RESEARCH ARTICLE

Research on fish identification in tropical waters under unconstrained environment based on transfer learning

Shan Zhang[1] · Weifang Liu[1] · Yupeng Zhu[2] · Wei Han[1] · Yi Huang[1] · Jianqiang Li[1]

Abstract

Seafloor unconstrained environment video is shot in uncontrolled open sea area. There are multiple backgrounds, complex illumination and weather changes, and rapid growth of algae and attached lenses, which affect the stability of video shooting quality, resulting in difficulty in image recognition. At present, there is still no algorithm that is superior to other methods generally, and it is necessary to build a model combined with specific scenes and applications. In this paper, a fish identification method in tropical waters based on transfer learning under unconstrained environment is proposed. Firstly, the image is pre-processed by affine transformation to realize data enhancement. Furthermore, RestNet50 deep convolutional neural network is constructed based on transfer learning to compare the effect of fish recognition before and after transfer learning. The results show that, the accuracy and loss indicators are better than those of non-transfer learning when the trained model of imagenet is introduced as the initial weight of the network. When the model is trained to 150 epochs, the indicators begin to converge, which can better complete the fish identification task in tropical waters under unconstrained environment.

Keywords unconstrained environment · transfer learning · data enhancement · fish identification · tropical waters

China is one of major marine countries in the world with an area of more than 3 million square kilometers and rich marine resources. However, compared with developed countries, the development and utilization of marine resources in China still needs to be further improved. Fish identification is the first step in the investigation of marine fish resources and an important basis for the effective development and utilization of marine biological resources. Tropical fishes usually have bright appearance and are one of the most energetic marine organisms. Their growth status, species richness, and activity trajectory can directly or indirectly reflect the biodiversity of the waters. Scientists can also understand the underwater environment, the behavior of different marine animals and their mutual influence through the analysis of fish resources. Thus, the classification and identification of fish in tropical waters has important strategic significance for the effective development and utilization of fishery resources and marine environmental protection in China.

In the past, the monitoring of marine fish resources was often carried out by field sampling method (Li et al. 2011), fishing method (Hodgson 2001) and hydroacoustics telemetry. These methods have some shortcomings such as dependence on expert experience, destruction of biological resources and inability to determine the type of fish (Yang 2015). Thanks to the improvement of data processing technology, evaluation model and method, monitoring and evaluation of marine biological resources step into the age of precision and nondestructive (Zhang and Zhu 2017). The establishment of underwater video deep learning system will be the most effective method for intelligent monitoring of marine fish resources.

As a non-destructive technique, computer vision capture and analysis technology can maintain the integrity of marine

✉ Shan Zhang
zhangshan@126.com

Weifang Liu
2638940612@qq.com

Yupeng Zhu
xiaopeng@126.com

Wei Han
291254631@qq.com

[1] School of Tourism, Hainan University, Haikou city, China

[2] Hainan Chaotu Technology Co., Ltd, Hainan, China

biological resources and bring the least environmental disturbance in monitoring, which is widely used in marine biological monitoring research. Seabed images are acquired by photography and video and ocean biological characteristics are studied by image processing (Evans 2003; Hsiao et al. 2013; Zhang and Zhu 2018; Zhang et al., 2019, 2020). Compared with traditional methods, using computer technology for marine biological monitoring is an automatic, non-invasive, economic and effective method.

Although the image-based target detection technology has been widely studied in recent years, the target detection technology based on images in seabed unconstrained environment still faces great challenges. Due to the fast and free movement of fish, weather, illumination and other external factors change frequently (Toh et al. 2009), resulting in low contrast and fuzzy texture details of video images, which is not conducive to train network to extract image features.

Fish image recognition algorithm

The basic process of fish image recognition: input fish image, select fish feature, build classifier, classification and recognition. Early methods over-rely on artificial selection features, which depend mainly on expert prior knowledge, through fish contour, shape, color and other features to detect (Zion et al. 2008; Chomtip et al. 2016; Alsmadi et al. 2016), the number of parameters allowed in feature design is fairly limited, and it is difficult to accurately obtain high-level features (Hsiao et al. 2013), thus affecting the accuracy of classification and recognition results. With the development of artificial intelligence technology, deep learning has achieved great success in the field of computer vision. The traditional machine learning method is gradually being replaced by the method based on deep learning. Deep learning can automatically learn the representation of features from big data and contain thousands of parameters. Mapping data from pixel level low dimensional features to semantic level high dimensional features through deep neural network structure, which makes it have prominent advantages in extracting the global features and context information of the image (Yu et al. 2013) and brings new ideas to solve the traditional computer vision problems (such as image segmentation and key point detection).

After 2012, especially the neural network model represented by Convolutional Neural Network (CNN) has been widely used (Frederic et al. 2015; Xu et al. 2015; Li et al. 2019). CNN is a popular deep learning architecture composed of alternating convolution layer and sub-sampling layer, since it can extract complex and high-dimensional image features, it has been proved to be effective in many computer vision applications, such as segmentation (Girshick et al. 2014; Ciresan et al.

2012), object recognition and classification (Szegedy et al. 2013), target tracking (Zhou et al. 2015). In 2017, Google and NASA found a new planet in Kepler-90 massive images through CNN technology.

Frederic et al. (2015) also found endangered species of manatee from hundreds of thousands of satellite images through CNN. Spampinato et al. (2012) proposed an algorithm for tracking fish in unconstrained environment, the detected object was described based on the covariance model, and the covariance matrix of the current object was compared with that of the candidate object. The trajectory was generated by template matching, and the average accuracy was about 80%, which needs to be further improved. To solve these problems, Chuang et al. (2015) further proposed a tracking algorithm based on deformable multi-core. Through the mean shift algorithm of color and texture features, the motion of the kernel was effectively estimated and the fish tracking was realized. However, the specific test results were not given in this paper.

