Hybrid Sequence Networks for Unsupervised Water Properties Estimation From Hyperspectral Imagery

Jiahao Qi [ID], Wei Xue [ID], Zhiqiang Gong [ID], Shaoquan Zhang [ID], Aihuan Yao,
and Ping Zhong [ID], *Senior Member, IEEE*

Abstract—Generating accurate estimation of water inherent optical properties (IOPs) from hyperspectral images plays a significant role in marine exploration. Traditional methods mainly adopt bathymetric models and numerical optimization algorithms to deal with this problem. However, these methods usually tend to simplify the bathymetric models and lack the capability of describing the discrepancy between reference spectrum and estimation spectrum, resulting in a limited estimation performance. To get a more precise result, in this work, we propose a novel network based on deep learning to retrieve the IOPs. The proposed network, named as IOPs estimation network (IOPE-Net), explores a hybrid sequence structure to establish IOPs estimation module that acquires high-dimensional nonlinear features of water body spectrums for water IOPs estimation. Moreover, considering the insufficiency of labeled training samples, we design a spectrum reconstruction module combined with classical bathymetric model to train the proposed network in an unsupervised manner. Then, aiming at further promoting the estimation performance, a multicriterion loss is developed as the objective function of IOPE-Net. In particular, we construct a hierarchical multiscale sequence loss as the key component to retain the details of original spectral information. Thus, the discrepancy between different spectrums can be better described by the obtained learning model. Experimental results on both simulated and real datasets demonstrate the effectiveness and efficiency of our method in comparison with the state-of-the-art water IOPs retrieving methods.

Index Terms—Hierarchical multiscale sequence (HMS) loss, hybrid sequence structure, inherent optical properties (IOPs), unsupervised methodology.

Jiahao Qi, Aihuan Yao, and Ping Zhong are with the National Key Laboratory of Science and Technology on Automatic Target Recognition, National University of Defense Technology, Changsha 410073, China (e-mail: qijiahao1996@163.com; aihy677@163.com; zhongping@nudt.edu.cn).

Wei Xue is with the National Key Laboratory of Science and Technology on Automatic Target Recognition, National University of Defense Technology, Changsha 410073, China, and also with the School of Computer Science and Technology, Anhui University of Technology, Maanshan 243032, China (e-mail: cswxue@ahut.edu.cn).

Zhiqiang Gong is with the National Innovation Institute of Defense Technology, Chinese Academy of Military Science, Beijing 100000, China (e-mail: gongzhiqiang13@nudt.edu.cn).

Shaoquan Zhang is with the Jiangxi Province Key Laboratory of Water Information Cooperative Sensing and Intelligent Processing, Nanchang Institute of Technology, Nanchang, China (e-mail: zhangshaoquan1@163.com).

Digital Object Identifier 10.1109/JSTARS.2021.3068727

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) possess abundant spectral information owing to its high spectral resolution [1], [2], [6], which can be used to solve some complex problems in a high-dimensional feature space [3], [12], [13]. Recent years have witnessed a growing academic interest in employing hyperspectral data to settle various intricate geoscience issues in remote sensing field, such as hyperspectral classification [7], [8], [14]–[16], hyperspectral target detection [4], [5], [17], [19], and hyperspectral unmixing [20], [21].

Recently, there have been increasing efforts to introduce HSI in marine exploration, since it contributes a lot to the earth climate, ecosystem, and marine habitats [22], [23], [33], [38], [40]. However, retrieving the inherent optical properties (IOPs) of marine environment remains a challenging problem. To occupy this challenge, the common method is to construct the multidimensional, nonlinear bathymetric models at first [26], [28]–[31]. After that, the estimation of IOPs is derived by solving above models with some numerical optimization methods [34]–[36].

Lots of methods have been proposed to solve the bathymetric models in recent years, which can be roughly divided into three categories: 1) Look-up table approaches (LUTs) [40]–[42]; 2) Semianalytical inversion approaches (SAAs) [37]–[39]; and 3) Empirical approaches (EPs) [44]–[47]. LUTs are the most fundamental algorithms in this research field. These algorithms prefer to construct a parameter table by using various combinations of water properties at the very beginning. Whereas, owing to the drastic variation contained with oceans in the world, extremely huge parameter tables are required for attaining accurate estimation. These parameter tables may lead to expensive storage space and make the solution progress time-consuming [43].

The SAAs, retrieving marine IOPs using both bathymetric models and empiricism, manage to estimate the magnitudes of IOPs simultaneously. They make an underlying assumption that high-dimensional nonlinear equations can always be found to correlate the bathymetric model with sensor-observed spectrums. After that, optimization algorithms (such as Levenberg–Marquardt method) are utilized to solve the mentioned equations. These methods are capable of finding out the solution in a few steps. However, due to the limitation of optimization algorithms, the estimation results can always be suboptimum. Moreover, if SAAs are employed, numerous parameters of

bathymetric models are supposed to be determinate beforehand, which may be invalid in practical applications.

As for EPs, they exploit statistical learning or machine learning method to derive the marine IOPs from a large-scale hyperspectral water dataset. These methods hold the same idea that ample spectral information of HSI make it possible to precisely determine water properties. EPs can take full advantage of the spectral information and make a tradeoff between the statistical information and bathymetric model. Consequently, compared with the methods mentioned earlier, EPs could probably achieve better performance and possess superior generalization when the training data is sufficient. However, "curse of dimensionality" problem may blur the final results for simply employing the whole HSI as training data. Moreover, EPs can not take account of the correlation among adjacent spectral bands, which has contributed a lot in other hyperspectral geoscience missions [48]. On the other hand, traditional statistical learning or machine learning methods might not be capable of extracting the appropriate high-level spectral features, which will undermine the accuracy of retrieving result.

In this article, a novel IOPs estimation network (IOPE-Net) is proposed to estimate the desired IOPs from HSIs. The proposed network composes of two crucial modules: IOPs estimation module and spectrum reconstruction module. To acquire the discriminative spectral features and reduce the dimension of original data, a hybrid sequence structure is developed in IOPs estimation module. Furthermore, it is so hard to collect the ground truths of water IOPs that few public datasets contain the labeled training data. Taking this adverse phenomenon into consideration, a specific spectrum reconstruction module is designed for allowing the IOPE-Net to be trained in an unsupervised manner. In addition, the classical bathymetric model is also embedded into the spectrum reconstruction module, thus IOPE-Net follows the same physical background as the existing methods. Finally, we describe the reconstruction error function with a multicriterion loss to build a more appropriate objective function. The multicriterion loss employs mean square error (MSE) and spectral angel (SA) loss as the basic elements. To obtain an more accurate estimation, a new loss function named hierarchical multiscale sequence (HMS) loss is also explored as an indispensable part of the multicriterion loss. HMS loss calculates the feature discrepancy in different feature scales to further depict the differences between disparate spectrums. In addition, the proposed method does not require any IOPs information beforehand and possesses a better generalization performance. The major contributions of this article are listed as follows.

