


Peer-Reviewed Technical Communication

Vision-Based *In Situ* Monitoring of Plankton Size Spectra Via a Convolutional Neural Network

Nan Wang, *Member, IEEE*, Jia Yu, *Member, IEEE*, Biao Yang, *Member, IEEE*,
Haiyong Zheng , *Member, IEEE*, and Bing Zheng, *Member, IEEE*

Abstract—Plankton size spectra monitoring is crucial for managing and conserving aquatic ecosystems. Thus, we develop an *in situ* size spectra monitoring system to obtain the size spectra of plankton and the information of their living status underwater. The system consists of an imaging unit and an information processing unit. The imaging part applies a darkfield illumination to enhance the image contrast. Three lenses with different magnifications are alternated by a motor automatically to capture sizes of plankton from 3 μm to 3 mm. Moreover, the system can analyze the captured images in real time using the proposed multitask size spectra convolutional neural network, obtaining size spectra and density distribution of plankton simultaneously. Field test confirms that our system performs well both in imaging and information processing. Furthermore, the system can provide the living behavior of plankton, thereby helping biologists to study the aquatic ecosystem effectively and precisely.

Index Terms—Convolutional neural network (CNN), plankton monitoring, size spectra, underwater vision.

I. INTRODUCTION

PLANKTON are the floating communities of plants and animals that live in large bodies of water. These plankton typically free float with water currents [1]. In addition to representing the bottom few levels of the food chain that supports commercially important fisheries, plankton ecosystems play an indispensable role in the biogeochemical cycles of many important chemical elements, including the carbon and oxygen production cycles of oceans [2]. Plankton abundance and distribution strongly depend on factors, such as ambient nutrient concentrations, the physical state of the water column, and the abundance of other plankton. Thus, plankton are sensitive to natural and man-made factors [3]. Monitoring the status of the plankton in aquatic environments has substantial value. This process may not only

TABLE I
SIZE RANGE OF PLANKTON DIVISION

Group	Size Range
Megaplankton	$> 20\text{cm}$
Macroplankton	$2 \sim 20\text{cm}$
Mesoplankton	$0.2 \sim 20\text{mm}$
Microplankton	$20 \sim 200\mu\text{m}$
Nanoplankton	$2 \sim 20\mu\text{m}$
Picoplankton	$0.2 \sim 2\mu\text{m}$
Femtoplankton	$< 0.2\mu\text{m}$

help people to estimate climate change or the quality of water, but also contribute to the study of the ecological environment and life. However, correctly sampling and measuring the biomass and distribution of the plankton, especially *in situ* and in real time, remains a long-standing challenge [4], [5].

Essentially, plankton are defined by their ecological niche rather than any phylogenetic or taxonomic classification. The determination of the trophic level of plankton is not always straightforward. For example, many plankton are both photosynthetic producers and heterotrophic consumers [6]. Thus, plankton are frequently described in terms of size (as listed in Table I). The existence and importance of nano and even smaller plankton ($< 20\mu\text{m}$) were only discovered during the 1980 s. However, they are considered to make up the largest proportion of all plankton in number and diversity.

Traditionally, plankton are sampled by a customized sample net. Such a net is often a multiple opening and closing net. When necessary, a boat is equipped with it and towed up and down to obtain replicate samples from destination depth intervals. Several popular plankton samplers are multiple opening/closing net and environmental sensing system [7], Bedford Institute of Oceanography Net and Environmental Sensing System [8], MultiNet [9], and Gulf-V [10], to name a few. When the sample is collected by the sampler, a standard set of procedures, including separation, taxonomic analyses, and dry, is processed by researchers. The procedures are always carried out in the laboratory or on the deck of a boat [11].

Plankton sampling and analysis are both time consuming and expensive. The collection of relatively complete samples from a sea zone always requires several days of work on a boat, followed by months of sample processing and analysis. Moreover, considerable active information, such as the living status of plankton, is lost because the sample analysis is not performed *in situ*. Furthermore, given the extrusion and

Manuscript received March 6, 2018; revised August 6, 2018 and October 25, 2018; accepted November 7, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61703381 and Grant 41776113, in part by the Natural Science Foundation of Shandong Province under Grant ZR2017BF006, in part by the China Postdoctoral Science Foundation under Grant 2016M590658, and in part by the Fundamental Research Funds for the Central Universities under Grant 201713017. (*Corresponding author: Haiyong Zheng.*)

Associate Editor: B. Thornton.

N. Wang, H. Zheng, and B. Zheng are with the Department of Electronic Engineering, College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China (e-mail: wangnan@ouc.edu.cn; zhenghaiyong@ouc.edu.cn; bingzh@ouc.edu.cn).

J. Yu is with the Department of Physics, College of Information Science and Engineering, Ocean University of China, Qingdao 266100 China (e-mail: yujia2008@ouc.edu.cn).

B. Yang is with the School of Information Science and Engineering, Changzhou University, Changzhou 213164, China (e-mail: yb6864171@cczu.edu.cn).

Digital Object Identifier 10.1109/JOE.2018.2881387

avoidance of significantly small organisms regardless of the small mesh size in the sampler net, data obtained by the traditional sample method can not fully reflect the exact ecological situation [12].

Due to the shortcoming of towed sampling and the necessity to monitor the plankton *in situ*, automatic or semiautomatic plankton monitoring systems have been urgently developed [13]. The most remarkable one is optical plankton counter (OPC), originally designed at the Bedford Institute of Oceanography, Dartmouth, NS, Canada, as a remotely towed sensor providing continuous real-time information on the size and abundance of zooplankton [14], [15]. The OPC has been deployed on various platforms and utilized in numerous water investigations. However, the OPC has several limitations. The most dominant one is maintaining a significant probability of two or more particles that are present in the beam simultaneously and be counted as a big particle when the plankton density is large [16]. Herman developed a new generation of the OPC, that is, the laser OPC (LOPC) [17], to solve the limitations of the original OPC. The LOPC consists of a laser diode, a cylindrical lens, and a central processing unit. The basic principle of measurement is unchanged though. Both OPC and LOPC mainly aim at counting the plankton numbers and measuring the size. They are not imaging systems that provide visual information.

To obtain more information about the living status of plankton, several *in situ* imaging systems are developed, such as video plankton recorder (VPR) [18], underwater video profiler [19], shadowed image particle profiling evaluation recorder (SIPPER) [20], zooplankton visualization system (ZOOVIS) [21], *in situ* ichthyoplankton imaging system (ISIS) [22], scripps plankton camera (SPC) [23], electronic holographic camera [24], FlowCAM [25], and Imaging FlowCytobot (IFCB) [26]. The VPR is highlighted for its high motility and multi-functions, and the objective plankton size range is $100\ \mu\text{m}$ – $1\ \text{cm}$. The SPC is another notable plankton imaging system, which uses darkfield illumination to enhance the contrast of transparent objects. The development of the SPC processes several generations. Millions of images of plankton have been captured by SPC (or the SPC2) and have been published on the website: <http://spc.ucsd.edu>. Several new species of plankton have been recognized with the help of the SPC.