On this basis, Jager et al. (2017) realized fish tracking in unconstrained environment with activated CNN through two-stage graph method, which has high tracking accuracy and efficiency. However, the 14 super parameters of the model cannot be adjusted automatically, thus reducing the portability of the model. Wang et al. (2017) customized CNN for the head image of each fish in the video frames and adaptively updated CNN parameters in the tracking process to adapt to slight illumination and fish appearance changes, this method has been proved to have good recognition and tracking effect for zebrafish, but the experiment is still limited to the aquarium environment.

Data sources

As the artificial fish reef has a good recovery effect on the marine resources of Wuzhizhou Island, the fish population is relatively dense in this area, and the resources are relatively abundant, so we selected sea areas where the artificial reefs were placed at Wuzhizhou island in the early stage for monitoring, in order to identify more tropical fish in the unconstrained marine environment. Since 2017, the project team has selected five artificial reefs, as presented in Fig. 1, to film and observe marine fish around Wuzhizhou Island in Sanya City, Hainan Province. The location, water depth and artificial reef type of these five monitoring points are shown in Table 1.

The underwater electronic cabin is arranged in five observation points, and two photoelectric HB-HM-02 underwater cameras are arranged in one cabin for data acquisition from different angles. Video resolution is 648 × 480, width and height, frame rate is 50. At present, all five sets of equipment have been arranged and operated

Table 1 Illustration of underwater video surveillance in Wuzhizhou Island

Monitoring points	Location	Water Depth	Artificial reef type	biological resources characteristics
1	Lover Island (109.768°E, 18.316°N)	13 m	Steel ship reef	Many corals, clams and fishes
2	SunView Rock (109.768°E, 18.312°N)	14 m	Cement circular truncated cone reef	Plenty hard corals and fishes
3	Tea Restaurant (109.763°E, 18.307°N)	25 m	Cement cube reef	Plenty soft corals and fishes
4	Winter Reefs (109.760°E, 18.308°N)	20 m	Cement cube reef and circular truncated cone reef	Plenty corals and fishes
5	Seabed Garden (109.765°E, 18.316°N)	12 m	Steel ship reef and Cement cube reef	Many corals, clams and fishes

well, and continuous and uninterrupted recording and transmission of video data and water quality monitoring data have been realized.

Ten cameras generate 150 + MB video data per minute (the camera does not shoot video during poor night illumination). Under the monitoring of 12 h, more than 100 GB video data are generated every day and 35 + TB video data are generated every year. This area has rich fish biological resources, which provides a solid guarantee for the data source of this paper. In addition, the Fish4Knowledge large underwater video library derived from LifeCLEF 2015 fish task 1 (Fisher et al. 2016) is also used as the experimental data. The video library contains more than 700,000 submarine videos, mainly fish images taken from the underwater landscape platforms of Taiwan's Nanwan Strait, Lantau Island and Hubi Lake, the videos are taken from September 30, 2013 to October 1, 2019.

Technical information on the above two datasets are presented in Table 1 below. The LCF-15 dataset has been marked by experts, with a total of 42,778 marked video frames. Selected 24 video clips of different time periods in Wuzhizhou Island database, a total of 3100 labeled video frames. 45,878 video frames were marked totally, including 32,290 for training set and 13,588 for test set (Table 2).

Fig. 1 Location map of underwater video surveillance in Wuzhizhou Island

Table 2 Overview of LCF-15 and Wuzhizhou Island database

Database name	Video quantity	Video resolution	FPS	Marked video frames	Training set samples	Test set samples
LCF-15	93	640×480	24	42,778	29,965	12,813
Wuzhizhou Island	24	640×480	50	3100	2325	775

Seabed video reflects the surrounding marine environment, marine fish color, shape, type and other information. Figure 2 is shown the video frame in LCF-15 database.

The video frame effect of Wuzhizhou Island dataset is shown in Fig. 3, which marks 3100 video frames.

Table 3 shows the information and video maps of some species appearing in the above dataset. The comparison shows that the image quality is low in the seabed unconstrained environment, and it is difficult to obtain high performance for image detection and classification tasks.