1) We introduce a novel neural network named IOPE-Net into hyperspectral underwater research filed. The proposed network explores a hybrid CNN-RNN structure for making good use of the abundant hyperspectral information. To the best of our knowledge, it is the first time that deep learning methods are applied to retrieving the IOPs from HSIs.

2) In order to address the insufficiency of labeled training samples, we construct spectrum reconstruction module with classical bathymetric model. This module make it
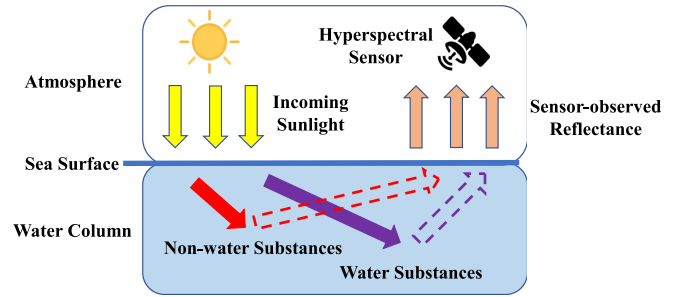

Fig. 1. Process of generating the reflectance spectrum from water.

possible to train IOPE-Net in an unsupervised methodology. In addition, with spectrum reconstruction module, IOPE-Net is guaranteed to follow the same physical background as the existing methods.

3) To obtain a more exact IOPs estimation result in complicated and changeable scene, we construct a multicriterion loss in this work. The proposed HMS loss is a vital element of this multicriterion loss. Multicriterion loss can further represent the reconstruction errors for establishing a more comprehensive objective function with HMS loss.

The remainder of this article is organized as follows. In Section II, we briefly review the models used in our research. Section III introduces all the details about the proposed IOPE-Net. In Section IV, we show the performance of our proposed method on both simulated and real datasets. Section V concludes this article.

II. PRELIMINARIES

In this section, we shortly introduce the general bathymetric model developed in oceanology. Additionally, some essential basic concepts are also defined in this section.

A. General Bathymetric Model

Hyperspectral sensor captures the reflected spectrums generated by the sunlight passing through water body. Albeit being intuitive, the process of generating reflectance spectrum from water body is illustrated in Fig. 1. Such process can be further divided into two separate procedures: downwelling and upwelling. First, the downwelling procedure describes how sunlight reaches the undersurface substances through water column. At the same time, the upwelling procedure denotes as the course of capturing reflected spectrums by hyperspectral sensors. It is noticeable that plenty of light attenuation occurs during the light transmission process.

According to above description, reflectance spectrum mainly depends on the light reflected from undersurface substances and then the impact of water column attenuation. Specifically speaking, the reflectance spectrum can be considered as the linear combination of various reflectance spectrums with coefficients determined by water column attenuation. Then, the

general bathymetric model is defined as

$$r(\lambda) = r_\infty(\lambda)\Big(1 - e^{-(k_d(\lambda)+k_u^c(\lambda))H}\Big)$$
$$+ \frac{r_B(\lambda)}{\pi}e^{-(k_d(\lambda)+k_u^b(\lambda))H} \tag{1}$$

where $r(\lambda)$ represents the sensor-observed spectrum, $r_\infty(\lambda)$ and $r_B(\lambda)$ denote as the reflectance spectrums of water column and undersurface substances, respectively. H is the depth information of undersurface substances. $K_d(\lambda)$ and $K_u(\lambda)$ are the attenuation coefficients in the downwelling and upwelling directions, respectively. In addition, the upwelling attention coefficient $K_u(\lambda)$ can be further divided into two components $K_u^c(\lambda)$ and $K_u^b(\lambda)$. More precisely, $K_u^c(\lambda)$ is related to water column and $K_u^b(\lambda)$ is connected with undersurface substances. As for undersurface substances, we may only pay attention to the targets of interests in water body. Consequently, the undersurface substances denote as the underwater targets in the rest of this article.

With (1), the mechanism of general bathymetric model is clear and distinct. However, if we would like to use this model, there exists lots of unknown parameters required to be identified in advance. Recent research works have found that all these unknown parameters are associated with the high-level IOPs absorption rate $a(\lambda)$ and scattering rate $b_b(\lambda)$ [49]. Let θ be the subsurface solar zenith angle whose value will be collected while obtaining hyperspectral datasets, and the relationship between $a(\lambda)$, $b_b(\lambda)$ and unknown parameters can be described as follows:

$$r_\infty(\lambda) \approx (0.084 + 0.170u)u \tag{2}$$

$$k_u^c(\lambda) \approx 1.03(1 + 2.4u)^{0.5}k \tag{3}$$

$$k_u^b(\lambda) \approx 1.04(1 + 5.4u)^{0.5}k \tag{4}$$

$$k_d(\lambda) = \frac{a(\lambda) + b_b(\lambda)}{\cos\theta} \tag{5}$$

where

$$u = \frac{b_b(\lambda)}{(a(\lambda) + b_b(\lambda))} \tag{6}$$

$$k = a(\lambda) + b_b(\lambda). \tag{7}$$

From (2) to (7), we can find that the unknown parameters all highly depend on the $a(\lambda)$ and $b_b(\lambda)$. More specifically, if we can find out the values of $a(\lambda)$ and $b_b(\lambda)$, all terms of the water IOPs will be determined. Consequently, so as to retrieve the water IOPs, in this work our dominant idea is to retrieve the crucial parameters $a(\lambda)$ and $b_b(\lambda)$ directly based on the spectral features of hyperspectral data.

B. Optically Deep and Optically Shallow

As mentioned above, the depth of underwater target has a profound impact on the sensor-observed spectrum $r(\lambda)$. In other words, $r_B(\lambda)$ would have less influence on $r(\lambda)$ with the depth of target gradually increasing. When the water column is too deep to permit any light from reaching the target, the reflected light captured by hyperspectral sensor is only the spectra scattered by

the water column. In this case, we can ignore the influence of underwater targets and (1) will degrade into a simple form

$$r(\lambda) = r_\infty(\lambda). \tag{8}$$

We denote the water body in such case as optically deep. On the contrary, if the water column is shallow enough, the light scattered by underwater targets will contribute to $r(\lambda)$. In this situation, the water body is denoted as the optically shallow. Optically deep and optically shallow are two crucial terms in our research. As the depth of target increases, the water body will eventually turn to optically deep and the targets will become indistinguishable from the background [43]. Optically deep is an ideal phenomenon, which will assist us to simplify the bathymetric model. Therefore, in remainder of this article, we assume background pixels can ignore the impact of targets, which is equivalent to an optically deep scenario.

III. IOPE-NET: THE WATER IOPs ESTIMATION NETWORK

In this section, we describe the proposed network for water IOPs estimation in detail. The IOPE-Net is composed of two imperative parts and the corresponding diagram is demonstrated in Fig. 2.

The first part contains a IOPs estimation module, which is used for estimating absorption rate $a(\lambda)$ and scattering rate $b_b(\lambda)$ from the input hyperspectral imagine concurrently. Consequently, the outputs of this part are the desired IOPs estimation results. The second part is designed for unsupervised training methodology. In this part, a spectrum reconstruction module has been established based on the classic bathymetric model, and we employ it to reconstruct sensor-observed spectrum with above IOPs estimation results. At the same time, the reconstruction errors evaluated subsequently are utilized to adjust the weight parameters of IOPs estimation module.