Most aforementioned plankton imaging systems provide the corresponding processing program, such as plankton extraction, identification, and classification [13]. Tang *et al.* [27] designed an automatic recognition system combining moment invariants and Fourier descriptor with granulometric features using a learning vector quantization neural network to classify plankton images detected by VPR. Luo *et al.* [28] presented a system to recognize SIPPER underwater plankton images by combining moment invariants and granulometric features with certain specific features (e.g., size and convex ratio) via active learning combined with support vector machine. Hu and Davis [29] developed a dual-classification method in which each VPR plankton image is identified first using a shape-based feature set and a neural network classifier, and then using a texture-based feature set and a support vector machine classifier. Zhao *et al.* [30] classified the binary SIPPER plankton image via random subspace and verified that the combining multiple stable classifiers are better than a single classifier. Sosik and Olson [31] developed an approach for IFCB plankton image classification, including feature extraction, feature selection, and support vector machine classifier. Vandromme *et al.* [32] accessed the biases in computing the size spectra of automatically classified zooplankton from a ZooScan integrated system in the laboratory. Li *et al.* [33] proposed a pairwise nonparametric discriminant analysis for recognizing a SIPPER binary plankton image. Bi *et al.* [34] developed a semiautomated method to analyze plankton taxa from images acquired by ZOOVIS. Faillettaz *et al.* [35] postprocessed a computer-generated classifica-

tion for images collected by the ISIS using random forest to describe plankton distribution patterns. In addition, automatic identification algorithms also attract more attention from the plankton analysis and computer vision interdisciplinary researchers, such as ADIAC [36], DiCANN [37], and so on [38]–[41].

However, nearly all the preceding processing programs work independently from the imaging instrument, the underwater part takes series of pictures, and then the automated plankton image processing softwares analyze the images afterward on a computer (on board or in the lab). So technically, the aforementioned plankton imaging systems are more like plankton cameras than complicated digital instruments. Moreover, biologists believe that the size of planktonic organisms and the relationship between their abundance and the size, rather than their taxonomy, are more appropriate to indicate the situation of the ecological environment [42], [43]. Academically, the relationship between abundance and size is called size spectra. The analysis of the slopes of size spectra is widely used at present to assess the state of marine ecosystems at regional and global scales [44]. Observed size spectra typically become steep (negative) following exploitation (mainly of fishes). Several researchers have also studied the relationship between size spectra and climate change [45]. Given these reasons, considerable plankton monitoring equipment provides the function to obtain the size and abundance information. However, nearly all the vision-based systems first classify the plankton and then measure the size of every plankton individually. Recognizing all kinds of plankton is nearly impossible because plankton are variously oriented in three dimensions and live in different shapes. Thus, the method is not only time consuming but also inaccurate. Furthermore, the existing plankton imaging systems aim at achieving the complete picture of individual planktonic organism to help the classification. The dedicated image processors scan for the region of interested image contained objects that are large enough for identification. This is an effective effort to reduce workload. However, the active living information of the biological communities, which need to be reflected by the whole picture, is lost.

Further, most of the existing underwater imaging systems use composite cable to transmit the images and supply power. Limited by the transmission speed, the quantity and quality of the captured images must be reduced. With the development of embedded computing unit, *in situ* information processing comes true and becomes an effective method nowadays. Images can be processed on site and real time, thus, more useful information can be well handled timely, leaving only the processed valid data to be transmitted.

Therefore, according to the oceanography need and the development of technology, a new monitoring system that can measure the size spectra and record the community behavior is required. In this paper, we present an *in situ* plankton monitoring system. Different from the existing systems, the proposed system focuses on detecting the biomass and distribution accurately, not via the traditional plankton classification procedure, but using a multitask size spectra convolutional neural network (MSCNN). The imaging unit and an embedded computation unit are combined to monitor the plankton community and analyze the information intelligently in real time. Our *in situ* monitor can record the size spectra of aquatic biota continuously, and obtain the density map simultaneously. A darkfield multiple magnifications microscope system is designed to capture plankton images, ranging from nanoplankton to small mesoplankton ($3\ \mu\text{m}$ \rightarrow $3\ \text{mm}$).

The remainder of this paper is organized as follows. In Section II, the main schematic of our *in situ* plankton size spectra monitoring is introduced. The MSCNN is explained in Section III. Implementation details and experimental results are given in Section IV. Section V presents the conclusion and discussion.

TABLE II
SPECIFIC PARAMETERS OF MICROSCOPIC IMAGING LENSES

Item	Large format tele-centric lens	2× microscope objective	10× microscope objective
Magnification	0.9×	2×	10×
Numerical Aperture	0.045	0.055	0.42
Working Distance (<i>mm</i>)	111	34	15
Field of View (<i>mm</i> × <i>mm</i>)	3.1 × 3.1	2.2 × 2.2	1 × 1
Focal Length (<i>mm</i>)	-	100	20
Resolving Power (μm)	7	5	0.6

TABLE III
SPECIFIC PARAMETERS OF THE CAMERA

Item	Parameter
Pixels	2048 × 2048
Pixel Size (μm)	5.5 × 5.5
Pixel Depth	12bit
Frame Rate (fps)	75
Video Output	USB 3.0
Resolution (MegaPixels)	4.2

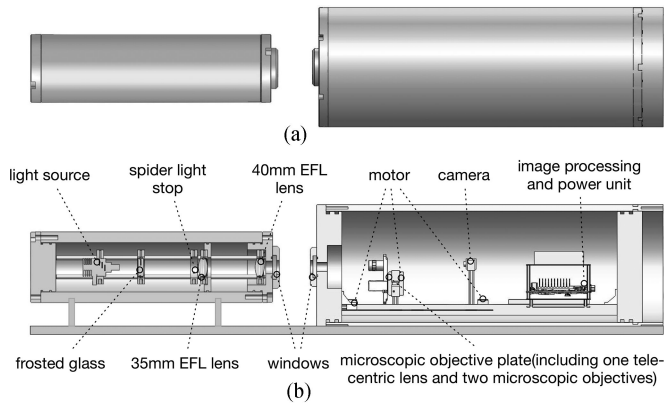


Fig. 1. Design structure of our *in situ* plankton size spectra monitoring system. (a) Top view. (b) Internal structure.

II. *In Situ* PLANKTON SIZE SPECTRA MONITORING SYSTEM

All the components of our *in situ* plankton size spectra monitoring system are housed within a cylindrical-shaped hull, which can be further divided into two concentric cabins, as illustrated in Fig. 1. The left cabin houses the light source, whereas the right cabin contains the imaging lenses and information processing unit. A power unit is also mounted on the right cabin. Simultaneously, a multiple-function interface is set at the bottom of the whole system to facilitate the data communication and power connection between the system and a shore-based station. The entire apparatus is 97 cm long. Here, we introduce our system from the perspective of function, including the imaging subsystem (unit) and the information processing subsystem (unit).

A. Imaging Subsystem

The imaging system uses darkfield illumination, formed by a white color LED and two aspherical achromatic lenses, to enhance the contrast of transparent objects. The imaging unit is more complicated. As previously mentioned, the plankton size varies over a wide range. Thus, we design an auto-modified multiple magnification imaging system to cover the size range of $3 \mu\text{m} \rightarrow 3 \text{ mm}$, thereby indicating that the proposed system can monitor from nanoplankton to small mesoplankton. Three magnifications, that are 0.9×, 2×, and 10×, are obtained: wherein 0.9× is achieved via a large format telecentric lens, which can supply a large field view, whereas 2× and 10× microscope objectives are used to realize the corresponding magnifications for microscopic imaging. The detailed informa-

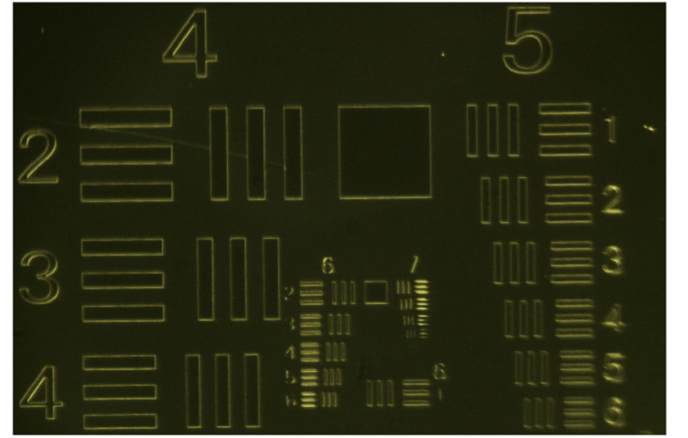


Fig. 2. 1951 USAF resolution test chart captured underwater by the proposed system.

tion of the three lenses and the camera is listed in Tables II and III, respectively.

