

## Multiple Plankton Detection and Recognition in Microscopic Images with Homogeneous Clumping and Heterogeneous Interspersion

Youngsung Soh<sup>1\*</sup>, Jaehyun Song<sup>2</sup>, and Yongsuk Hae<sup>2</sup>

Department of Information and Communication Eng., Myongji University<sup>1</sup>,  
Artificial Intelligence Lab in DIPS Co<sup>2</sup>.

**Abstract** The analysis of plankton species distribution in sea or fresh water is very important in preserving marine ecosystem health. Since manual analysis is infeasible, many automatic approaches were proposed. They usually use images from in situ towed underwater imaging sensor or specially designed, lab mounted microscopic imaging system. Normally they assume that only single plankton is present in an image so that, if there is a clumping among multiple plankton of same species (homogeneous clumping) or if there are multiple plankton of different species scattered in an image (heterogeneous interspersion), they have a difficulty in recognition. In this work, we propose a deep learning based method that can detect and recognize individual plankton in images with homogeneous clumping, heterogeneous interspersion, or combination of both.

- Key Words : Plankton, Recognition, Deep learning, Convolution layer, Detection

### I. INTRODUCTION

Plankton play an important role in marine ecosystem as they occupy the base of natural food chain. Some plankton species are harmful and others are not. To keep marine ecosystem healthy, it is critical to monitor the distribution of plankton species and generate alarm if needed. Monitoring can be done either manually or automatically. As can be expected, manual approaches based on human expertise are tiring, time consuming, and error prone. Thus many automatic approaches having recognition functionalities were proposed in the literature.

Until recent years recognition is mostly based on feature extraction followed by pattern recognition or

classical machine learning method. Features used include shape, texture, invariant moments, etc., and support vector machine (SVM), ensemble classifier, and shallow neural network (NN) were frequently used for recognition. Only a few adopt deep learning framework to recognize plankton species (classes) and outperform conventional approaches. However almost all deep learning based methods assume that single image has only one plankton so that they cannot handle the case where multiple plankton of same species are clumped (homogeneous clumping) or there are multiple plankton of different species scattered in an image (heterogeneous interspersion). They do not have localization or detection ability either.

In this paper, we propose a deep learning based

Received 18 May 2018, Revised 20 June 2018, Accepted 24 June 2018

\* Corresponding Author Youngsung Soh, Department of Information and Communication Engineering, Myongji University in Korea  
E-mail: soh@mju.ac.kr

method that can detect individual plankton in an image with homogeneous clumping, heterogeneous interspersions, or combination of both and recognize their classes. Our work was inspired by the system named “You Only Look Once (YOLO)” [1], a descendant of R-CNN family [2,3]. We applied the method to many real flow cytometer and microscope (FlowCam) [4] images and manmade examples and obtained almost perfect results.

## II. RELATED WORKS

Luo et al. [5] proposed an active learning approach where they used shadow image particle profiling evaluation recorder (SIPPER II) images. They first extracted 49 features including normal/weighted moment invariants, Fourier descriptors, and textural features and apply multi-class SVM for recognition. The number of samples tested was 8,440 of only 5 species and resulting recognition accuracy was 61%.

Hu et al. [6] used co-occurrence matrices as features obtained from images captured by video plankton recorder (VPR). SVM was used as a classifier to recognize 20,000 examples of 7 species. They reported that the accuracy was 72% which is better than [5].

Blaschko et al. [7] also extracted various features such as moment, texture, contour, and differential features from images captured by low resolution in situ FlowCam image sensor and used a single or ensemble classifiers for recognition. They tested 982 examples of 13 species and reported 53% to 73% accuracy depending on ensemble method.

Cowen et al. [8] used high resolution line scan camera system called ISIS to get images. They allow multiple organism present in single image. They used mixture of Gaussian based background models to detect objects. Next they extracted features including shape histogram, blob solidity, Hu moments up to third order, Fourier descriptors, and circular projection

descriptors. Finally multiclass SVM was applied for recognition. They tested 1,110 examples of 5 species and reported 61% to 97% accuracy depending on taxa.