To further introduce the content of IOPE-Net, we first introduce a target-free bathymetric model and describe the general idea about how to develop the proposed network. Then, we will show the details of IOPE-Net in rest of this section.

A. Water IOPs Estimation

The complete form of bathymetric model is described in (1). It is commonly acknowledged that the target pixels always occupy a tiny percent of the whole HSI [50]. Consequently, the assumption that training data is target-free can be made while retrieving the IOPs. As is mentioned before, the background pixels can be regraded as optically deep and the bathymetric model will degrade into a simple form as (8). In this situation, if the $a(\lambda)$ and $b_b(\lambda)$ are determined, the sensor-observed spectrum $r(\lambda)$ can be well-reconstructed [30]. According to (2)–(8), the relationship among sensor-observed spectrum $r(\lambda)$, absorption rate $a(\lambda)$, and scattering rate $b_b(\lambda)$ can be summarized into a mathematical model. We denote this model as target-free model and it will be depicted as follows:

$$r(\lambda) = r_\infty(\lambda) \approx (0.084 + 0.170u)u \tag{9}$$

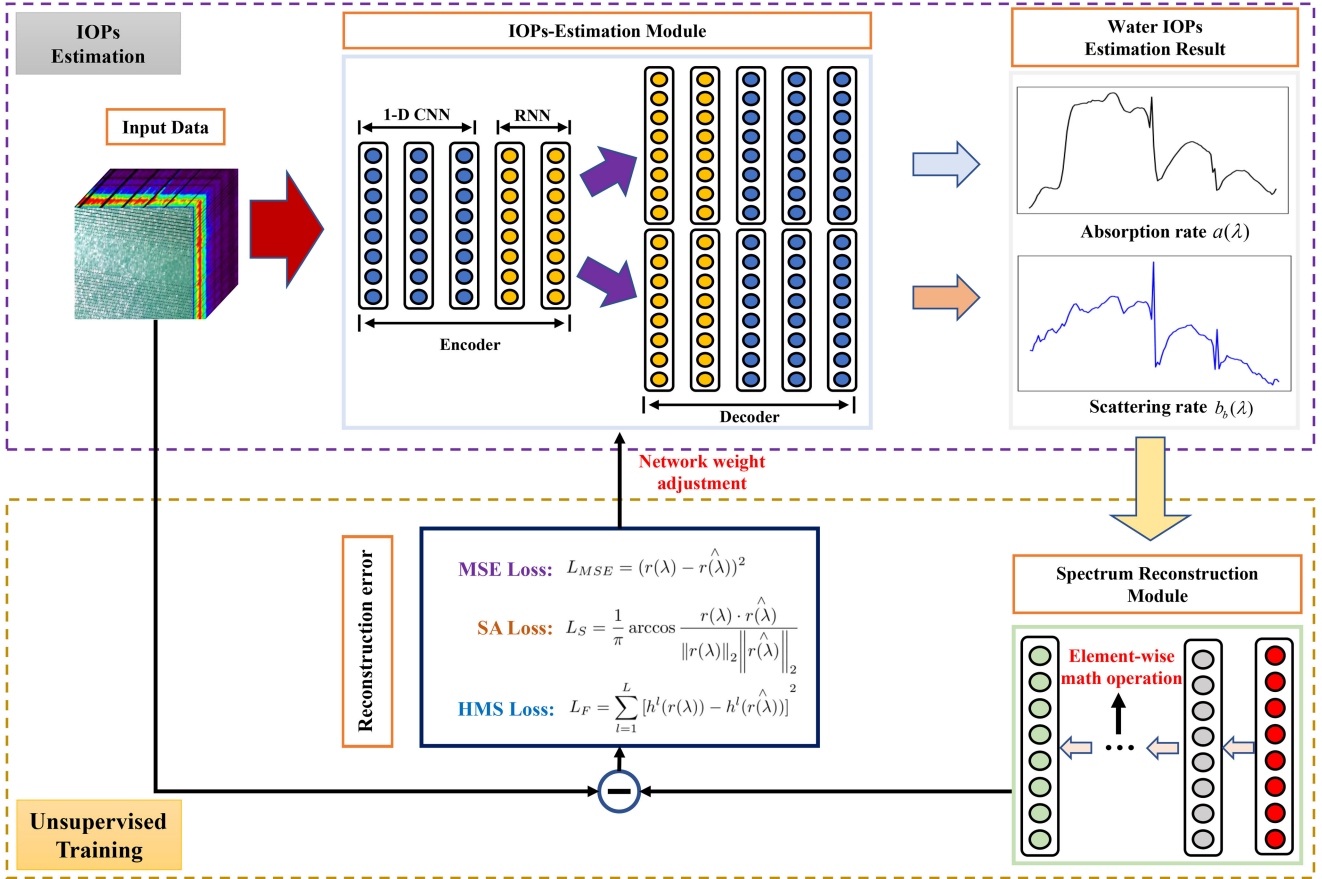$$u = \frac{b_b(\lambda)}{(a(\lambda) + b_b(\lambda))}. \tag{10}$$

Fig. 2.   Diagram of the proposed IOPE-Net.

Obviously, the target-free model tends to simplify the water IOPs retrieving problem by ignoring the effect of underwater targets. Based on this model, we can reconstruct the sensor-observed spectrum with the given absorption rate $a(\lambda)$ and scattering rate $b_b(\lambda)$. In other words, these high-level IOPs are the only required components to reconstruct the sensor-observed spectrum and the desired parameters we devote to attaining for water IOPs retrieving. Following this standpoint, in this article, we will estimate the $a(\lambda)$ and $b_b(\lambda)$ from the HSIs directly for water IOPs estimation.

Furthermore, the reconstruction error can be employed as a criterion to measure the accuracy of estimation result. It is self-evident that the solution, which can minimize the reconstruction error is bound to be the optimal solution. Considering this crucial relationship, we develop a novel IOPs estimation model, which employs the reconstruction error as objective function. This objective function can be optimized to seek out the optimal solutions of $a(\lambda)$ and $b_b(\lambda)$ simultaneously.

At the same time, inspired by the design philosophy of generative adversarial networks (GANs), the degradation model (8) is embedded subsequently as a discriminator. The task of this discriminator is to assist the IOPs estimation model in attaining a better estimation result. It also makes our method obey the same physical background as existing works.

B. IOPs Estimation Network

There is no doubt that the IOPs estimation network is the most significant part of our research work for its IOPs estimation assignment. Inspired by U-Net [53], we come up with the perspective that an encoder–decoder form network structure can separate the desired components from the original data. From (9) and (10), we can know that the absorption rate and scattering rate are the main components of the sensor-observed spectrum. Therefore, a IOPs estimation module is developed with hybrid one-dimension convolutional neural networks (1-D CNNs) and recurrent neural networks (RNNs) sequence structure in an encoder–decoder form, which is illustrated Fig. 3.