To indicate the resolving power, a 1951 USAF resolution test chart, which is conformed to the MIL-STD-150 A standard, is applied as a shooting target. Fig. 2 depicts the captured image. The scales and dimensions of the bars are listed in Table IV. Clearly, the fourth element of Group 7 is captured well, showing the resolution of the proposed system down to $3 \mu\text{m}$.

Specially, the three objective lenses are set on a disk, which can be rotated by a motor. The motor enables the three lenses to shift alternatively. Moreover, a location sensor is used to position every lens appropriately at the center of the light path in each turn. Two

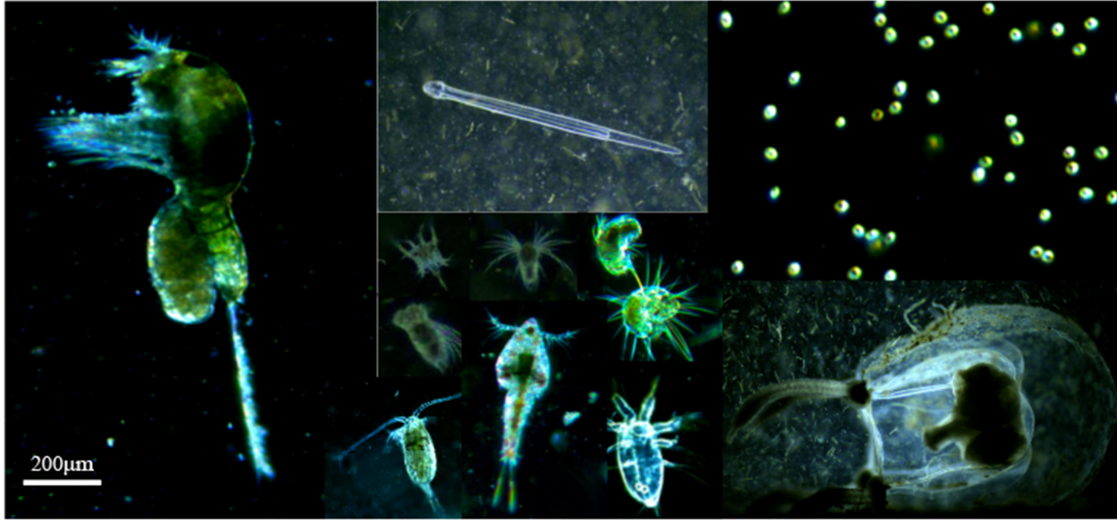


Fig. 3. Samples of plankton captured by the proposed system.

TABLE IV
WIDTH OF ONE LINE IN MICROMETERS IN USAF RESOLVING
POWER TEST TARGET 1951

Element	Group Number			
	4	5	6	7
1	31.25	15.36	7.81	3.91
2	27.84	13.92	6.96	3.48
3	24.80	12.40	6.20	3.10
4	22.10	11.05	5.52	2.76
5	19.69	9.84	4.92	NA
6	17.54	8.77	4.38	NA

extra motors are used to change the location of the objective lens and camera because different magnifications require various imaging distances. The three motors and camera are controlled by an embedded Advanced RISC Machine platform. It is set to take pictures every 10 min, and three magnifications at each run. The interval time and number of images can be set according to practical needs. Thus, the subsequent information processing unit has abundant time to process the images. Pictures of plankton captured by the system are shown in Fig. 3.

B. Information Processing Subsystem

The schematic procedure of the information processing subsystem is shown in Fig. 4. The quality of the captured images remain low given the underwater light attenuation and suspended substance, although we use darkfield illumination to enhance the outline of the plankton. Thus, the image enhancement step is necessary. Illumination normalization and edge enhancement are realized in this subsystem. The system mainly aims to monitor the biomass size spectra of the ambient environment but not the plankton classification. Thus, we propose the MSCNN to detect the individual shape of every plankton to count the biomass size spectra. Moreover, the density distribution of the ambient plankton is achieved to show the real living status of the community. Furthermore, a judgment unit is added to monitor whether the biomass size spectra

is normal. If something abnormal occurred (which indicates that the distribution of the biomass size spectra has changed significantly), then the frame is recorded, and an abnormal flag will be set to facilitate the further exploration or detection.

III. MULTITASK SIZE SPECTRA CONVOLUTIONAL NEURAL NETWORK

A. Biomass Size Spectra

Empiricists and modelers currently adopt multiple definitions to calculate the biomass size spectra explicitly. The difference is shown in various combinations of normalized or nonnormalized biomass distribution B , or abundance N , or energy E , versus individual length l , weight w , or volume v , in a log-log space. Typically, 2 is applied as the base, which means each size class width is doubled with respect to the previous one. According to different regression modes, the size spectra can be plotted as linear or dynamic one.

In this research, we perform several simplifications because we must obtain the size spectra from the image (in a 2-D form). The numbers of different size classes are counted as N , and each particle is handled as ellipses shape to calculate the volume. The continued frames are used to refine the size spectra. Thus, the size spectra can be recognized as the ratio between N and v .

B. Multitask Size Spectra Convolutional Neural Network

Two coefficients, namely the number and the shape of the plankton, should be computed to obtain the size spectra of plankton. Traditionally, these two goals can be realized by object detection method based on handcrafted features [46] or saliency [47], [48]. Recently, deep convolutional neural network (CNN) have achieved great development in detection field [49]. After trained by large-scale data, CNN yields a really high accuracy [50]. As CNNs have proved their efficiencies in natural language processing and computer vision fields, researchers explore the method on many real applications [51]–[53]. A kind of successful application is applying CNN to count the density map of a crowd to prevent the extreme event from occurring. For example, Zhang *et al.* [54] proposed a multicolumn CNN, which stacks the feature maps generated by filters of different sizes and combines them to generate the final number prediction; Oñoro-Rubio *et al.* [55] presented two models, Counting CNN and Hydra CNN, to count object instances

