Deep convolutional neural network and data augamentation for fish detection and classifaction

Ziqiang Zheng
College of Information Science and Engineering
Ocean University of China
Qingdao 266100, China
Email: zhengziqiang1@gmail.com

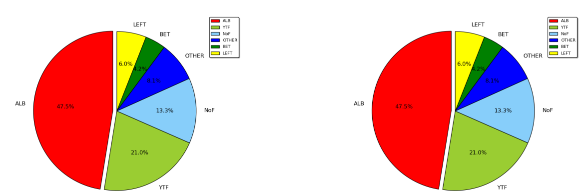
***Abstract*—Fish recognition is one of the most significance of fish survey application[1], and the fish classification is crucially required. Different from the best known and the most well investigated object detection-face detection, fish usually occupy a smaller area in an image than the face of one person. Besides some pictures are taken under water and the pictures are not clear and dark for the light is absorbed by water. For these two reasons, it's a challenge for us to detect the fish and do the classification. In this paper, We introduced two methods to improve our accuracy for the fish classification. In the western and central pacific, where 60% of the world's tuna is caught, illegal, unreported , and unregulated fishing practices are threatening marine ecosystems. The Nature Conservancy is inviting the Kaggle community to develop the algorithms to automatically detect and classify species of tunas. The Kaggle competition aims to detect and classify the species of fish. The competition provides a train dataset which contains six species of tuna fish ,fish but not tuna and no fish. Eight target categories are available in the dataset. The dataset is critically imbalanced, the Albacore tuna has thousands of images while the Opah has only sixty images. Fig. I shows the imbalance of the Kaggle dataset. Our approach can avoid over-fitting caused by the imbalance of training dataset. We employ the Alexnet, GoogleNet, and VggNet neural network for the classification. As a result of the small dataset, the VggNet architecture performs worse than AlexNet. In this paper we proposed a local region based the fish area modelling approch so that the local feature can be modelled. In order to obtain more image information and realize imbalanced datasets classification, we have done some data augmentation. In the futrue, we will use the py-faster-rcnn to do the detection[2].**

I. Introduction

In this paper, we introduced two methods to improve the accuracy. On the one hand, the dataset includes Albacore tuna, Bigeye tuna, Yellowfin tuna, Mahi, Opah, Shark, Other(meaning that there are fish present but not in the above six categories), and No Fish(meaning that no fish in the picture). Each image has only one fish category. In consideration of the imbalance of the dataset, we got a new picture by rotating the selected image. Through this method, we can increase the number of some species of fish images, which can help avoid the over-fitting. And the detection of rotation angle of the fish has a certain degree of robustness. We doubled our datasets by this way and the accuracy increases two percent. On the other hand, we used the faster-rcnn to achieve the detection and classification. There have been less efforts made on solving fish detection and classification simultaneously. The second method is to pick up some pictures to do image processing. We have made a mosaic of the area where the fish are located, and we used both the processed images and the raw images to train our model. Fig. I shows the differences between the two images. This method can force our neural network to concentrate on the area where the fish present. And the features are learned from comparing the raw pictures with the pictures with mosaic. In this method, We modified the Caffe to get the pooled image data to find the area that contributes to the classification most. Through comparing the pooling gray image with the raw image, we found that the area fish present is brighter than other areas, which proves that the fish area contributes to the classification most. Fig. II-A illustrates that the image information of the region where the fish are located has greater significance than other regions. The pooling image data can help judge if our model has learned the object features. In addition, we use the processed images to finetune our pre-trained model and it performs better when testing.

TABLE I
KAGGLE DATABASE

classes	Total	Training	Valiation
ALB	45.5%	43.1%	47.5%
YFT	19.4%	20.6%	21.0%
NoF	12.3%	13.8%	13.3%
OTHER	7.9%	7.6%	8.1%
BET	5.2%	5.4%	4.2%
Sum	90.3%	90.5%	94.0%



(a) Training database (b) Testing database

Fig. 1. Data distribution of Kaggle

II. Related Work

Recently, several papers have brought some methods for fish detection and fish classifaction[3]. Spatial constraints are

TABLE II
SINGAL MODEL RESULT ON FULL DATA

database	model	iteration	accuracy
full	alexnet	70k	0.94
full	googlenet	40k	0.942
full	vgg16	34k	0.44

TABLE III
KAGGLE DATABASE

database	model	iteration	accuracy
full+rotating	alexnet	78k	0.9673
full+rotating	googlenet	80k	0.964
full+rotating	vgg16	20k	0.45

TABLE IV
KAGGLE DATABASE

database	model	iteration	accuracy
full+rotating+mosaic	alexnet	60k	0.9703
full+rotating+mosaic	googlenet	56k	0.954

are introduced by dividing the image to regions and learning the feature region[4]. In our method, we use the image processing method to dividing the image to feature regions and background regions. And we used the mosaic images to lead our model to fit the feature regions and ignore the other regions.

A. Archicture

The architecture of our model used in the experiment is the convolutional neural network(CNN).

III. CONCLUSION

We have used two methods to help our model to detect the fish and avoid the over fitting.

REFERENCES

[1] H. Qin, X. Li, J. Liang, Y. Peng, and C. Zhang, "Deepfish: accurate underwater live fish recognition with a deep architecture," *Neurocomputing*, 2015.

[2] Z. Ge, C. McCool, C. Sanderson, and P. Corke, "Modelling local deep convolutional neural network features to improve fine-grained image classification," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 4112–4116.

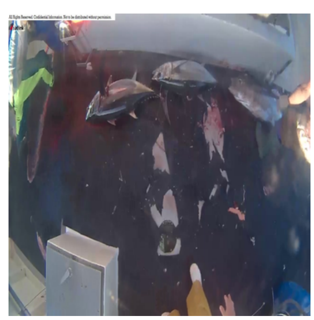
[3] M.-C. Chuang, J.-N. Hwang, and K. Williams, "A feature learning and object recognition framework for underwater fish images," *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1862–1872, 2016.

[4] K. Anantharajah, Z. Ge, C. McCool, S. Denman, C. Fookes, P. Corke, D. Tjondronegoro, and S. Sridharan, "Local inter-session variability modelling for object classification," in *IEEE Winter Conference on Applications of Computer Vision*. IEEE, 2014, pp. 309–316.

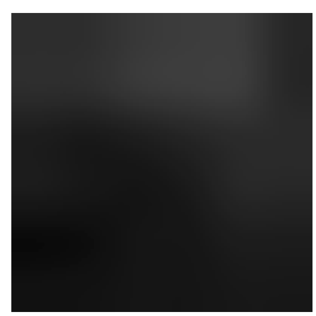
(a) the raw image

(b) image with mosaic

Fig. 2. We make a mosaic of the area where the fish are located. we have trained our model useing both the two images at the same time, which can realize the system performance and improve the robust.



(a) the image



(b) the pooling image

Fig. 3. We make a mosaic of the area where the fish are located,and the mosaic can help our model detect the fish area.