This module is considered as a predictor to tackle the simultaneous estimation of absorption rate and scattering rate by separating them from the sensor-observe spectrums. Moreover, the reason why we employ hybrid CNN–RNN sequence structure as the basic structure depends on these factors. Primarily, 1-D CNNs have been proved to achieve superior performance in terms of dealing with sequence data sets, which are capable of extracting middle-level, locally invariant and discriminative features from the input spectrum [56]. However, 1-D CNNs might be insensitive about the location information of spectral features. Where two spectral curves have identical shape but different phase positions would acquire similar convolution
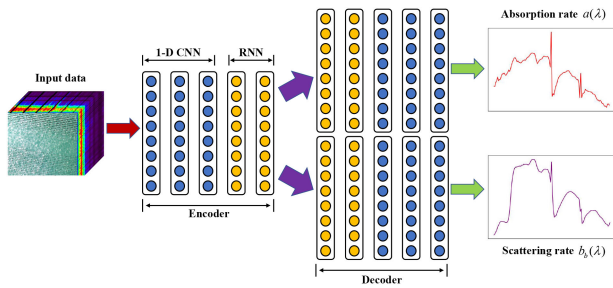
Fig. 3.   Schematic of the proposed IOPs-estimation module.

features under 1-D CNNs. This phenomenon can bring strong interference to IOPs estimation for two shape-approximate but value-different spectrums will have nearly the same estimation results.

To address this issue, we resort to the RNN blocks to dispose feature sequences derived from previous 1-D CNN blocks with the consideration of spectral contextual information among different bands. The spectral contextual information represents the correlation between different bands in HSIs, which is helpful for IOPs estimation by distinguishing shape-approximate spectrums. Then, this specific hybrid sequence structure can obtain the remarkable spectral features and contextual information synchronously.

According to Fig. 3, 1-D CNN block and RNN block are basic constituents of the proposed hybrid sequence structure. A 1-D CNN block is generally composed of several stacks with two prime parts: 1-D convolution layers and 1-D pooling layers. In convolution layers, the input sequence data are convoluted with various learnable filters to generate corresponding feature sequences. Specifically, let $\mathbf{x} \in \mathbb{R}^l$ be the input sequence and its length is l. Assuming there exists m learnable filters at this convolution layer, while weight w_j and bias b_j are frequently referred to as the parameters of jth filter. In this way, the jth output feature sequence of this convolution layer is defined as follows:

$$\mathbf{y}_j = f\left(\mathbf{x} * \mathbf{w}_j + b_j\right), j = 1, 2, \dots, m \qquad (11)$$

where $*$ denotes the convolution operation and $f(\cdot)$ refers to the activation functions, which are used to impose the nonlinearity on the network.

In our work, ReLU [24] is utilized as the activation function for it possesses two remarkable advantages: robustness for gradient vanishing and speediness for convergence. In addition, the learnable filters are employed with a receptive field 1×3 (which is the smallest size to capture the notion of left/right) for reducing the amount of network parameters. As for the 1-D pooling layers, they are periodically embedded after the convolution layers to refine the sequence features owing to the redundant information of input sequence data. At the same time, the amount of parameters and computational workload are also decreased after pooling operation.

Furthermore, we insert RNN block after 1-D CNN block for capturing the contextual information within the feature sequences. RNN architecture can identify the patterns and dynamic

temporal characteristics by using the information from current input and the activation of previous hidden states concurrently. The process of transmitting sequence information between RNN structures can generally be considered as the update of hidden state, which can be described as follows:

$$\mathbf{h}^{<t>} = \begin{cases} 0 & \text{if } t = 0 \\ f\left(\mathbf{h}^{<t-1>}, \mathbf{x}^{<t>}\right) & \text{otherwise} \end{cases} \qquad (12)$$

and

$$f\left(\mathbf{h}^{<t-1>}, \mathbf{x}^{<t>}\right) = f\left(\mathbf{w}\mathbf{x}^{<t>} + \mathbf{u}\mathbf{h}^{<t-1>} + \mathbf{b}_h\right) \qquad (13)$$

where $\mathbf{x}^{<t>} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$ is the current input feature sequence. $\mathbf{w}$ and $\mathbf{u}$ are weight matrices for current input and the activation of previous hidden state separately. $\mathbf{b}_h$ shows the corresponding bias vector and $\mathbf{h}^{<t>}$ represents the hidden state in tth step. Similarly, $f(\cdot)$ denotes the activation function and tangent function *tanh* is the most popular activation function for RNN architecture.

Due to the hidden states term, the output of RNN structure depends on both current input and previous inputs, so that the contextual information of sequence data can be collected while updating hidden states. However, the performance of RNN structure will be downgraded when the length of input sequence is too long, owing to the vanishing gradient or exploding gradient problems [27]. So, as to tackle this problem, long short-term memory (LSTM) structure is introduced, which has more parameters compared with the conventional RNN structures. In this work, the lengths of input hyperspectral curve are not too long (all below 330) and the pooling operation can also reduce the sequence length (finally below 83). In consequence, the vanishing gradient or exploding gradient problems would not take place in most cases and conventional RNN can exhibit the similar performance as LSTM under this situation [25]. To considerate the number of network parameters and the complexity of dataset, we will use the conventional RNN to deal with simulated dataset and dispose the real datasets with LSTM.

In conclusion, we have demonstrated the details of the proposed hybrid sequence structure. To give a distinct explanation of the IOPs estimation module, the flowchart of this module can be summarized as follows.

1) First, the input hyperspectral curve is fed into the CNN blocks of encoder to yield feature sequences.
2) Then, the RNN block of encoder receives the feature sequences and computes the hidden states to acquire contextual information.
3) Subsequently, the RNN block of decoder accomplishes the identical operation as prior step to get more contextual information.
4) Finally, the CNN block of the decoder will apply deconvolution operation (identical math operation as convolution) on feature sequences to estimate the water IOPs.

C. Unsupervised Training Method

1) Spectrum Reconstruction Module: Following the IOPs estimation module, a spectrum reconstruction module consisting of several element-wise math operations layers is devised in the

second part of IOPE-Net. The element-wise math operations contained in this module are designed as a reconstruction function based on classical bathymetric model. Consequently, the spectrum reconstruction module devotes to reconstructing the sensor-observed curve with the outputs of the preceding module. Moreover, another critical contribution of spectrum reconstruction module is allowing IOPE-Net to learn model parameters in an unsupervised fashion. It is commonly acknowledged that sufficient training labels are necessary for optimizing a deep neural network [9]–[11]. Nevertheless, the training labels of water IOPs are so difficult to collect that none of public available datasets contain them. Owing to lacking of ground truths, an unsupervised training method based on reconstruction error is proposed to tackle this issue. For the sake of achieving this goal, the sensor-observed curve is used as the input of network and training label simultaneously. Therefore, the reconstruction error can stem from the output of spectrum reconstruction module and sensor-observed spectrum, which is defined as

$$e_{\text{reconstruction}} = L(r(\lambda), r(\hat{\lambda})) \tag{14}$$

where $r(\lambda)$ and $r(\hat{\lambda})$ represent the sensor-observed spectrum and the output of spectrum reconstruction module, respectively. $L(\cdot)$ denotes the loss function employed in this work.