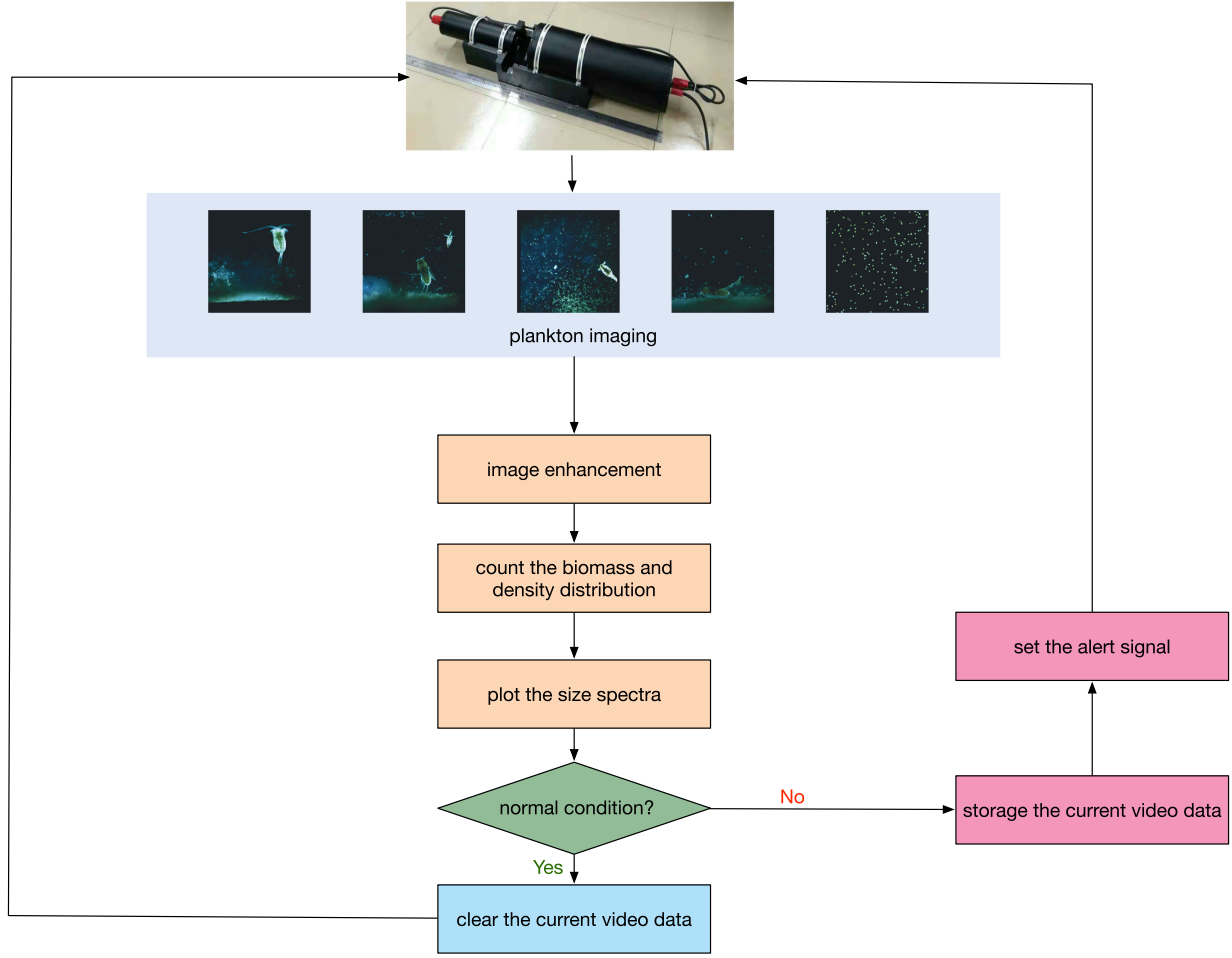


Fig. 4. Main schematic of the information processing subsystem.

in images, and no geometric information of the scene must be provided to the Hydra CNN. Both of them reported high prediction accuracy.

Inspired by above-mentioned works, we design the MSCNN to obtain the plankton detection and observe the living status simultaneously, and Fig. 5 shows the corresponding model. The proposed MSCNN is a dual-task network. The first function of the work is to detect the shape of each plankton, and this information will be used in the future to calculate the size spectra. The other function is to obtain the density map of plankton. In contrast to the detection result, the density map shows the distribution situation of the plankton (e.g., whether these plankton are inclined to cluster together or just wandering in solitude). This information may help biologists in studying the behavior of the plankton.

Fig. 5 presents the three parts of the network. The first part (blue color) consists of three convolutional layers. Conv1 and Conv2 layers have filters of size 7×7 with a depth of 32, and the stride and pad are 1 and 3, respectively. Both are followed by a max-pooling layer, with a 2×2 kernel size. The Conv3 layer has 5×5 filters with a depth of 64. All the convolutional layers are followed by a rectified linear unit. The three layers are used to learn the low-level features.

The second part (green color) is composed of two deconvolutional layers and a softmax layer. The deconvolutional layers are used to upsample the output of the former layer. The kernel size and stride of both the two deconvolutional layers are set to 2. The output of the softmax layer is the separate result of background and foreground. In this case, the output is the plankton detection result.

The third part (pink color) also consists of three convolutional layers, with the depth of 1000, 400, and 1, respectively. The stride and kernel size of these layers are all 1. In this way, we use a fully convolutional network, which allows the different sizes of the input image. The output of the last layer is the density map.

In consideration of the compatibility of the computing unit (NVIDIA Jetson TX2), we constrain the input size to 72×72 , although the network allows any input size. The original image is divided into several patches. Each patch passes forward the network separately, and they are combined together to obtain the entire result image.

C. Loss Function for Multitask

The proposed MSCNN is trained in the manner of multitask learning. The loss between the estimated density map and its ground truth is calculated using the Euclidean loss, defined as L_{density}

$$L_{\text{density}} = \frac{1}{2N_u} \sum_{i=1}^{N_u} \|F_{\text{den}}(P_i, O) - D_1(P_i)\|_2 \quad (1)$$

where N_u is the number of training samples, and O is a set of network parameters. P_i is the i th patch, $F_{\text{den}}(P_i, O)$ is the estimated density map of P_i , and $D_1(P_i)$ is the ground truth of $F_{\text{den}}(P_i, O)$.

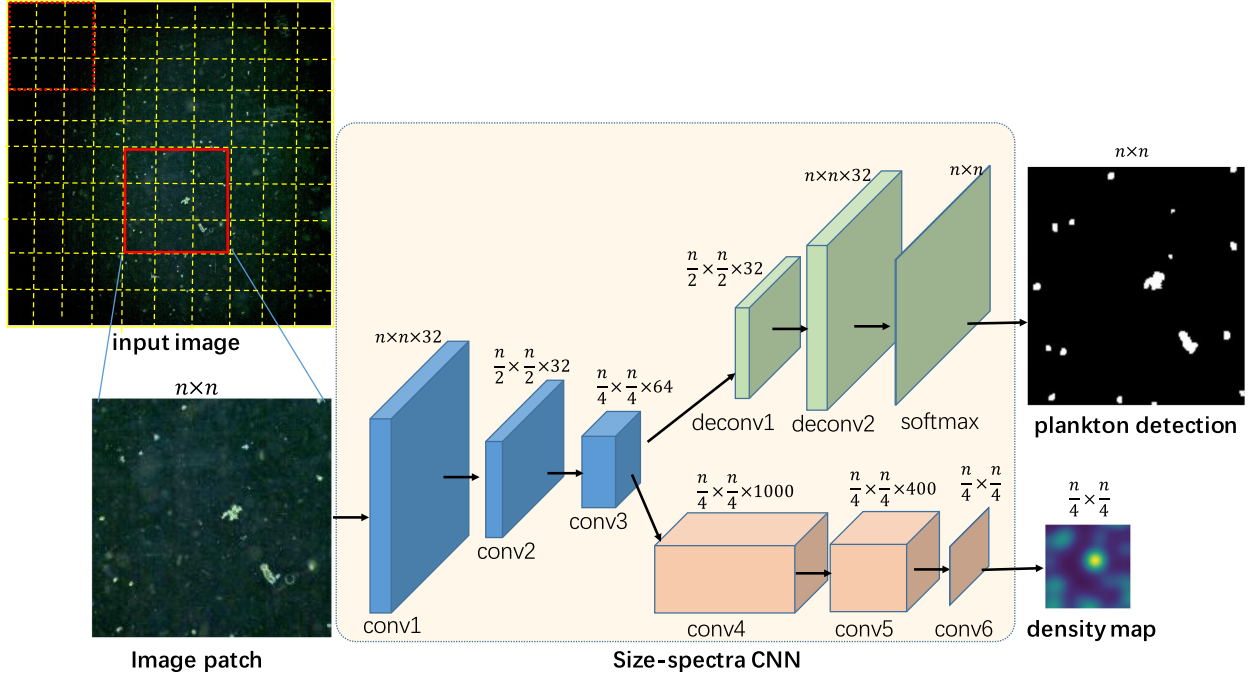


Fig. 5. Proposed MSCNN model.