Schulze et al. [9] used Sobel operator and histogram normalization for preprocessing and adopted region growing method to segment the plankton from images captured by FlowCam imaging sensor. Next they extracted features including elliptic Fourier descriptors, co-occurrence matrices, directionality histogram, moments, rotation invariant local binary pattern, etc. Finally two stage shallow neural networks were used for plankton/non-plankton separation and recognition. They used 1,418 plankton images of 10 species and reported 94.7% accuracy in average. This system was developed as an integrated system known as PlanktoVision.

There was a competition for plankton classification presented by National Data Science Bowl (NDSB) [10] from Dec. 2014 to Mar. 2015. More than 1,000 teams participated and the team named “Deep Sea” won the prize. There were 30,000 examples of 121 species and the task was to correctly classify them. Most top rankers adopted deep learning architectures and Deep Sea reported log loss of 0.595671.

Almost all the previous studies do not allow homogeneous clumping and heterogeneous interspersions that frequently happen in real environment. In the next section we propose a deep learning based method that overcomes these difficulties and produces high detection and recognition accuracies.

## III. THE PROPOSED METHOD

As was mentioned in Chapter 1, our work was inspired by YOLO [1] which is a descendant of R-CNN family [2,3]. YOLO was originally developed to detect and classify objects in image database such as ImageNet [11], Coco [12], etc. Later it was used in other applications including traffic sign detection [13].

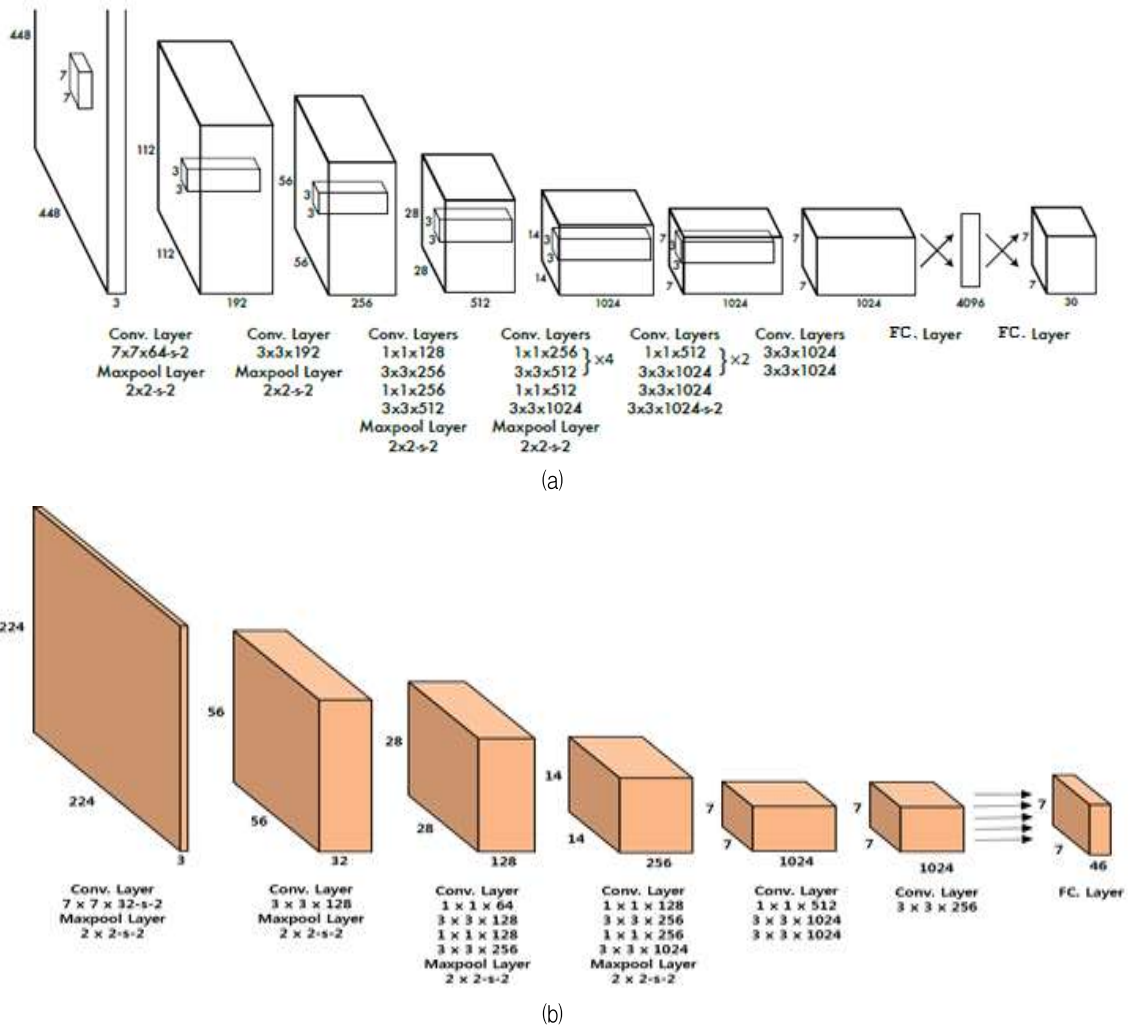
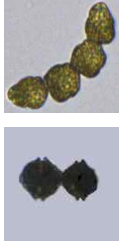