Data pre-processing

Convert video to picture

The traditional remote video surveillance system requires large hardware and software resources, needs high bandwidth and has long delay for data transmission (Xiao and Cai 2015). The seabed video of Wuzhizhou Island is compressed using H. 264 coding standard. On the basis of retaining the advantages of the original video coding, it has the advantages of
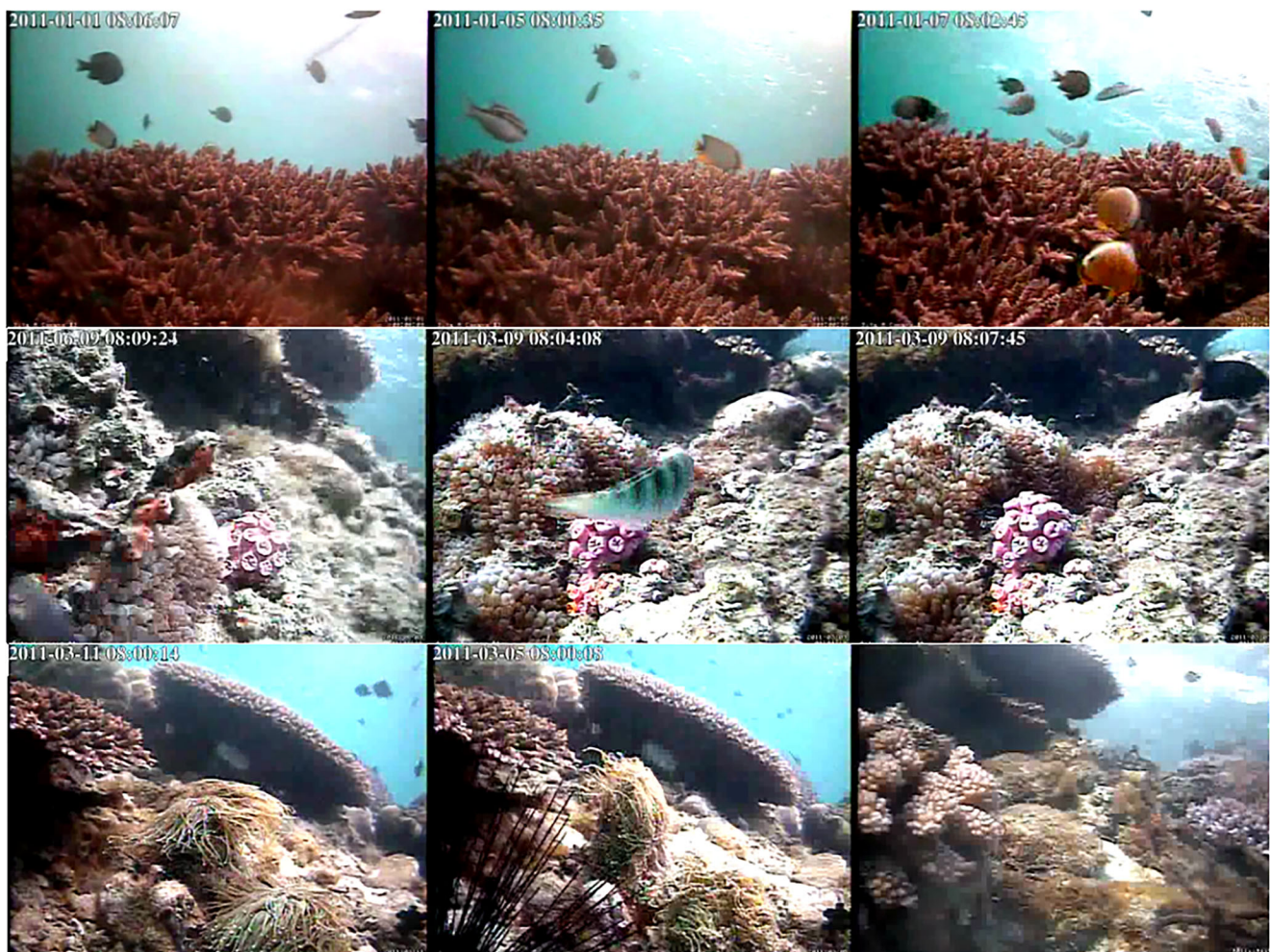
Fig. 2 Visual effect diagram of video frame in LCF-15 database

Fig. 3 Visual effect diagram of video frame in Wuzhizhou Island database

high compression rate, strong network adaptability and good anti-interference performance. FFmpeg is an open source solution for video codec. The system uses the ffmepg method to obtain the frame number, frame rate, video time and other parameters of the video file through the imageio library, then reads the video frames and stores it as pictures based on OpenCV according to user needs.

The flowchart of video to picture is shown in the Fig. 4.

Data enhancement

In order to obtain more image samples, expand the sample size, and improve the accuracy and generalization ability of the model, this paper mainly uses ImageDataGenerator based on Keras image processing to enhance the original training data. This method mainly expands the data set by rotating the image between −40° and 40°, zooming in the proportion of 0.2–1 / 0.2, shifting / clipping 0–0.2 percentage along the horizontal or vertical direction, and filling the nearest neighbor with dew. The processing flow is shown in Fig. 5 and the data enhancement effect is shown in Fig. 6.

Transfer learning
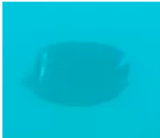
Data is the key of deep learning, and having enough available training samples is the premise of computer learning to obtain good classifier. Most of the available data are designed for

specific subjects, such as handwritten digital database MINIST, image dataset ImageNet, image semantic understanding dataset COCO and so on. These open source data cannot fully meet the requirements of deep learning training data required in this study, so it is necessary to supplement the training data for specific fish. However, manual image annotation is very time-consuming, it takes an average of 1.5 h to label a single image in the driverless dataset Cityscapes at the fine pixel level. In addition, mislabeling are prone to occur due to subjective factors. Therefore, the project introduces transfer learning to solve the problem that there is only a small amount of labeled sample data in the current target area.

Transfer learning is a new machine learning method that uses trained knowledge to solve problems in different but related fields, so as to solve the problem of only a small number of labeled sample data in the target field. In the past, deep learning models need to be retrained for different fields. On the one hand, it takes a lot of time and computing resources. On the other hand, in the field of image recognition, the underlying features of the image are highly similar, and combined into different and more complex features for image classification and recognition. Accordingly, the network model after pre-training of other large training sets can be used as the underlying image feature extractor (Dai et al. 2007; Lawrence and Plat 2004; Lee et al. 2007), so as to realize the transfer of general image feature extraction knowledge to the field of tropical fish classification and recognition, and reduce the dependence on labeled data. " Transfer learning

Table 3 Comparison of different fish between data maps and video maps.