The physical essence of the reconstruction error can be interrupted as the similarity between model-retrieved curve and sensor-measured curve. Obviously, if the derived IOPs estimation results are capable of minimizing the reconstruction error, they would be deemed as the optimal solutions. In consequence, we regard reconstruction error as the objective function to adjust the parameters of neural network in IOP estimation module, which guarantees IOPE-Net to be self-sufficient and to fit the training datasets in an unsupervised methodology. Moreover, with the classical bathymetric model embedded into spectrum reconstruction module, the proposed method will follow the identical physical background as existing methods. In other words, IOPE-Net belongs to data-driven and model-driven method at the same time. Due to the data-driven property, our method is competent for yielding optimal IOPs estimation when the training samples are adequate. Meanwhile, model-driven property makes the relationship between the network topology and performance explainable and predictable [51].

2) Loss Function of Proposed Method: In general, loss functions play an important role in training the deep neural network. To further train the proposed network, we develop a multicriterion loss which is a linear combination of three significantly different loss terms. These three loss terms describe the differences between reconstructed curve and input curve in different aspects.

First, it is generally acknowledged that MSE loss is the basic measurement for prediction and regression problems. Consequently, we use it to represent the reconstruction error in the aspect of numerical value

$$L_{\text{MSE}} = ||r(\lambda) - r(\hat{\lambda})||_2 \tag{15}$$

where $r(\lambda)$ and $r(\hat{\lambda})$ denote the input curve and the output of spectrum reconstruction module, $|| \cdot ||_2$ represents the l_2 norm.

Owing to the favorable derivative characteristic, (15) is readily achievable by deep learning framework.

Apart from paying attention to the similarity about numerical value, we also place emphasis on the similarity of curve shape. To achieve this aim, we use the SA loss to describe the spectral similarity between $r(\hat{\lambda})$ and $r(\lambda)$. The SA loss is demonstrated as follows:

$$L_S = \frac{1}{\pi} \arccos \frac{r(\lambda) \cdot r(\hat{\lambda})}{||r(\lambda)||_2 ||r(\hat{\lambda})||_2}. \tag{16}$$

It is evident that the SA loss term will penalize the discrepancy of the spectral shape. At the same time, the SA loss makes contribution to accelerate the progress of finding optimal solution. Compared with a random initial curve, it can provide a better initial result for MSE loss to reach the global optimum easily. In addition, the SA loss can guarantee better convergence performance by constraining the direction of gradient and contributing the gradient value while meeting the stationary point or local optimum. Last but not least, to ensure different loss functions have identical scale, the spectral loss should be divided by the constant π to limit the value from 0 to 1 [17].

However, note that the Euclidean loss (MSE loss) is likely to blur the final result due to the deficiency of feature details [54]. To cope with this problem, a novel detail-preserving loss named HMS loss has been proposed. We hold the viewpoint that the details are highly relevant to the spectral features of input spectrums. If the spectral features can be preserved as much as possible, the inaccuracy incurred by MSE loss will be eased or even eliminated. Following this consideration, we design the HMS loss to minimize the reconstructed error in spectral feature aspect. Drawing inspiration from Siamese network [55], we employ identical pretrained neural network as feature extractors to get multiscale spectral features from $r(\lambda)$ and $r(\hat{\lambda})$ separately. The HMS loss represents the Euclidean distance among the extracted spectral features

$$L_F = \sum_{l=1}^{L} ||h^l(r(\lambda)) - h^l(r(\hat{\lambda}))||_2 \tag{17}$$

where $h^l(\cdot)$ denotes the math operation of the lth layer in the extractor neural network. It is worth mentioning that the spectral discrepancies in different layers of the feature extractors have been evaluated and the final outcome represents the sum of these spectral discrepancies. Consequently, HMS loss calculated by (17) has taken the different scales of spectral features into account. Besides, the pretrained weight derives from an autoencoder [57], which possesses the identical structure as the network proposed in IOPs estimation module. The details of the proposed loss function are illustrated in Fig. 4.

Putting everything together, the IOPE-Net work is trained by the following multicriterion loss:

$$L = L_{\text{MSE}} + \lambda_s L_S + \lambda_F L_F \tag{18}$$

where λ_S and λ_F are constants, which represent the weights of L_S and L_F in the overall loss function respectively. Both of these parameters are hyperparameters determined by the given dataset.
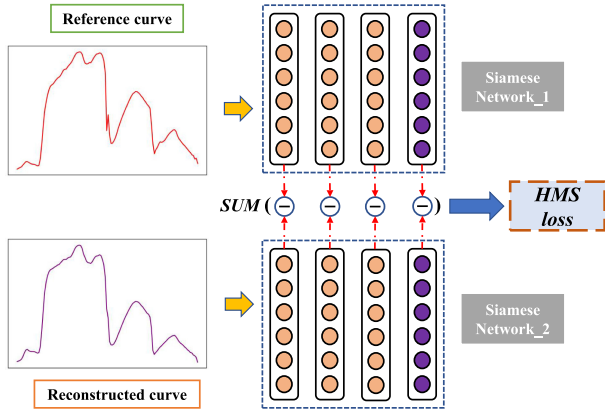
Fig. 4. Details of the proposed loss function.

So far, we are capable of figuring out an accurate and physical interpretable spectral reconstruction error with the spectrum reconstruction module and multicriterion loss. Intuitively, only if this particular reconstruction error has been minimized can we achieve a promising water IOPs retrieving result. Consequently, we could tackle the estimation of water IOPs by finding out an optimal solution for the following optimization problem:

$$\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{b}}\right) = \arg\min_{\boldsymbol{w},\boldsymbol{b}}\{L_{\mathrm{MSE}}(r(\lambda), r(\hat{\lambda}))$$

$$+ \lambda_s L_S(r(\lambda), r(\hat{\lambda})) + \lambda_F L_F(r(\lambda), r(\hat{\lambda}))\} \quad (19)$$

where $\boldsymbol{w}$ and $\boldsymbol{b}$ refers to the weight matrixes and bias vectors of IOPs estimation network. Finally, a pseudocode about the training process of IOPE-Net is given in Algorithm 1.

IV. EXPERIMENTS

In this section, we accomplish several experiments on both simulated and real hyperspectral datasets to confirm the outperformance of our proposed method. First of all, we introduce the details of employed hyperspectral datasets briefly. Second, the content of the next subsection *experiment details* is mainly about evaluation criteria and parameter settings. Then, the effectiveness of hybrid sequence structure and HMS loss proposed in Section III are tested in subsection *component analysis*. Finally, the experiment results and their corresponding analyses are detailed in remainder of this section.

A. Hyperspectral Datasets

The IOPs retrieving experiments are performed on one simulated dataset and three real hyperspectral datasets. In terms of the real hyperspectral datasets, to evaluate the performance in different scenarios, we collect the datasets from sea, lake, and river.