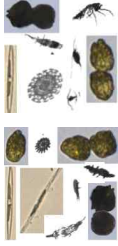


Fig. 1. Deep neural network architecture for plankton detection and recognition: (a) YOLO and (b) PROPOSED

A prominent advantage of YOLO over other deep learning based object detection scheme is execution speed with a bit more sacrifice on performance. However, since YOLO has too many unnecessary convolution layers for our application, we modify it to have a lot less convolution layers with smaller spatial dimension. Furthermore, we changed one of two fully connected layers with convolution layer to reduce the number of connections that makes training and testing faster. The original YOLO (YOLO) and the proposed network (PROPOSED) are depicted in Fig. 1(a) and 1(b) respectively.

Convolution layers (CONV) are responsible for feature learning, max pool layers (MP) for data size reduction and translation invariance, and fully connected layers (FC) for recognition. Among various convolution filter sizes, 1x1 is used to either increase or decrease size of data (feature maps). YOLO has 24 CONVs, 4 MPs, and 2 FCs, whereas PROPOSED has 14 CONVs, 4 MPs, and 1 FC. The numbers of CONV and FC are cut down to almost half which makes both training and testing a lot faster while maintaining the similar performance.

Table 1. The number of samples (training + validation) and example images in each category

Harmful species		Non-harmful species		Harmful + Non-harmful		Total
Multiple	Single	Multiple	Single	Multiple	Single	
90	47	20	772	20	0	949
						