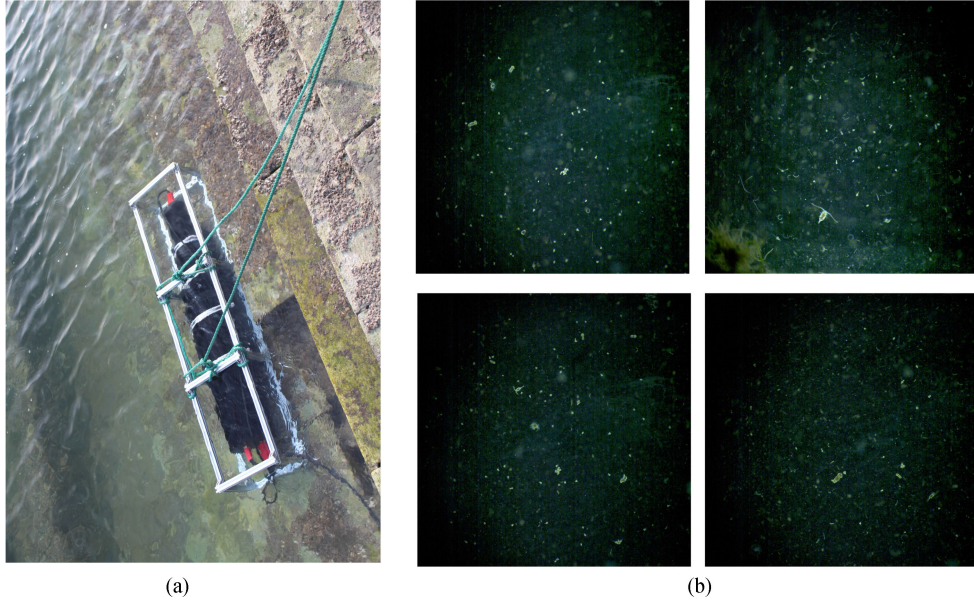


Fig. 6. Preliminary experiment for data collection. (a) Offshore test. (b) Several selected training images.

Meanwhile, the Euclidean loss between plankton detection result and the ground truth is defined as $L_{\text{detection}}$

$$L_{\text{detection}} = \frac{1}{2N_u} \sum_{i=1}^{N_u} \|F_{\text{det}}(P_i, O) - D_2(P_i)\|_2 \quad (2)$$

where $F_{\text{det}}(P_i, O)$ is the predicated detection result of P_i , and $D_2(P_i)$ is the ground truth of $F_{\text{det}}(P_i, O)$.

Finally, the total loss of the MSCNN is defined as follows:

$$L_{\text{total}} = \lambda_1 L_{\text{density}} + \lambda_2 L_{\text{detection}}$$

where λ_1 and λ_2 are weights to balance the contribution of the two loss functions. In this paper, we set $\lambda_1 = 0.3$ and $\lambda_2 = 0.7$, for we mainly aim to get the plankton detection information precisely.

IV. EXPERIMENTAL RESULTS

A. Data Set

To collect abundant training data, we apply the plankton monitoring system in waters for several days. Many underwater videos are collected. Then, the videos are screened manually to select the frames that

contain different kinds and situations of plankton as the training data. In this application, the number of remaining frames is 5964. Fig. 6 shows the experimental field and several selected images.

The remaining images are then annotated as follows by three people: a single red dot marked at the center and a white area labeled for the shape of each plankton. The labeled white ground truth is used to train the plankton detection task. To get the ground truth density map, for an image I , $D_{1,I}$ is defined as a sum of Gaussian functions centered on each red dot annotation

$$D_{1,I}(p) = \sum_{p \in A_I} N(p|\mu, \delta) \quad (4)$$

where A_I is the set of 2-D points annotated for the image I , and $N(p|\mu, \delta)$ represents the evaluation of a normalized 2-D Gaussian function, with μ and δ representing the mean and variance, respectively, at the coordinate defined by p . In this paper, δ is set to 15 empirically and μ is zero. Further, the total object count N_I can be directly computed by integrating the density map values in $D_{1,I}$ over the entire image as follows:

$$N_I = \sum_{p \in I} D_{1,I}(p) \quad (5)$$

where N_I can be further used to rectify the counting number of the plankton detection.

B. Training and Testing

Training and evaluation of the MSCNN are performed on a workstation with two NVIDIA GTX 1070 Graphics Processing Units (8G) using the Caffe framework and a cuDNN library. The pretrained model is then transferred to the embedded NVIDIA Jetson TX2 for an underwater *in situ* detection task.

In the training procedure, we followed the experimental setup proposed in [55], by using only the training and validation sets for learning. For each original input image, we randomly extract 100 patches with the size of 115×115 , and then resize them to 72×72 for feeding them into the network.

C. Subjective Evaluation

Fig. 7 depicts several original underwater images captured by our imaging system. Group (a) and (b) images reflect the situation of a fresh pool, and the original image in Group (c) is captured in the nearby river of the aforementioned pool. To illustrate the performance of MSCNN, the ground truths of the plankton detection and distribution density map are displayed together with the predicated result of the MSCNN. We compare the result with the traditional adaptive thresholding method because no other similar network-based plankton detection work is available.

Plenty of plankton live in the field as illustrated in all the original images, whereas the small-sized plankton are more than the large-sized plankton. Moreover, constantly massive suspended particles make the plankton detection mission a difficult one. This “marine snow” from the plankton even by human eyes is hard to separate. However, with several deep CNNs, the MSCNN works well in the background and foreground separation. It is clear that our method detects most of the plankton from the background, even those in the dark area. By contrast, the adaptive thresholding method only successfully detects the apparent plankton, thereby omitting the plankton in the dim area and not very bright. Notably, an unknown kind of plankton, with a chain shape (marked in the yellow box), is presented in the images of Group (c). Considering that this kind is not labeled in the learning data set, the corresponding detec-