Class name	data image	Video image
Abudefduf vaigiensis		
Coris gaimard		
Stethojulis strigiventer		
Dascyllus trimaculatus		
Caranx melampygus		
Chaetodon auripes		
Scarus dimidiatus		
Ringtailwrasse		
Lutjanus Quiquelineatus		

will be the driving force for the next machine learning business success following supervised learning." Professor Wu Enda of Stanford University said at the 2016 NIPS conference.

The ImageNet image dataset, which began in 2009, is the world 's largest image recognition database currently, it was built by Stanford computer scientists simulate human recognition system. Including about 15 million pictures, 22,000 types, Each picture has been rigorously screened and marked manually. Since 2010, the large-scale visual knowledge challenge (ILSVRC) around the world has been held annually based on ImageNet. Research teams around the world evaluate their algorithms on a given dataset and compete for higher accuracy in visual recognition tasks. After

Fig. 4 Flowchart of video to picture
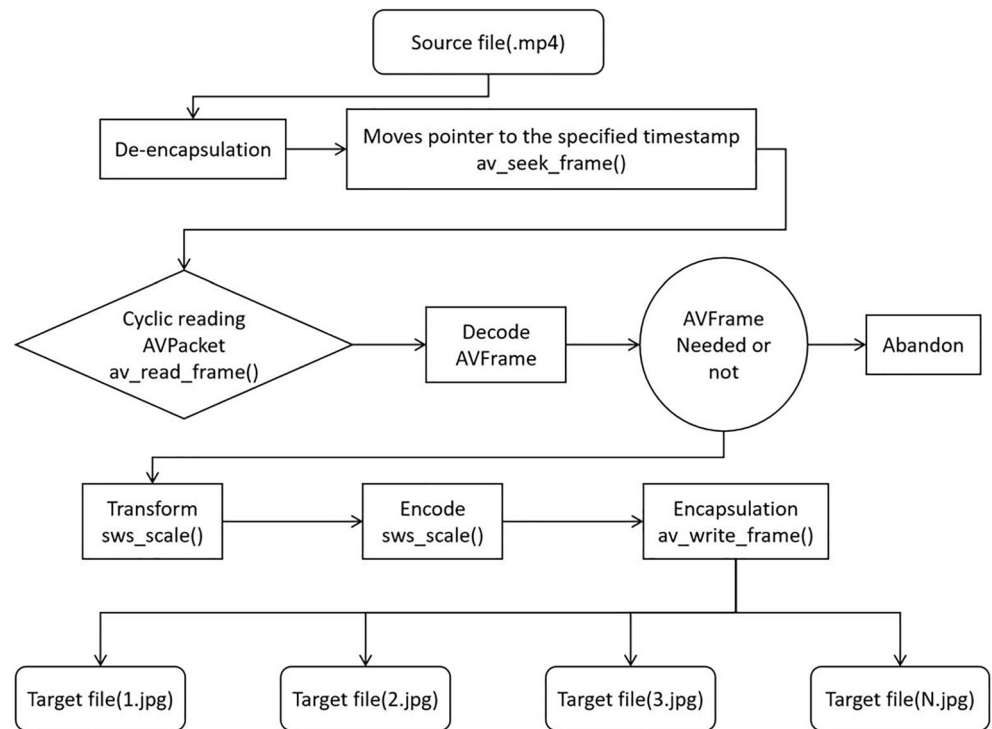
2017, the competition is hosted by the Kaggle community. In terms of image classification, ImageNet competition has become a de facto evaluation standard for computer vision classification algorithms. Keras is an open source artificial neural network library written in Python, which mainly provides five open-box ready-to-use CNN networks based on ImageNet training (VGG16, VGG19, ResNet50, Inception V3, Xception). Unlike traditional sequential network architectures such as AlexNet, OverFeat and VGG, ResNet50 's network architecture relies on microarchitecture modules because

global average pooling is used instead of full connection layer. Although ResNet50 is deeper than VGG16 and VGG19, weights is smaller. In this paper, the ResNet50 network in the Keras library is used to initialize training target model parameters by using the trained imagenet as the weight, so as to realize the sharing of learned knowledge to the new model. Then, the output layer is adjusted according to the target task, and the network model is trained by using the tropical fish image samples to improve the accuracy and validity of the detection algorithm. Finally, the fish classification

Fig. 5 Data enhancement flowchart
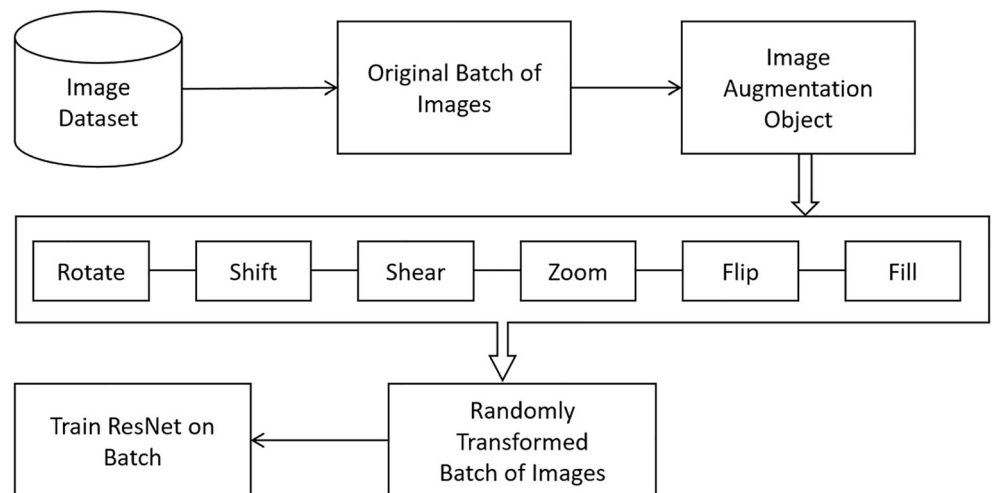
Fig. 6 Diagram training data enhancement (horizontally flip, rotate, zoom)