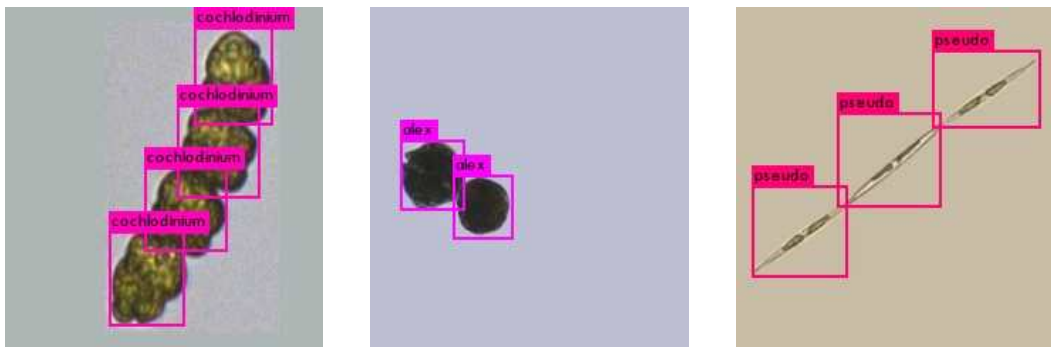
#### IV. EXPERIMENTAL RESULTS

We used total 949 images (samples) of 18 species. Out of 18, 3 are harmful species that cause red tide, blood poisoning, etc. and 15 are non-harmful ones. We divide samples into these two groups since harmful species tend to clump a lot more than non-harmful ones. Some of the images in both groups have only one plankton body while others have homogeneous clumping, heterogeneous interspersion, or combination of both. Since it was not easy to get FlowCam images of all categories, we combine what we already have to produce various heterogeneous interspersion cases. Table 1 shows the number of samples in each category and some example images of each. Among all images in the database, 80% were used for training and remaining 20% for validation. Before dividing them into training and validation sets, we extracted their ground truth locations and classes manually. Batch size of 32 was used and 8 times of data augmentation was done by applying 6 geometric

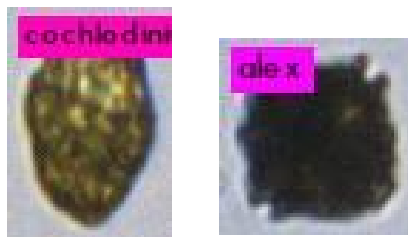
distortions and 2 pixel value modifications. Thus  $949 * 0.8 * 8 = 6,074$  images were used for training. There were  $949 * 0.2 = 189$  samples in validation set. Some samples have only one plankton while others have multiple plankton. The total number of plankton in all validation set images were 305.

For learning, we adopt leaky ReLU activation and batch normalization to prevent gradient vanishing and explosion. We use softmax function as a cost function. Learning was performed on PC with one i7 CPU and one NVIDIA GTX980 GPU board. It took around 20 hours for 20,000 epochs. For testing, the same H/W was used and the processing speed was around 50 fps.

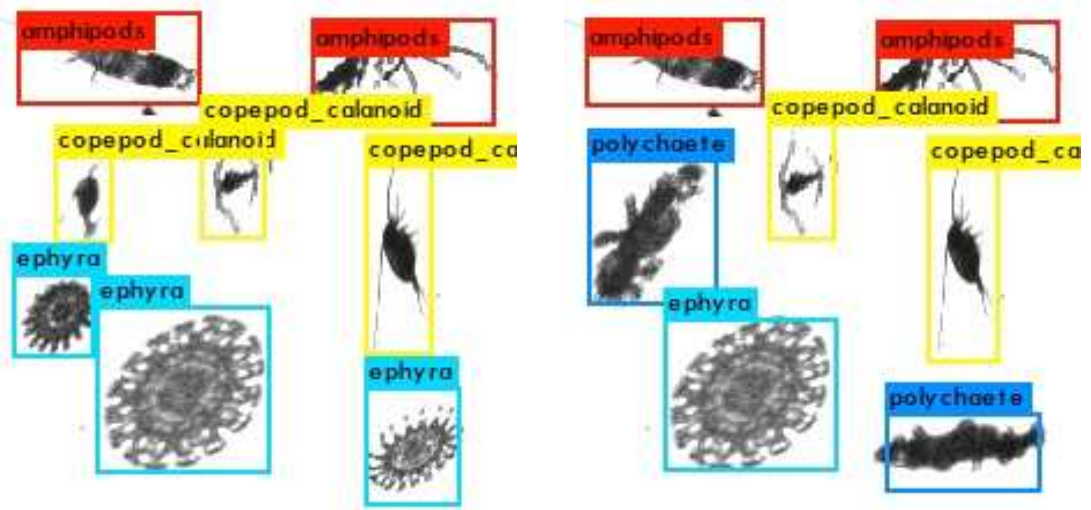
Fig. 2 shows some examples of plankton detection and recognition results. Fig. 2(a) is for Harmful species (Multiple), Fig. 2(b) for Harmful species (Single), Fig. 2(c) for Non-harmful species (Multiple), Fig. 2(d) for Non-harmful species (Single), and Fig. 2(e) for Harmful + Non-harmful (Multiple).