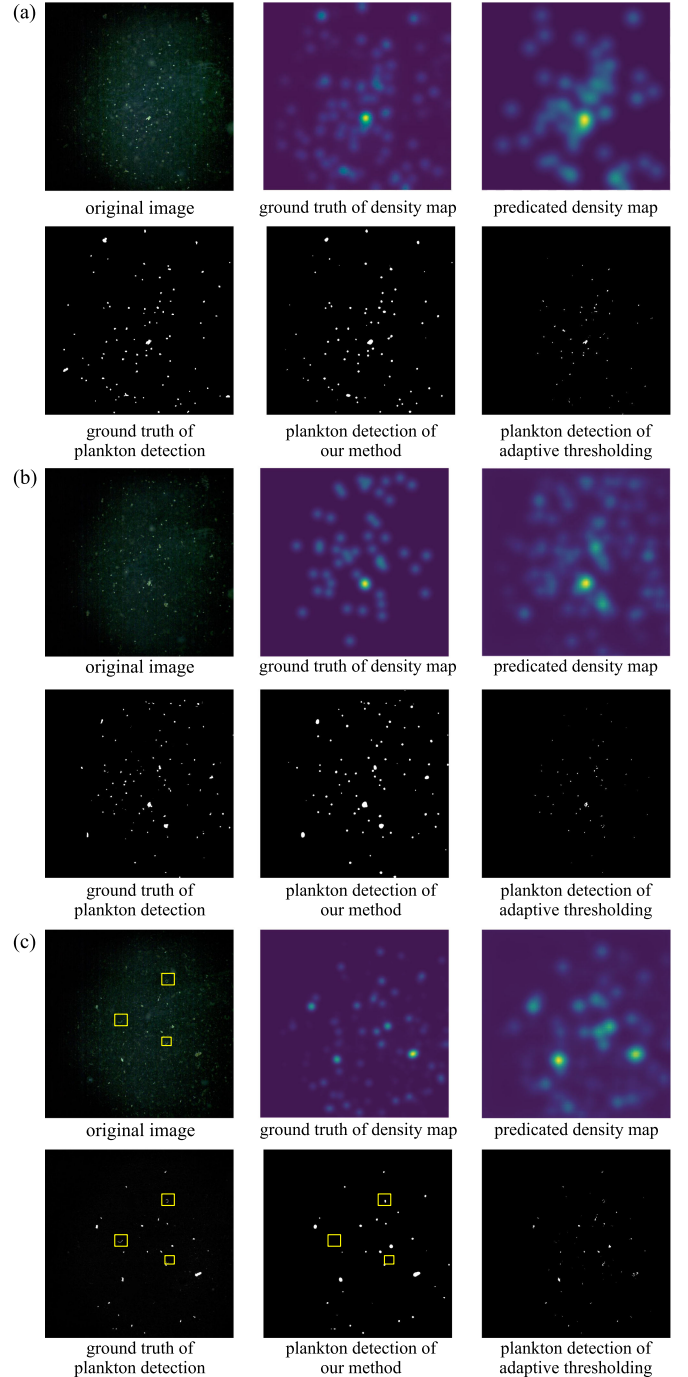


Fig. 7. Results of the proposed MSCNN model and adaptive thresholding method.

tion result omits such kind of plankton. This indicates the importance of the abundant data set for deep learning.

Moreover, the density map clearly shows the living status and trend of the plankton. The middle of the first two group of images has some food, which may be preferred by small-sized plankton. Thus, the small-sized plankton are clustered around the food. This behavior is well indicated by the density map (the cluster center is most brilliant in the density map), but cannot be illustrated by the detection result. No food is presented in the images of Group (c), so few plankton are imaged in this field. Furthermore, we can deduce that the plankton are just

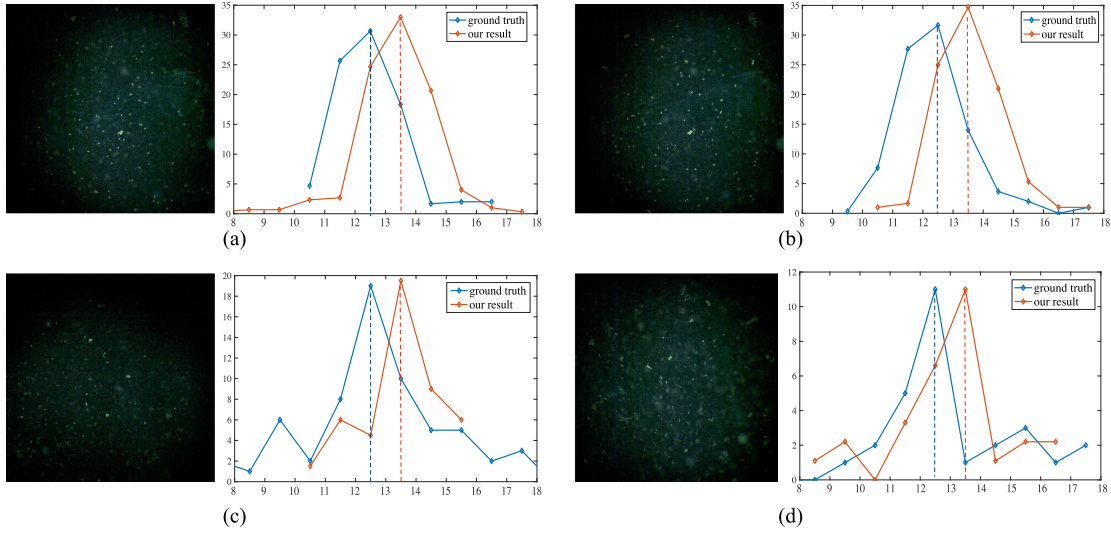


Fig. 8. Ground truth and plankton size spectra calculated by the proposed system at different times of the same plankton group. The x -axis and y -axis indicate the volume of the plankton and the counting number of the plankton that belong to such size, respectively. The interval among (a), (b), (c), and (d) is 30 min.

TABLE V
OBJECTIVE EVALUATION OF THE PLANKTON DETECTION

Method	MAE	MSE	PSNR
Adaptive thresholding	1.43	1.49	46.41
MSCNN	0.53	0.62	51.78

wandering around with the help of the density map, because there is no clustering center.

D. Objective Evaluation

An accurate detection of the plankton is the key point of the work, which will impact the calculation of the biomass size spectra directly. We evaluate the work with three metrics to adjust the accuracy of the plankton detection more precisely. These metrics are mean absolute error (MAE), mean square error (MSE), and peak signal-to-noise ratio (PSNR), which are commonly used for object counting [54], [56], [57]. They are defined as follows:

$$\text{MAE} = \frac{1}{W} \sum_{I=1}^W |D^d - D^{\text{gt}}| \quad (6)$$

$$\text{MSE} = \frac{1}{W} \sum_{I=1}^W (D^d - D^{\text{gt}})^2 \quad (7)$$

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}} \right) \quad (8)$$

where W is the total number of testing frames, and D^d and D^{gt} are the predicated and ground truth images, respectively. All the three metrics indicate the difference between the ground truth and the predicated results. Small MAE and MSE and large PSNR indicate an improved performance.

In our practice, we randomly select 100 pictures to calculate the three measures. The results are summarized in Table V. Both the MAE and MSE measures of our results are lower than 1 and nearly one-third of

the traditional adaptive thresholding method. If we combine this index with the subjective result shown in Fig. 7, then we can deduce that the difference between the result of the MSCNN and the ground truth is caused because the detection result of the MSCNN is slightly plump. A deeper network, which is our further work, may solve the problem.

E. Size Spectra

Once plankton are correctly detected by the MSCNN, we can use such information to count and compute the final size spectra. Furthermore, in consideration of the shooting angle, the same plankton is variable in different images. For example, a long ellipse-shaped plankton may be small when it is vertical to the camera or large when it is parallel to the camera. Thus, we calculate the average of the three sequential images to plot the size spectra. Moreover, the three sequential images are captured by the three different magnification lenses. Resizing is preformed to normalize the resolution.

Fig. 8 exhibits a series of calculated size spectra of the fresh pool, calculating every half hour in the same afternoon. The plankton, with the size of approximately $12.5 (\log_2(\mu\text{m}^3))$ and the body length of about $22 \mu\text{m}$, are predominant in the ecosystem. In the first two images, some food is found in the center. Thus, the overall biomass is large. We can find in the corresponding size spectra figure that the curve is plump. Afterward, the biomass has decreased when the food is eaten up. Few plankton are captured in the view. The corresponding size spectra figure is steep up and down.

