A ConvNet-based Marine Animals Automatic Detection and Removal System in Underwater Photogrammetry

Jiageng Zhong1, Ming Li1*, Jiangying Qin1, Yuxin Cui2

[1]Wuhan University (LIESMARS),
Luoyu Road NO.129, Wuhan, China. zhongjiageng and jy_qin@whu.edu.cn, mingli39@ethz.ch
[2]Hongyi Honor College of Wuhan University,
Bayi Road NO. 299, Wuhan, China. yuxincui@whu.edu.cn

Abstract: In recent years, with the advancement of marine resources and environment research, the ecological functions of reef-building coral reef ecosystems distributed in warm shallow waters of the ocean are being continuously discovered and valued by people. Because it is closely related to global climate warming and the development of marine resources, the demand for its dynamic and normalized monitoring is increasingly urgent. In order to achieve efficient, accurate, and automated large-scale coral surveys, traditional water surface photogrammetry and remote sensing technologies (satellite remote sensing, drone remote sensing, and surface ship remote sensing) cannot meet the requirements. Underwater photogrammetry technology must be used to achieve this goal. Millions of coral images acquired by underwater ROV/AUS/UUV every year, but in many images, marine animals such as fish or turtles appear to block coral on the seabed in a large area. This situation is very similar to traditional close-range photogrammetry or nap-of-the-object photogrammetry. Therefore, in order to reduce the interference of moving objects on the observation objects, these moving objects must be handled properly. Therefore, how to efficiently check occluded images and eliminate invalid images has become a difficult task at present, and traditional manual recognition and processing cannot handle the ever-increasing massive data processing.

In view of the difficulty, an underwater image sieving and processing scheme is proposed, the key technology of which is the object detection model and the salient object detection model. For underwater images, it is necessary to determine the type and number of the objects at first, so the underwater object detection model is used to extract all the marine animals in the images, such as fish and turtles. Aiming at the characteristics of underwater vision, on the basis of preprocessing, an improved algorithm based on YOLO is used to achieve efficient and accurate underwater object detection. According to the type and number output from object detection, the area of the image occluded by marine animals can be quickly estimated, and then the image is initially sieved based on the estimation result. Specifically, it is necessary to remove images which are covered by a large area of occlusions or contain too many occlusions. If the number of marine animals and the occluded area in the image is small, and other parts contain important information, the occluded area of the image will be cropped first. The salient object detection model is then used to segment its accurate boundary, and the occluded object will be erased, so as to provide better data for the subsequent underwater photogrammetry automatic interpretation tasks. To accurately eliminate the occlusion, it is necessary

to extract the precise boundary of the object as we told before. While the existing models usually pay more attention to the accuracy of the subject and ignore the importance of boundaries, a novel salient object detection model with multi-level feature fusion and emphasis on boundaries is proposed in this paper. It consists of a prediction network and a refinement network. The prediction network is a supervised Encoder-Decoder network which uses U-Net architecture, and it can output initial saliency map by combining high-level and low-level image features. The refinement network refines the saliency map via learning the residuals between the initial saliency map and the ground truth. The training loss function is a hybrid loss, which consists of binary cross-entropy loss and structural similarity loss. The former loss weights the foreground and background pixels equally so that it can evaluates the precision of predicted result from the whole. And the latter assigns higher weights to the boundary thus it helps to gain saliency maps with more accurate edges. By using the hybrid loss, the loss of the predicted saliency map can be measured from different levels, thereby the region of the salient object is accurately predicted while the fine structure is precisely segmented. Considering the lack of high-resolution underwater image dataset, a new underwater object detection dataset is constructed by this paper, which includes a fish detection dataset and an underwater salient object detection dataset.

The results of the experiment on the dataset show that our scheme can effectively sieve and process underwater images, as shown in Fig.1. The underwater object detection model can recognize the marine animals quickly and accurately, thus sieving the underwater image preliminarily, eliminating the images containing too many occlusions and cropping the occlusion from the remaining valid images. On this basis, the salient object detection model can identify the precise position and boundary of the object in the occluded area. And then the occlusions are erased, so that mismatches will be avoided in the photogrammetric process. It can be seen in Fig.1 that image containing quite many fish is eliminated while the image with only a few fish is preserved and the fish is erased. Furthermore, in this workflow, the object detection is fast and the salient object detection may take relative more time as it deals with boundaries. The statistical data shows that the efficiency of this method can meet the needs of real-time processing, so it can help to reduce workload.

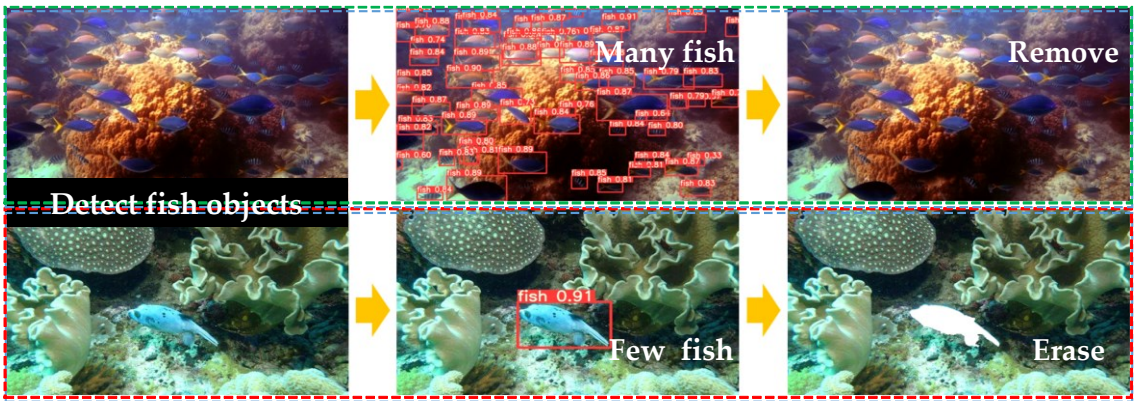

Fig.1.The process flow and effect of the method proposed in this article

In conclusion, the proposed underwater image sieving and processing scheme uses convolutional neural network to accurately locate the position of marine animals and segment them, and perform the required processing based on the analysis of the occlusion circumstances to reduce the influence of occlusions to the utmost. Its characteristics of high efficiency, accuracy and rapidness can bring quality assurance and high precision to underwater photogrammetry.

Keywords: coral images; object detection; salient object detection; underwater photogrammetry