and recognition model in the tropical sea area is obtained. The transfer learning process is shown in Fig. 7.
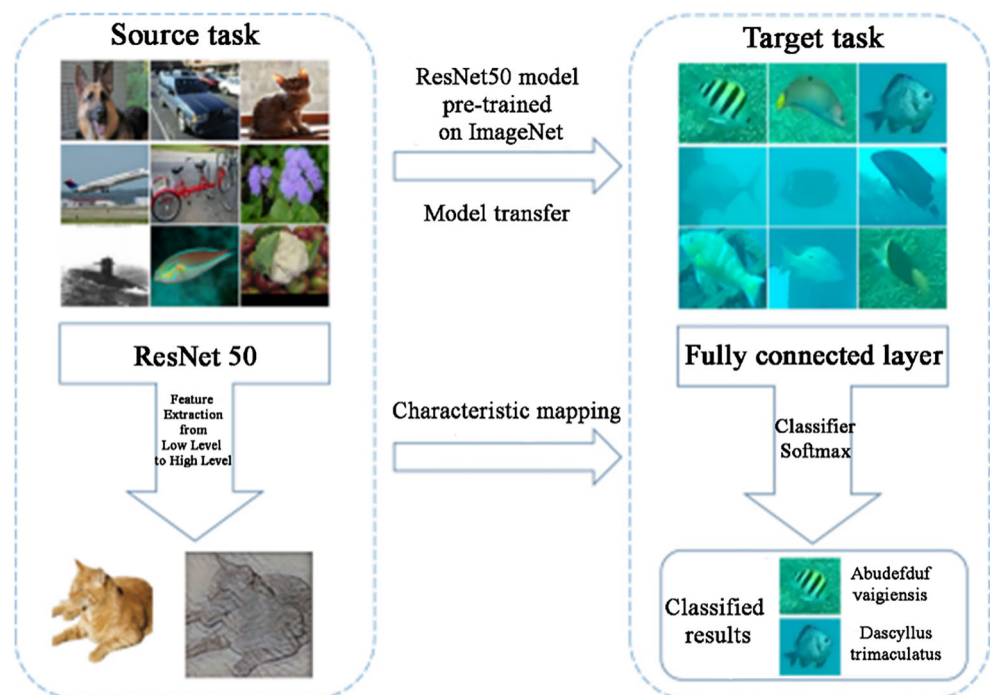
Model architecture

The entire ResNet deep convolution neural network consists of the initial convolution (Conv) layer, several Residual Blocks and the final fully connected layer (FC). The first Conv layer does not require normalization and ReLU nonlinear activation.

Conv layer in Residual Block normalizes images according to mean and standard deviation, taking into account the training time of gradient descent, the ReLU nonlinear activation which is several times faster than the equivalent tanh unit is used as the activation function (Hinton et al. 2012). Then the

first 3×3 convolution operation is performed, and the convolution results are batch normalization again. After ReLU nonlinear activation, the second 3×3 convolution operation is performed. Implement a long jump connection, pass the activation of the first layer quickly to the next Residual Block after the sixth layer. Total training times (train steps) of the model is 300, the image is convolved by a series of 3×3 filters, the convolution sliding step is fixed to 1, and the algorithm performance is optimized by mini _ batch. Considering the limited memory, the batch _ size of training set and test set are both set to 24, use padding size to fill the same image. The total parameters of the model are 2358 7712, the fixed parameters are 53,120, and the training parameters are 135 34,592. The model structure and parameter setting are shown in Fig. 8.