Notably, the size spectra calculated via our method is slightly larger than the ground truth. A one-size shift to the right from the ground truth is observed. The analysis of the detection result of the MSCNN (see Fig. 7) may provide the truth, which is that the predicted plankton shape is slightly plump. However, we can easily rectify the size spectra in practice.

V. CONCLUSION AND DISCUSSION

In this paper, an *in situ* plankton monitoring system is proposed. The system contains an imaging unit and an information processing unit. Darkfield illumination is applied to enhance the image contrast, and an automodified multiple magnification imaging system is designed to capture the image of a wide range of different-sized plankton. The

images are then subjected to the MSCNN to detect individual plankton and density distribution.

The captured images confirm the effectiveness of the darkfield illumination and multiple magnifications. The size spectra figures for the fresh pool are presented to show the performance of the whole system. In addition, the introduction of the distribution density provides vivid information on the plankton community, which could be very helpful in studying the behavior of aquatic organisms. However, the accuracy relies heavily on the abundant data set because the processing method is based on the CNN. Additional images must be collected to improve the performance. That is, using added data indicates an improved performance of this system.

REFERENCES

- [1] C. Saret, *Plankton: Wonders of the Drifting World*. Chicago, IL, USA: Univ. Chicago Press, 2015.
- [2] C. B. Miller and P. A. Wheeler, *Biological Oceanography*. Hoboken, NJ, USA: Wiley, 2012.
- [3] P. Falkowski, "The power of plankton," *Nature*, vol. 483, no. Suppl. 7387, pp. S17–S20, 2012.
- [4] S. C. Doney, "Oceanography: Plankton in a warmer world," *Nature*, vol. 444, no. 7120, pp. 695–696, 2006.
- [5] P. Wang, X. Tian, T. Peng, and Y. Luo, "A review of the state-of-the-art developments in the field monitoring of offshore structures," *Ocean Eng.*, vol. 147, pp. 148–164, 2018.
- [6] G. Wang, X. Wang, X. Liu, and Q. Li, "Diversity and biogeochemical function of planktonic fungi in the ocean," *Prog. Mol. Subcellular Biol.*, vol. 53, no. 53, pp. 71–88, 2012.
- [7] P. H. Wiebe *et al.*, "New development in the MOCNESS, an apparatus for sampling zooplankton and micronekton," *Mar. Biol.*, vol. 87, no. 3, pp. 313–323, 1985.
- [8] D. D. Sameoto, L. O. Jaroszynski, and W. B. Fraser, "BIONESE, a new design in multiple net zooplankton samplers," *Can. J. Fisheries Aquatic Sci.*, vol. 37, no. 4, pp. 722–724, 1980.
- [9] H. Weikert and H. John, "Experiences with a modified bé multiple opening-closing plankton net," *J. Plankton Res.*, vol. 3, no. 2, pp. 167–176, 1981.
- [10] D. Sameoto *et al.*, "Collecting zooplankton," *ICES Zooplankton Methodology Manual*, vol. 40, no. 2, pp. 55–81, 2000.
- [11] H. R. Skjoldal, P. H. Wiebe, L. Postel, T. Knutsen, S. Kaartvedt, and D. D. Sameoto, "Intercomparison of zooplankton (net) sampling systems: Results from the ICES/GLOBEC sea-going workshop," *Prog. Oceanogr.*, vol. 108, no. 1, pp. 1–42, 2013.
- [12] N. MacLeod, M. Benfield, and P. Culverhouse, "Time to automate identification," *Nature*, vol. 467, no. 7312, pp. 154–155, 2010.
- [13] M. C. Benfield *et al.*, "RAPID: Research on automated plankton identification," *Oceanography*, vol. 20, no. 2, pp. 172–187, 2007.
- [14] A. W. Herman, "Simultaneous measurement of zooplankton and light attenuation with a new optical plankton counter," *Continental Shelf Res.*, vol. 8, no. 2, pp. 205–221, 1988.
- [15] A. W. Herman, "Design and calibration of a new optical plankton counter capable of sizing small zooplankton," *Deep Sea Res., A Oceanogr. Res. Papers*, vol. 39, no. 3–4A, pp. 395–415, 1992.
- [16] A. Remsen, T. L. Hopkins, and S. Samson, "What you see is not what you catch: A comparison of concurrently collected net, optical plankton counter, and shadowed image particle profiling evaluation recorder data from the northeast Gulf of Mexico," *Deep Sea Res., I, Oceanogr. Res. Papers*, vol. 51, no. 1, pp. 129–151, 2004.
- [17] A. W. Herman, B. Beanlands, and E. F. Phillips, "The next generation of optical plankton counter: The Laser-OPC," *J. Plankton Res.*, vol. 26, no. 10, pp. 1135–1145, 2004.
- [18] C. S. Davis, F. T. Thwaites, S. M. Gallager, and Q. Hu, "A three-axis fast-tow digital video plankton recorder for rapid surveys of plankton taxa and hydrography," *Limnology Oceanogr. Methods*, vol. 3, no. 2, pp. 59–74, 2005.
- [19] M. Picheral, L. Guidi, L. Stemann, D. M. Karl, G. Iddoud, and G. Gorsky, "The Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton," *Limnology Oceanogr. Methods*, vol. 8, no. 9, pp. 462–473, 2010.
- [20] S. Samson, T. Hopkins, A. Remsen, L. Langebrake, T. Sutton, and J. Patten, "A system for high-resolution zooplankton imaging," *IEEE J. Ocean. Eng.*, vol. 26, no. 4, pp. 671–676, Oct. 2001.
- [21] M. Benfield, C. Schwehm, and S. Keenan, "ZOOVIS: A high resolution digital camera system for quantifying zooplankton abundance and environmental data," in *Proc. Amer. Soc. Limnology Oceanogr.*, 2001, pp. 12–17.
- [22] R. K. Cowen and C. M. Guigand, "In Situ Ichthyoplankton Imaging System (ISIS): System design and preliminary results," *Limnology Oceanogr. Methods*, vol. 6, no. 2, pp. 126–132, 2008.
- [23] J. S. Jaffe, "To sea and to see: That is the answer," *Methods Oceanogr.*, vol. 15, pp. 3–20, 2016.
- [24] H. Sun, D. C. Hendry, M. A. Player, and J. Watson, "In Situ underwater electronic holographic camera for studies of plankton," *IEEE J. Ocean. Eng.*, vol. 32, no. 2, pp. 373–382, Apr. 2007.
- [25] C. K. Sieracki, M. E. Sieracki, and C. S. Yentsch, "An imaging-in-flow system for automated analysis of marine microplankton," *Mar. Ecology Prog. Ser.*, vol. 168, no. 1, pp. 285–296, 1998.
- [26] R. J. Olson and H. M. Sosik, "A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot," *Limnology Oceanogr. Methods*, vol. 5, no. 6, pp. 195–203, 2007.
- [27] X. Tang *et al.*, "Automatic plankton image recognition," *Artif. Intell. Rev.*, vol. 12, no. 1–3, pp. 177–199, 1998.
- [28] T. Luo *et al.*, "Active learning to recognize multiple types of plankton," *J. Mach. Learn. Res.*, vol. 6, no. Apr., pp. 589–613, 2005.
- [29] Q. Hu and C. Davis, "Accurate automatic quantification of taxa-specific plankton abundance using dual classification with correction," *Mar. Ecology Prog. Ser.*, vol. 306, no. 8, pp. 51–61, 2006.
- [30] F. Zhao, F. Lin, and H. S. Seah, "Binary SIPPER plankton image classification using random subspace," *Neurocomputing*, vol. 73, no. 10–12, pp. 1853–1860, 2010.
- [31] H. M. Sosik and R. J. Olson, "Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry," *Limnology Oceanogr. Methods*, vol. 5, no. 6, pp. 204–216, 2007.
- [32] P. Vandromme *et al.*, "Assessing biases in computing size spectra of automatically classified zooplankton from imaging systems: A case study with the ZooScan integrated system," *Methods Oceanogr.*, vol. 1–2, pp. 3–21, 2012.
- [33] Z. Li, F. Zhao, J. Liu, and Y. Qiao, "Pairwise nonparametric discriminant analysis for binary plankton image recognition," *IEEE J. Ocean. Eng.*, vol. 39, no. 4, pp. 695–701, Oct. 2014.
- [34] H. Bi *et al.*, "A semi-automated image analysis procedure for in situ plankton imaging systems," *PLoS One*, vol. 10, no. 5, 2015, Art. no. e0127121.
- [35] R. Faillietaz, M. Picheral, J. Y. Luo, C. Guigand, R. K. Cowen, and J.-O. Irisson, "Imperfect automatic image classification successfully describes plankton distribution patterns," *Methods Oceanogr.*, vol. 15, pp. 60–77, 2016.
- [36] H. du Buf and M. M. Bayer, *Automatic Diatom Identification*. Singapore: World Scientific, 2002.
- [37] P. F. Culverhouse *et al.*, "Automatic classification of field-collected dinoflagellates by artificial neural network," *Mar. Ecology Prog. Ser.*, vol. 139, no. 1/3, pp. 281–287, 1996.
- [38] M. A. Mosleh, H. Manssor, S. Malek, P. Milow, and A. Salleh, "A preliminary study on automated freshwater algae recognition and classification system," *BMC Bioinf.*, vol. 13, Suppl. 17, 2012, Art. no. S25.
- [39] N. Santhi, C. Pradeepa, P. Subashini, and S. Kalaiselvi, "Automatic identification of algal community from microscopic images," *Bioinf. Biol. Insights*, vol. 7, pp. 327–334, 2013.
- [40] A. Verikas, A. Gelzinis, M. Bacauskiene, I. Olenina, and E. Vaiciukynas, "An integrated approach to analysis of phytoplankton images," *IEEE J. Ocean. Eng.*, vol. 40, no. 2, pp. 315–326, Apr. 2015.
- [41] H. Zheng, R. Wang, Z. Yu, N. Wang, Z. Gu, and B. Zheng, "Automatic plankton image classification combining multiple view features via multiple kernel learning," *BMC Bioinf.*, vol. 18, no. Suppl. 16, 2017, Art. no. 570.
- [42] Z. V. Finkel, J. Beardall, K. J. Flynn, A. Quigg, T. A. V. Rees, and J. A. Raven, "Phytoplankton in a changing world: Cell size and elemental stoichiometry," *J. Plankton Res.*, vol. 32, no. 1, pp. 119–137, 2009.
- [43] S. R. Kerr and L. M. Dickie, *The Biomass Spectrum: A Predator-Prey Theory of Aquatic Production*. New York, NY, USA: Columbia Univ. Press, 2001.