1) Simulated Dataset: This data is derived from the water bathymetric model [30], with specific water constituent concentration of the turbid water mentioned in [58]. A simulated water body spectrum is developed with above condition and its wavelength is fixed at 400–700 nm (150 bands at 2 nm resolution). Furthermore, a white Gaussian noise ($\sigma = 0.01$)

Algorithm 1: The Training Process of IOPE-Net.

Input: HSI $\boldsymbol{X}$, convergence condition τ.
Output: Water IOPs estimation result.
Initialization: 1) Initialize all the weight parameters in IOPE-Net with Kaiming initialization [18]; 2) Set condition variable $t \geq 0$;

1: **while** t **do**
2:     Estimate water IOPs $a(\lambda)$ and $b_b(\lambda)$ with IOPs estimation module;
3:     Reconstruct the sensor-observed spectrum with $a(\lambda)$ and $b_b(\lambda)$ via Eq. (2) to Eq. (7);
4:     Calculate the reconstruction error $e_{reconstruction}$ with Eq. (14) and Eq. (18);
5:     **if** $e_{reconstruction} \leq \tau$ **then**
6:         Set condition variable $t \leq 0$;
7:     **else**
8:         Update $\boldsymbol{w}$ and $\boldsymbol{b}$ to minimize the Eq. (19).
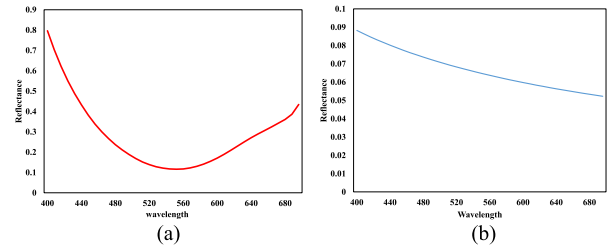9:   **end if**
10: **end while**

Fig. 5. Ground truths of simulated dataset. (a) The ground truth of absorption rate $a(\lambda)$. (b) The ground truth of absorption rate $b_b(\lambda)$.

is carried out on the generated spectrum to create 400 similar spectrums with a certain intraclass variability. After that, we use these similar spectrums to construct a simulated dataset. The ground-truth of absorption rate $a(\lambda)$ and scattering rate $b_b(\lambda)$ are illustrated in Fig. 5(a) and (b), which can be derived by the method mentioned in [58] readily.

2) Real Datasets: All the real datasets are captured by various sensors in different scenes. At the same time, atmospheric correction is implemented by the ATCOR model [59].

The first dataset was collected by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) from a gullet in Galveston Bay, Texas, America. The spectral ranges of this dataset are 224 bands over 366–2495 nm at 9.5 nm resolution. We select a 342×342 experiment chip in this dataset as the sea scenario, where the water is optically deep (over 350 m). The AVIRIS dataset and the corresponding experiment chip information are demonstrated in Fig. 6.

The second dataset is a scene of Dongting Lake with the size of 500×500 pixels in Yueyang City, Hunan Province, China. It was collected by Gaofen-5 satellite with Advanced Hyperspectral Imagery (AHSI) in 2020. The spectral resolution is about 5 nm in visible and near-infrared band and 10 nm in short-wave infrared band while the HSI has 330 spectral channels covering from 400 to 2500 nm. It is worthwhile to
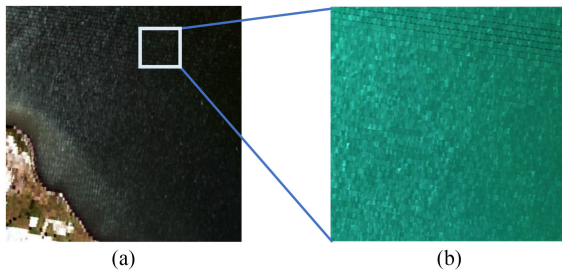
Fig. 6. Detailed information of dataset Sea. (a) The panoramic picture of dataset. (b) The experiment chip.
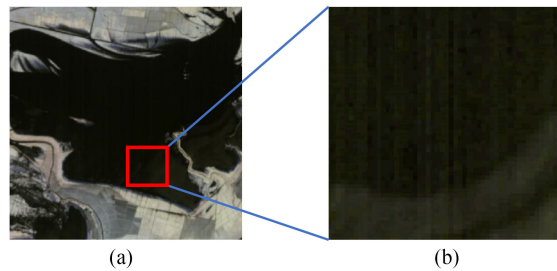

Fig. 7. Detailed information of dataset Lake. (a) The panoramic picture of dataset. (b) The experiment chip.
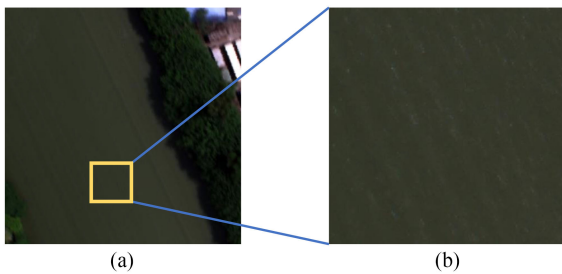

Fig. 8. Detailed information of dataset River. (a) The panoramic picture of dataset. (b) The experiment chip.

notice the water in lake is so turbid that sunlight can not reach to the bottom. In other words, everywhere in this lake can be regraded as optically deep. Similarly, a 60×60 testing chip is sampled for our experiments. The detailed information of the second dataset is illustrated in Fig. 7.

Finally, the last dataset was captured by Nano-Hyperspec and the experiment location is Nangang River, which is the typical river scene and locates in Guangzhou City, Guangdong Province, China. Compared with previous datasets, this one covers a narrow spectral range, whose wavelengths range from 400 to 1000 nm at 2.22 nm spectral resolution. Furthermore, the water in the river is also muddy enough to be considered as optical deep. A 180×180 pixels chip is sampled in our experiment, which is showed in Fig. 8.

B. Experiment Details

To further demonstrate experiment results about our research work, we will shortly introduce some additional information about subsequent experiments. The evaluation criteria employed to show the performance of different methods are exhibited firstly. Moreover, we also list the experiment settings about all the experiments.

1) Evaluation Criteria: To evaluate the performance of estimate results, we take the relative error percentage (REP) [43] of retrieving curve as one criterion, defined as follows:

$$E(r) = \sum_{l=1}^{L} \frac{\left| r_l(\overset{\sim}{\lambda}) - r_l(\lambda) \right|}{r_l(\lambda)} \tag{20}$$

where $r_l(\hat{\lambda})$ and $r_l(\lambda)$ represent the value of estimate result and ground-truth in lth band, respectively. L denotes the total number of wavelengths in testing data. Compared with other distance-based metrics, REP takes the scale information into consideration to provide a more comprehensive comparison result. Moreover, the REP can always be a positive value which is in favor of the accumulation operation.

It is worth noting that, for simulated data, we can get the ground-truth of IOP spectrums $[a(\lambda)$ and $b_b(\lambda)]$ easily following the methods mentioned in [58]. Consequently, we calculate the REP on both IOP spectrums and sensor-observed spectrums to quantify the estimation error.