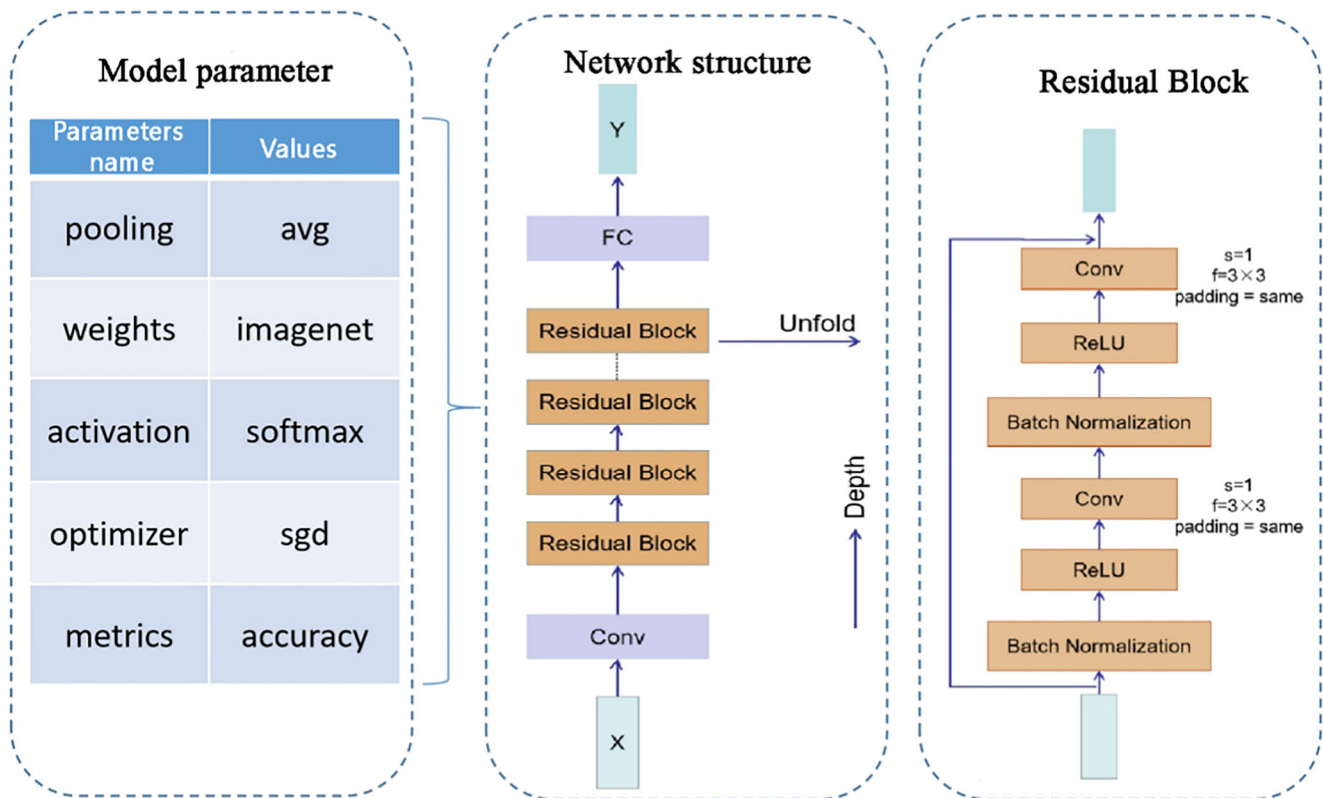**Fig. 7** Transfer learning flow chart

Fig. 8 Architecture of ResNet tropical fish identification model

Experiments and results analysis

Undersea video data in LCF-15 datasets were acquired from south of Taiwan coral reef areas.The marine climate of Hainan and Taiwan is remarkable, and both have tropical climate. Due to its proximity to the equator and gentle terrain, Hainan has a tropical monsoon climate with long summer and short winter.

Taiwan has high latitude, complex terrain, large terrain drop, and tropical monsoon climate in the south, which is similar with Hainan province. In order to avoid the difference of fishery resources caused by the difference of natural conditions between the two places, this study selected the dominant fish species in Taiwan and Hainan to make the data set. There are mainly 9 categories as follows: *Abudefduf vaigiensis,Coris*

Fig. 9 Train loss and validation loss changes after transfer learning training
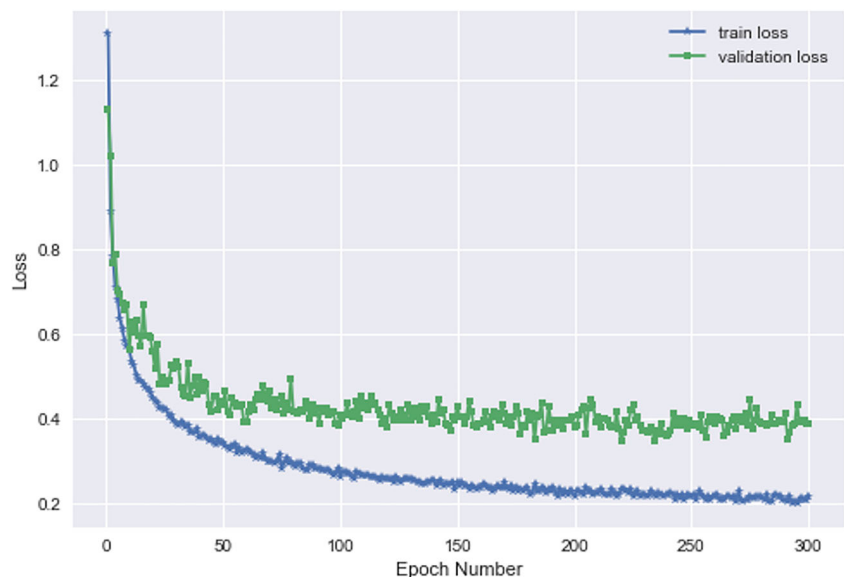
Fig. 10 The change of train accuracy and validation accuracy after transfer learning training
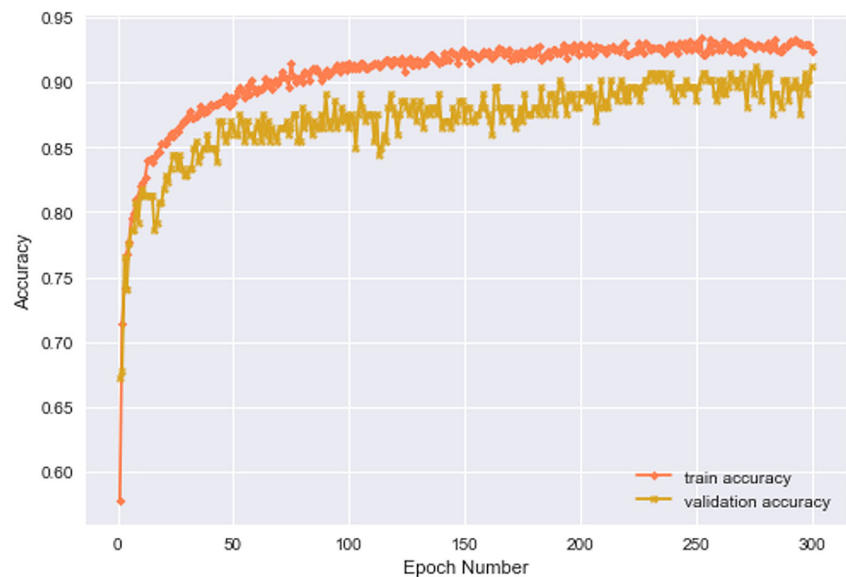
*gaimard,Stethojulis strigiventer,Dascyllus trimaculatus,Caranx melampygus,Chaetodon auripes,Scarus dimidiatus,*Ringta ilwrasse,Lutjanus Quiquelineatus.