(a)



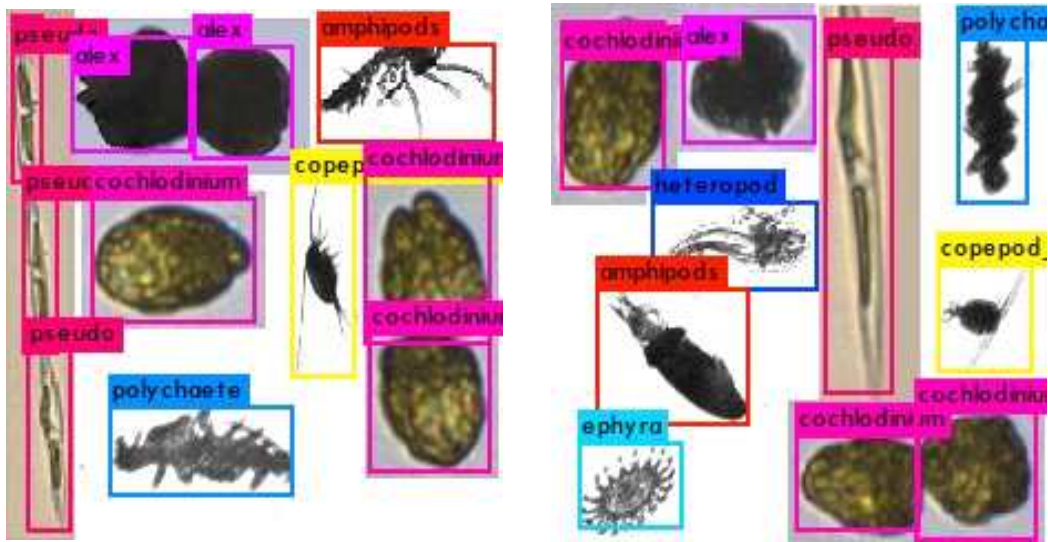
(b)



(c)



(d)



(e)

Fig. 2 Detection and classification results for each category: (a) Harmful species(multiple), (b) Harmful species(single), (c) Non-harmful species(multiple), (d) Non-harmful species(single), and (e) Harmful + Non-harmful(multiple)

Here boxes of various colors represent detection of plankton locations and the labels above them show recognition results. Our method detected all 305 plankton and recognized all of them right. To measure the detection accuracy, we computed IOU (Intersection of Union) between the detected and the ground truth boxes. The average IOU all over 305 detections was 0.86. This means that detected plankton location accuracy in average is 86%. While our method

detected all plankton, there was one false detection (false positive) in Harmful (Multiple) category and it is depicted in Fig. 3. Fig. 3(a) shows the original and Fig. 3(b) the detection result. The box indicated as “FP” in Fig. 3(b) is a false detection. There were 10 plankton named “pseudo-nitzchia” (marked as pseudo) in the image, but our method found 11 plankton.

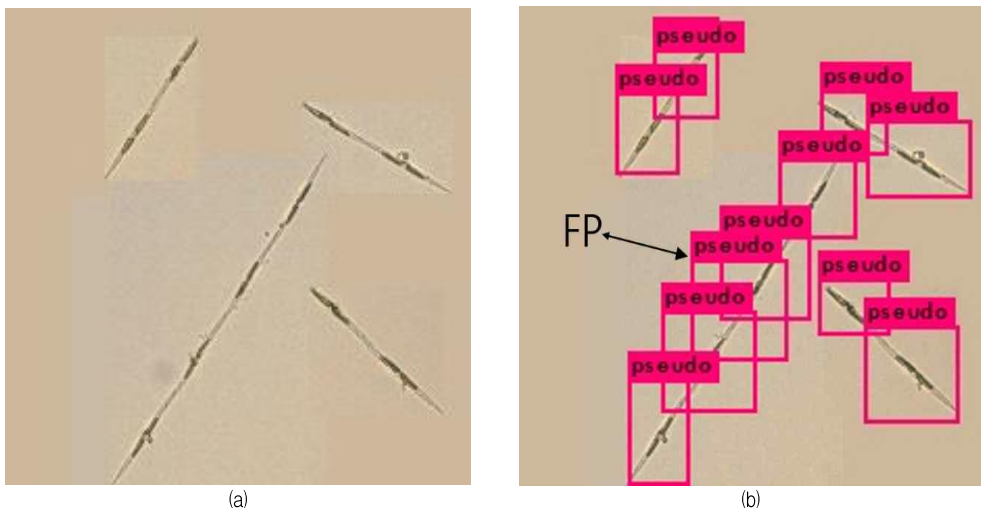


Fig. 3. False positive (FP) example: (a) Original image and (b) Detection results