- [44] Y. J. Shin, M. J. Rochet, S. Jennings, J. G. Field, and H. Gislason, "Using size-based indicators to evaluate the ecosystem effects of fishing," *ICES J. Mar. Sci.*, vol. 62, no. 3, pp. 384–396, 2005.
- [45] M. Barange *et al.*, "Impacts of climate change on marine ecosystem production in societies dependent on fisheries," *Nature Climate Change*, vol. 4, no. 3, pp. 211–216, 2014.
- [46] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.
- [47] C. Chen, S. Li, H. Qin, and A. Hao, "Structure-sensitive saliency detection via multilevel rank analysis in intrinsic feature space," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2303–2316, Aug. 2015.
- [48] C. Chen, S. Li, Y. Wang, H. Qin, and A. Hao, "Video saliency detection via spatial-temporal fusion and low-rank coherency diffusion," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3156–3170, Jul. 2017.
- [49] S. Ren, K. He, R. Girshick, X. Zhang, and J. Sun, "Object detection networks on convolutional feature maps," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 7, pp. 1476–1481, Jul. 2017.
- [50] O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.
- [51] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [52] Y. Zhou, X. Bai, W. Liu, and L. J. Latecki, "Similarity fusion for visual tracking," *Int. J. Comput. Vis.*, vol. 118, no. 3, pp. 337–363, 2016.
- [53] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 11, pp. 2298–2304, Nov. 2017.
- [54] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-image crowd counting via multi-column convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 589–597.
- [55] D. Oñoro-Rubio and R. J. López-Sastre, "Towards perspective-free object counting with deep learning," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 615–629.
- [56] C. Zhang, H. Li, X. Wang, and X. Yang, "Cross-scene crowd counting via deep convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 833–841.
- [57] Z. Shen, Y. Xu, B. Ni, M. Wang, J. Hu, and X. Yang, "Crowd counting via adversarial cross-scale consistency pursuit," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 5245–5254.



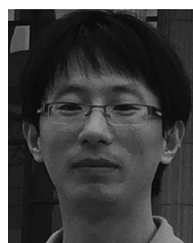
Jia Yu (M'16) received the B.S. degree in physics, the M.S. degree in optics engineering, and the Ph.D. degree in marine science, all from the Ocean University of China, Qingdao, China, in 2002, 2005, and 2008, respectively.

In 2008, she joined the Department of Physics, Ocean University of China, where she is currently an Associate Professor. Her research interests include underwater optical system development, digital holography, and microscopic system design.



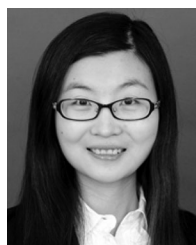
Biao Yang (M'17) received the B.S. degree from the College of Automation, Nanjing University of Technology, in 2009, the M.S. and the Ph.D. degrees from the College of Instrument Science and Technology, Southeast University, Nanjing, China, in 2011 and 2014, respectively.

He is currently with Changzhou University, Changzhou, China. His current research interests include target detection and behavior analysis based on computer vision.



Haiyong Zheng (M'12) received the B.S. degree in electronic information engineering and the Ph.D. degree in ocean information sensing and processing from the Ocean University of China, Qingdao, China, in 2004 and 2009, respectively.

In 2009, he joined the Department of Electronic Engineering, Ocean University of China, where he is currently an Associate Professor. His research interests include image processing, computer vision, and machine learning.



Nan Wang (M'16) received the B.S. degree in measurement and control technology and instrument, the M.S. degree in instrument science and technology, and the Ph.D. degree in instrument science and technology, all from Southeast University, Nanjing, China, in 2009, 2012, and 2015, respectively.

In 2016, she joined the Department of Electronic Engineering, Ocean University of China, Qingdao, China, where she is currently a Lecturer. Her research interests include underwater image processing, logical stochastic resonance, and robotics.



Bing Zheng (M'07) received the B.S. degree in electronics and information system, the M.S. degree in marine physics, and the Ph.D. degree in computer application technology from the Ocean University of China, Qingdao, China, in 1991, 1995, and 2013, respectively.

He is currently a Professor with the Department of Electronic Engineering, Ocean University of China. His research interests include ocean optics, underwater imaging, and optical detection.