In order to evaluate transfer learning, we conducted experiment on LCF-15 and Wuzhizhou Island datasets, and compared the learning indicators before and after transfer learning with the network structure, learning efficiency and loss function unchanged. An Intel (R) Xeon (R) X5675 3.06 GHz computer system processor, 32 GB memory and NVIDIA GeForce GTX 1050 Ti GPU were used in the experiment.

Train loss represents the loss of training data. Cross entropy is used as a loss function to measure the fitting ability of the training set. The lower the train loss value, the higher the fitting ability of the model on the training set. Validation loss represents the loss of the validation set. Cross entropy is used as the loss function to measure the fitting ability of the model on the unknown data, that is, the generalization ability of the model. The lower the validation loss value, the higher the

generalization ability of the model. It can be seen from Fig. 9 that under the introduction of the imagenet trained model as the initial weight of the network, the train loss value and the validation loss value decrease rapidly, and with the increase of training times, these two value maintain a continuous downward. After 150 times of training, the validation loss value oscillates between 0.41 and 0.37. With the deepening of training, the loss value further oscillates downward, indicating that this method has good fitting ability and generalization ability.

Train accuracy represents the accuracy on the training data, and validation accuracy represents the accuracy on the validation set. The higher the accuracy is, the higher the accuracy of the model for tropical fish classification and recognition is. It can be seen from Fig. 10 that under the introduction of imagenet trained model as the initial weight of the network, the values of train accuracy and validation accuracy increase rapidly, and with the increase of training times, the values of train accuracy and validation accuracy keep increasing trend.

Fig. 11 Comparison of train loss and validation loss before and after transfer learning

Fig. 12 Comparison of train accuracy and validation accuracy before and after transfer learning
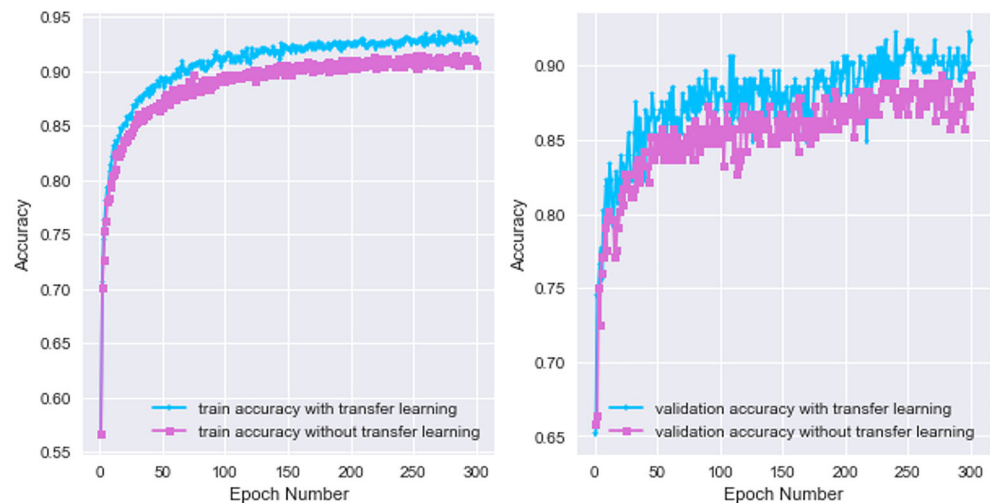
After 150 times of training, the validation accuracy value oscillates between 0.85 and 0.90. With the deepening of training, the accuracy value further oscillates upward, indicating that this method has good recognition accuracy.

While maintaining the same network structure and other hyper parameters, this paper compares the changes of train loss and validation loss indexes before and after using transfer learning. From Fig. 11, it can be seen that the train loss values decrease to about 0.2–0.3 with the increase of training times before and after using transfer learning. At the initial stage of training, the loss will decrease significantly. After about 100 training sessions, the train loss will basically decrease to below 30%, and the validation loss value will oscillate between 0.25 and 0.45. Loss after transfer learning is significantly lower than before transfer learning.

In the case of maintaining the same network structure and other hyper parameters, this paper compares the changes of train accuracy and validation accuracy indexes before and after using transfer learning. It can be seen from Fig. 12 that before and after using transfer learning, train accuracy values increased to about 0.9 with the increase of training times. In the initial stage of training, accuracy will have a significant increase. When the training is about 50 times, train accuracy is basically stable above 0.8. The accuracy after transfer learning is obviously higher than that before transfer learning.

This paper takes tropical fish in the marine video surveillance data under unconstrained conditions as the research object, constructs a tropical fish recognition model based on transfer learning with ResNet50 as the basic training network. The results show that, transfer learning has a good effect on the computer recognition of tropical fish in unconstrained environment, which is helpful to improve the recognition accuracy and shorten the learning time. In the future, we will further obtain ultrahigh-definition video data and optimize the algorithm to solve the problem of convergence to local optimum rather than global optimum. At the same time, we will

improve the hardware conditions, further increase the processing accuracy.