## V. CONCLUSION

Distribution of harmful plankton species in ocean is an important indicative of marine ecosystem health. There were several deep learning based approaches that recognized plankton images. However almost all of them assume that there is only one plankton in an image and thus no detection is performed. In this paper we allow multiple plankton present in an image in clumped or scattered fashion and propose a fully automatic way of detecting and recognizing them. The proposed method produced 100% detection and recognition accuracies with 1 false positive in detection.

Currently the proposed method uses data set with 18 plankton species. We try to extend the number of species as we get more data and this is intended for future research.

## REFERENCES

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," arXiv preprint arXiv:1506.02640, 2015.
- [2] R. Girshick, "Fast R-CNN," *Proc. of IEEE International Conference on Computer Vision (ICCV)*, 2015.
- [3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," arXiv preprint arXiv:1506.01497, 2015.
- [4] <http://www.fluidimaging.com/>
- [5] T. Luo, K. Kramer, D. B. Goldgof, L. O. Hall, S. Samson, A. Rensen, and T. Hopkins, "Active Learning to Recognize Multiple Types of Plankton," *Journal of Machine Learning Research*, pp.589–613, vol. 6, 2005.
- [6] Q. Hu1 and C. Davis, "Automatic plankton image recognition with co-occurrence matrices and support vector machine," *Marine Ecology Progress Series*, Mar, pp.21–31, vol. 295, 2005.
- [7] M. B. Blaschko, G. Holness, M. A. Mattar, D. Lisin, P. E. Utgoff, A. R. Hanson, H. Schultz, and E. M. Riseman, "Automatic In Situ Identification of Plankton,"

*Proc. of Seventh IEEE Workshops on Application of Computer Vision*, pp.79-86, vol. 1, 2005.

- [8] R. K. Cowen\* and C. M. Guigand, "In situ ichthyoplankton imaging system (ISIS): system design and preliminary results," *Limnology and Oceanography: Methods*, pp.126-132, vol. 6, Issue, 22008.
- [9] K. Schulze, U. M. Tillich, T. Dandekar, and M. Frohme, "PlanktoVision - an automated analysis system for the identification of phytoplankton," *BMC Bioinformatics*, 2013:115, 2013.
- [10] <https://www.kaggle.com/c/datascienceowl>
- [11] <http://image-net.org/>
- [12] <http://mscoco.org/>
- [13] [http://cs231n.stanford.edu/reports2016/263\\_Report.pdf](http://cs231n.stanford.edu/reports2016/263_Report.pdf)

## AUTHOR INTRODUCTION

### Youngsung Soh



1978.2 : BS in electrical engineering from Seoul National University  
 1986.2 : MS in computer science from the University of South Carolina in USA  
 1989.2 : Ph.D in computer science from the University of South Carolina in USA  
 1989~1991 : Systems Engineering Research Institute in Korea as a senior researcher  
 1991 ~ : Professor in Myongji University in Korea  
 <Interest of research : object tracking, deep learning, parallel algorithms>

### Jaehyun Song



2003.2 : BS from Myongji University  
 2005.8 : MS in information and communication engineering from Myongji University  
 2006~2017 : Senior researcher in Nain information systems in Korea  
 2018~ : CEO of DIPS in Korea  
 <Interest of research : object tracking, deep learning, parallel algorithms>

### Yongsuk Hae



2009.2 : BS in Myongji University  
 2012.2 : MS in information and communication engineering from Myongji University  
 2012~2017 : Senior researcher in Nain information systems in Korea  
 2018~ : CTO of DIPS in Korea  
 <Interest of research : object tracking, stereo vision, deep learning, parallel algorithms>