As for the real datasets, we could use the same criterion to evaluate the final result as simulated data in an ideal world. Unfortunately, there is no public available data containing the ground-truth of IOP curves. We have no choice but to compute the REP only on sensor-observed spectrums for real datasets. However, to assess the performance on real data more preferably, how the accuracy of estimation results affects subsequent target detection task is regarded as another objective criterion for real datasets. To achieve this goal, several pseudounderwater targets conducted by the way proposed in [52] are embedded into real datasets as underwater targets. The ground truth information of pseudotargets in different datasets are shown in Fig. 11(a) and (b). Then, in order to evaluate performance of targets detection, the receiver operating characteristic (ROC) curves and the area under ROC curves (AUC) are utilized as another two criteria.

2) Experiment Settings: For all the datasets, 70 % of the pixels are set aside to conduct training set and the rest are used for evaluating the performance of proposed method. To demonstrate the superiority of the proposed method, it is compared with the following existing research methods: 1) SAA [43]; 2) LUT [40]; and 3) EP [44]. For all the testing methods, they are carried out on identical datasets, which have been preprocessed to eliminate the influence of environment factors (such as atmospheric scattering). In the case of LUT, only with a good initial point can it find out a rational solution. However, there is little prior information about the IOPs of real datasets. Considering the fact that SAAs and LUTs are identical perspective solved by different optimization algorithms, we finally use the solutions of SAAs as the initial points for LUTs. Additionally, to evaluate the detection performance on real datasets, classical land-based detection methods constrained energy minimization (CEM) and Adaptive Matched Filter (AMF) are employed as detectors for
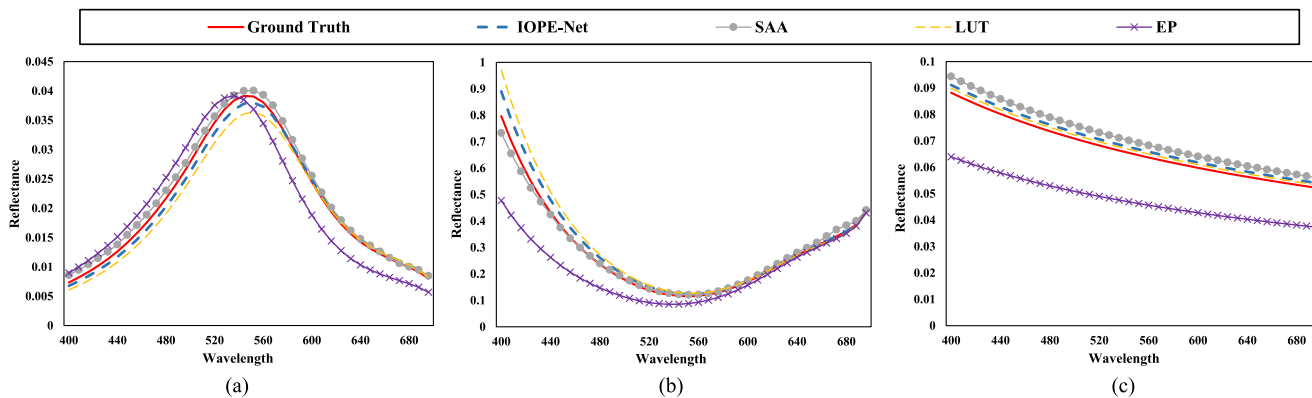
Fig. 9.   Estimation results of compared methods on the simulated data set. (a) Curves of sensor-observed reflectance $r(\lambda)$, (b) Curves of absorption rate $a(\lambda)$, (c) Curves of scattering rate $b_b(\lambda)$.
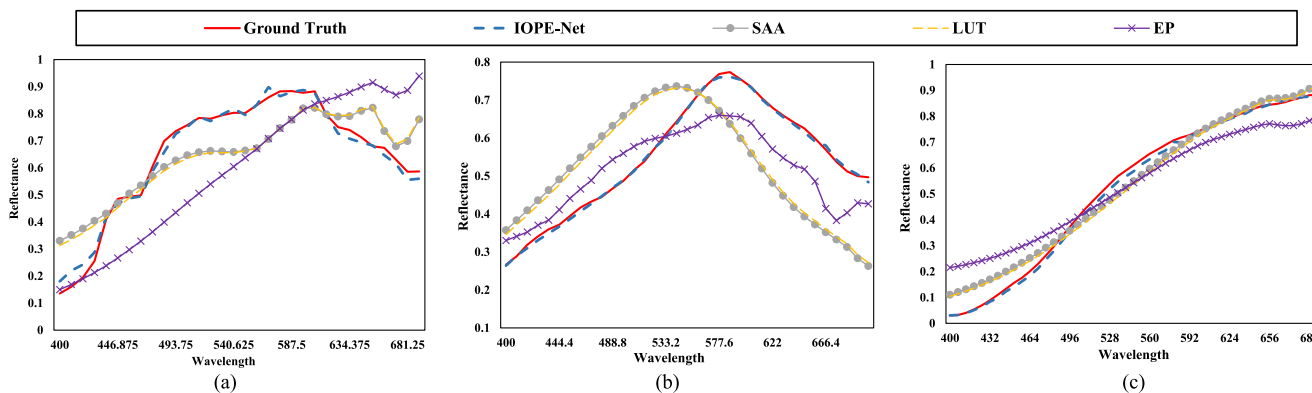

Fig. 10.   REPs results of sensor-observed reflectance $r(\lambda)$ for the real datasets. (a) The Sea scenario. (b) The River scenario. (c) The Lake scenario.
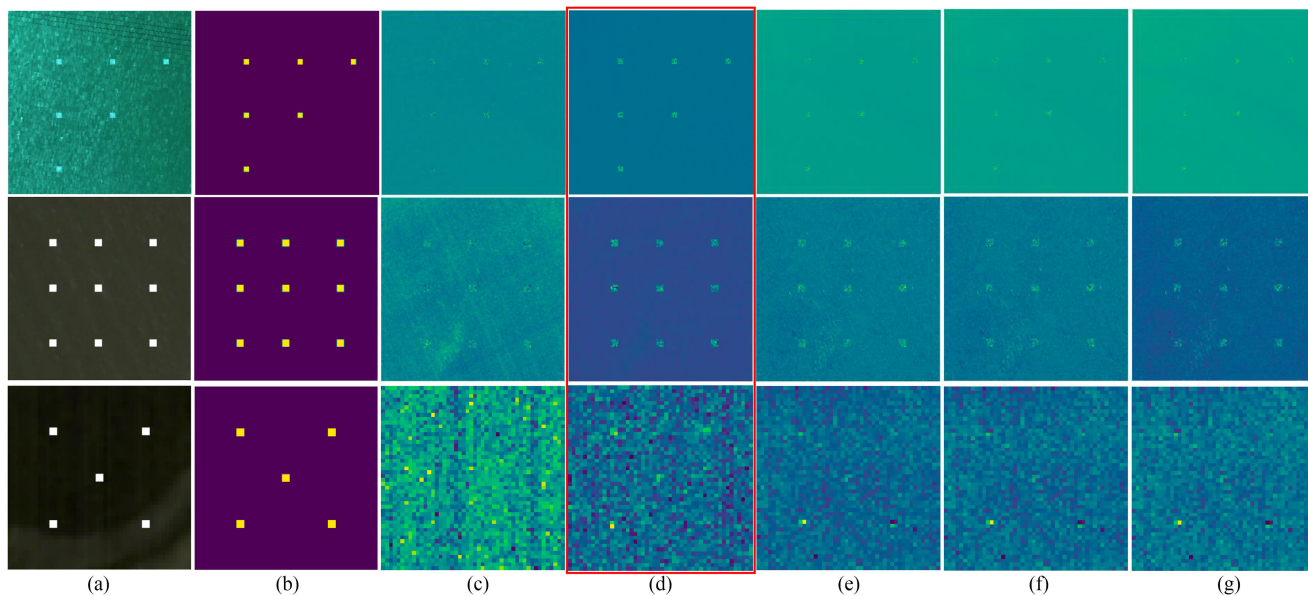

