-6

[7] Monkman, G. G., Hyder, K., Kaiser, M. J., et al., "Using machine vision to estimate fish length from images using regional convolutional neural networks". Methods in Ecology and Evolution, 2019. Vol.10-12. p.2045–2056. DOI: https://doi.org/10.1111/2041-210X.13282

[8] Tseng, C. H., Hsieh, C. L., Kuo, Y. F., "Automatic measurement of the body length of harvested fish using convolutional neural networks". Biosystems Engineering, 2020. Vol.189. p.36–47. DOI: https://doi.org/10.1016/j.biosystemseng.2019.11.002

[9] Lai, P.C., Lin, H.Y., Lin, J.Y., Hsu, H.C., Chu, Y.N., Liao, C.H., Kuo, Y.F., "Automatic measuring shrimp body length using CNN and an underwater imaging

system". Biosystems Engineering, 2022. Vol.221. p.224–235. DOI: https://doi.org/10.1016/j.biosystemseng.2022.07.006

[10] Setiawan, A., Hadiyanto, H., Widodo, C.E., "Shrimp body weight estimation in aquaculture ponds using morphometric features based on underwater image analysis and machine learning approach". Revue d'Intelligence Artificielle, 2022. Vol.36-6. p.905–912. DOI: https://doi.org/10.18280/ria.360611

[11] Bao, Z.N., Yu, Y., Li, F.H., "The establishment and application of a fast phenotypic determination technique based on Faster R-CNN for growth traits in shrimp". Acta Hydrobiologica Sinica, 2023. Vol.47-10. p.1576–1584 (in Chinese). DOI: https://doi.org/10.7541/2023.2022.0490

[12] Saleh A., Hasan, M.M., Raadsma, H.W., Khatkar M.S., Jerry, D.R., Azghadi, M.R., "Prawn morphometrics and weight estimation from images using deep learning for landmark localization". Aquacultural Engineering, 2024. Vol.106-8. p.102391. DOI: https://doi.org/10.1016/j.aquaeng.2024.102391

[13] Toshev, A., Szegedy. C., "Deeppose: human pose estimation via deep neural networks". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA:IEEE. 2014. p.1653–1660. DOI: https://doi.org/10.1109/CVPR.2014.214

[14] Park, S., Ji, M., Chun. J., "2D human pose estimation based on object detection using RGB-D information". KSII Transactions on Internet & Information Systems, 2018. Vol.12-2. p.800–816. DOI: https://doi.org/10.3837/tiis.2018.02.015

[15] Moccia, S., Migliorelli, L., Carnielli, V., Frontoni, E., "Preterm infants' pose estimation with spatio-temporal features". IEEE Transactions on Biomedical Engineering, 2020. Vol.67-8. p.2370–2380. DOI: https://doi.org/10.1109/TBME.2019.2961448

[16] Wang, S.C., Guo, C.X., Yang, R.F., Zhang, Q.C, Ren, H.Y., "A lightweight vision-based measurement for hand gesture information acquisition". IEEE Sensors, 2023. Vol.23-5. p.4964–4973. DOI: https://doi.org/10.1109/JSEN.2022.3204641

[17] Groos, D., Ramampiaro, H., Ihlen, E., "EfficientPose: scalable single-person pose estimation". Applied Intelligence, 2021. Vol.51-4. p.2518–2533. DOI: https://doi.org/10.1007/s10489-020-01918-7

[18] Xu, Y.F., Zhang, J., Zhang, Q.M., Tao, D.C., "Vitpose++: vision transform for generic body pose estimation". IEEE Transactions on Pattern Analysis Machine Intelligence, 2024. Vol.46-2. p.1212–1230. DOI: https://doi.org/10.1109/TPAMI.2023.3330016

[19] Huang, Y. W., Lin, Z. Q., Zhang, J., Chen, J. K., "Lightweight human pose estimation algorithm combined with coordinate transformer", Journal of Graphics, 2024. Vol.45-3. p.516–527 (in Chinese). DOI: https://doi.org/10.11996/JG.j.2095-302X.2024030516

[20] Yao, Y.P., Xu, C., Chen, H.J., Liu, Y., Xu, S.L., "Estimation of pig body measurements based on keypoint detection and multi-object tracking". Journal of South China Agricultural University, 2024. Vol.45-5. p.722–729 (in Chinese). DOI: https://doi.org/ 10.7671/j.issn.1001-411X.202404012

[21] Newell, A., Yang. K., Deng. J., "Stacked Hourglass Networks for Human Pose Estimation". 14th European Conference on Computer Vision, Amsterdam, Netherlands:ECCV.2016. p.483–499. DOI: https://doi.org/10.1007/978-3-319-46484-8_29

[22] Zhou, F.X., Jiang, Z.H., Liu, Z.H., Chen, F., Chen, L., Tong, L., Yang, Z.L., Wang, H.K., Fei, M.R., Li, L., Zhou, H.Y., "Structured context enhancement network for mouse pose estimation". IEEE Transactions on Circuits and Systems for Video Technology, 2022. Vol.32-5. p.2787–2801. DOI: https://doi.org/10.1109/TCSVT.2021.3098497

[23] Sun K, Xiao B, Liu D, et al. "Deep high-resolution representation learning for human pose estimation". Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Long Beach, USA: IEEE.2019. p.5693-5703. DOI: https://doi.org/10.1109/CVPR.2019.00584

[24] Zhang, H., Dun, Y.J., Pei, Y.X., Lai, S.Q., Liu, C.X., Zhang, K.P., Qian, X.M., "HF-HRNet: a simple hardware friendly high-resolution network", IEEE Transactions on Circuits and Systems for Video Technology, 2024. Vol.34-8. p.7699–7710. DOI: https://doi.org/10.1109/TCSVT.2024.3377365

[25] Tripathi, A., Singh, T., Nair, R., Duraisamy, P., "Improving early detection and classification of lung diseases with innovative MobileNetV2 framework", IEEE Access, 2024. Vol.12. p.116202–116217. DOI: https://doi.org/10.1109/ACCESS.2024.3440577

[26] Ji, Y.H., Xu, J.P., Yan, P.P., Xue, J.L., "A method for detecting quality and defects in raw coffee beans based on improved ResNet50 model". Journal of Chinese Agricultural Mechanization, 2024. Vol.45-4. p.237–243 (in Chinese). DOI: https://doi.org/10.13733/j.jcam.issn.2095-5553.2024.04.034

[27] Peng, H.X., He, H. J., Gao, Z.M., Tian, X.G., Deng, Q.T., Xian, C.L., "Litchi diseases and insect pests identification method based on improved shuffleNetV2". Transactions of the Chinese Society for Agricultural Machinery, 2022. Vol.53-12. p.290–300 (in Chinese). DOI: https://doi.org/10.6041/j.issn.1000-1298.2022.12.028

SUPPLEMENTARY MATERIAL

https://www.revistadyna.com/documentos/pdfs/_adic/11350-1_en.pdf