References

Alsmadi M K A,Omar K B,Noah S A, et al (2016) Fish recognition based on the combination between robust features selection,image segmentation and geometrical parameters techniques using artificial neural network and decision tree[J].International Journal of Computer Science and Information Security,2016,6(2):215–221

Chomtip P,Pimprapai L,Waranat K, et al (2016) Thai fish image recognition system[C]//The Proceedings of International Joint Conference on Computer Science and Software Engineering. Bangkok, Thailand.2016:115–120

Chuang MC, Hwang JN, Williams K et al (2015) Tracking live fish from low contrast and low-frame-rate stereo videos. IEEE Trans Circuits Syst Video Technol 25(1):167–179

Ciresan D, Giusti A, Gambardella LM et al (2012) Deep neural networks segment neuronal membranes in electron microscopy images. Adv Neural Inf Proces Syst:2843–2851

Dai W , Yang Q , Xue G R , et al. (2007) Boosting for Transfer Learning[C]// Machine Learning, Proceedings of the 24th

International Conference (ICML 2007), Corvallis, Oregon, USA, June 20-24, 2007. 193–200

Evans FH. 2003. Detecting fish in underwater video using the EM algorithm. International Conference on Image Processing, 3(III): 1029–1032

Fisher, R., Chen-Burger, Y., Giordano, D., Hardman, L., Lin, F., 2016. Fish4Knowledge: Collecting and Analyzing Massive Coral Reef Fish Video Data , Intelligent Systems Reference Library. Springer International Publishing URL. https://books.google. com.pk/books?id=j846jwEACAAJ

Frederic M, Alvarez LM, Hodgson A. 2015. Automating marine mammal detection in aerial images captured during wildlife surveys: a deep learning approach. Australasian Joint Conference on Artificial Intelligence, 379-385

Girshick R, Donahue J, Darrell T et al (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. IEEE Conference on computer vision and pattern recognition 2014:580–587

Hinton GE, Srivastava N, Krizhevsky A et al (2012) Improving neural networks by preventing co-adaptation of feature detectors[J]. arXiv e-prints

Hodgson G (2001) Reef Check: The first step in community-based management[J]. Bulletin of Marine Science -Miami-

Hsiao YH, Chen CC, Lin SI et al (2013) Real-world underwater fish recognition identification,using sparse representation[J]. Ecological Informatics 23(9):13–21

Jager J, Wolff V, Fricke NK et al (2017) Visual fish tracking: Combining a two-stage graph approach with CNN-features. IEEE OCEANS:1–6

Lawrence ND, Plat JC (2004) Learning to learn with the informative vector machine[C]//Machine Learning, Proceedings of the 21stst International Conference (ICML 2004). New York:ACM:65–72

Lee SI, Chatalbashev V, Vickrey D (2007) Koller D.Leaming a meta—level prior for feature relevance from multiple related tasks[C]//Machine Learning, Proceedings of the 24th International Conference (ICML 2007), Corvallis, Oregon, USA, June 20-24, 2007.489–496

Li HT, Liu TL, Shao ZD (2019) Application of a BP neural network based on decision tree and two partition algorithm in prediction of red tide severity[J]. Mar Sci 43(09):34–40

Li YZ, Shi ZR, Ai H et al (2011) Large scale distribution patterns of taxonomic diversity of fish in coral reef waters, South China Sea[J]. Journal of Fishery Sciences of China 8(3):619–628

Spampinato C., Palazzo S., Giordano D., et al.2012. Covariance based fish tracking in real-life underwater environment. International Conference on Computer Vision Theory and Applications, 409–414

Szegedy C, Toshev A, Erhan D (2013) Deep neural networks for object detection. Adv Neural Inf Proces Syst:2553–2561

Toh YH., Ng TM., Liew BK. 2009. Automated fish counting using image processing. International Conference on Computational Intelligence and Software Engineering, 1–5

Wang SH, Zhao JW, Chen YQ (2017) Robust tracking of fish schools using CNN for head identification. Multimed Tools Appl:86–105

Xiao YD, Cai SZ (2015) Design of Mobile Video Surveillance Software Based on iOS Platform[J]. Comput Eng 41(2):268–271

Xu Y, Du J, Dai LR, et al (2015) A regression approach to speech enhancement based on deep neural networks[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 7–19

Yang H S (2015) The principle and practice of constructing marine ranch[M].Beijing: science press.2015

Yu K., Jia L., Chen Y Q, et al (2013) Deep Learning: Yesterday, Today, and tomorrow[J]. Journal of Computer Research and development, 26.,2013, 50(9): 1799–1804

Zhang S, Zhu Y (2017) Research on the Construction of Underwater Video Surveillance and Water Environment Monitoring Information System for Marine Ranching[J]. Boletín Técnico 55(17):738–742

Zhang S, Zhu Y P (2018) Simulation of Regional Species Identification of Marine Organisms Based on 3D Vision[J]. Computer Simulation, 29.2018.35(5) : 402–406+443

Zhang S, Zhu Y, Bin Y (2019) Tropical coral reef fish identification method based on ResNet in unconstrained environment of the seabed[J], Revista Científica 2019(107):1157–1164

Zhang SM, Liu Y, Fan W et al (2020) Aquarium fish target detection APP development based on TensorFlow[J]. Fishery Modernization 2:60–67

Zhou X., Xie L., Zhang P. 2015. Online object tracking based on CNN with metropolis-hasting re-sampling. Proceedings of the ACM international conference on multimedia, 1–4

Zion B, Alchanatis V, Ostrovsky V, Barki A, Karplus I (2008) Classification of guppies Poecilia reticulata gender by computer vision[J]. Aquac Eng 38(2):97–104