Fig. 11.   Underwater target detection results of compared methods for the real data sets. (a) Color composites of Sea, River and Lake scenarios. (b) Reference map. (c) CEM. (d) IOPE-Net+CEM. (e) SAA+CEM. (f) LUT+CEM. (g) EP+CEM. The best performances are highlighted with red.

TABLE I
EXHAUSTIVE INFORMATION OF TESTING NEURAL
NETWORK STRUCTURES

MLP	1-D CNN	1-D CNN +RNN
FC-Input Dimension	Conv1x3-16	Conv1x3-16
FC-1024	Conv1x3-32	Conv1x3-32
FC-Output Dimension	Conv1x3-64	Conv1x3-64
		RNN

TABLE II
REPS OF SENSOR-OBSERVED SPECTRUMS FOR DIFFERENT NETWORK
STRUCTURES

data set	REPs of sensor-observed spectrums		
	MLP	1-D CNN	1-D CNN +RNN
Simulated	1.9273	1.6374	**1.4172**
Sea	7.3349	2.4791	**1.6668**
River	5.5953	1.0712	**0.3396**
Lake	9.3458	3.1795	**1.5124**

The bold entries represent the best performance in each row.

subsequent high-level target detection task. Finally, all the experiments are implemented with an Intel(R) Core (TM) i9-10920X CPU machine with 64 GB of RAM. And all the testing methods are performed on a system running Python 3.8.5 and Pytorch 1.7.0.

C. Component Analysis

In this section, we will carry out some extra tests to confirm that the perspectives proposed in previous sections are valid for the IOPE-Net.

1) *Effectiveness Evaluation of Hybrid Sequence Structure:* To further verify the validity of hybrid sequence structure, three different kinds of network structures are constructed. The first one is a simple multilayer perceptron (MLP) composing of three fully connected layers, which can be considered as the baseline. The second structure consists of three 1-D CNN layers with receptive field 1×3. As for the last structure, it merely adds a RNN layer at the subsequence of 1-D CNN layers compared with the second one. The exhaustive information of these network structures is shown in Table I. The network layers are denoted as "{structure type} $<$receptive field size $> - <$ number of channels$> ."$ Three different structures are tested on all the datasets and the REPs of sensor-observed spectrums are employed as the criterion. The corresponding experimental results are listed in Table II.

Apparently, the performance of hybrid sequence structure surpasses other two structures on all the datasets. But it is noticeable that the performance gap between hybrid sequence structure and 1-D CNN based structure is not so distinct. In addition, compared with 1-D CNN based structure, hybrid sequence structure has more parameters and requires a better computation capability. The decision on selecting these two different structure is determined by a tradeoff between performance and efficiency. To

TABLE III
REPS OF SENSOR-OBSERVED SPECTRUMS AND TRAINING TIME FOR
EFFECTIVENESS EVALUATION OF HMS LOSS

data set	REPs of sensor-observed spectrums		Training time (in seconds)	
	IOPE-Net(Y)	IOPE-Net(N)	IOPE-Net(Y)	IOPE-Net(N)
Simulated	1.4172	**1.217**	23.8235	**19.47**
Sea	**1.6668**	2.8711	221.8768	**216.3741**
River	**0.3996**	1.0427	73.9031	**69.1707**
Lake	**1.5124**	2.9483	16.6329	**14.9701**

The bold entries represent the best performance in each row.

better evaluate the performance of proposed network, the hybrid sequence network structure is adopted in our experiments.

2) *Effectiveness Evaluation of HMS Loss:* To demonstrate whether HMS loss contributes to the estimation results, the proposed model trained without it is also implemented on all the datasets. Two criteria, REPs and training time, are utilized in this comparative experiment, and Table III shows the detailed results. The characters (Y and N) in parentheses imply whether the objective function contains HMS loss.

From the last two columns in Table III, we can draw the conclusion that HMS loss will make the proposed method more time-consuming. But this negative influence is so slight that it can be ignored in practical application. In terms of the estimation performance, HMS loss has significantly promoted the estimation accuracy in real datasets (raised at least 40 %). This result confirms that HMS is helpful while employing the IOPE-Net to cope with real datasets. However, HMS loss has negative impact on simulated dataset for REPs and training time simultaneously. The reason accounted for this phenomenon is that the spectrums in simulated dataset seldom contain intricate features. Consequently, the MSE and SA loss are adequate to depict the characteristics of these simple spectrums. Under this condition, HMS loss might make IOPE-Net overfitting instead of promoting its estimation performance. Taken together, HMS loss can contribute to the IOPE-Net in real scenarios but blur the performance in simulated dataset. Following this conclusion, we will only apply HMS loss to real datasets in subsequent experiments.

D. IOPs Estimation Performance

In the section, the performance of the proposed IOPE-Net is evaluated, summarized, and analyzed. Besides, the comparisons with prevalent IOPs retrieving methods SAA, LUT, and EP are also implemented both on the simulated dataset and the real datasets.

1) *Performance on Simulated Dataset:* For the simulated dataset, the reference curves and estimation curves are illustrated in Fig. 9. According to visual judgement, apart from the curves of scattering rate $b_b(\lambda)$, it is hard to determine which algorithm achieves the best performance. That is to say, owing to the simplicity of the simulated dataset, all the testing methods can acquire an excellent solution. For the quantitative analysis of the compared methods, the REP is regarded as the criterion and the corresponding result is shown in Table IV.

TABLE IV
REP RESULTS OF SENSOR-OBSERVED REFLECTANCE, ABSORPTION RATE, SCATTERING RATE ON SIMULATED DATASET

Curves	REPs of the compared methods			
	SAA	LUT	EP	IOPE-Net
$r(\lambda)$	1.6292	2.4306	6.5778	**1.4173**
$a(\lambda)$	**1.1741**	2.9651	7.2272	1.9301
$b_b(\lambda)$	2.5188	**0.7003**	9.9041	1.1910
Average	1.7741	2.0320	7.9031	**1.5128**

The bold entries represent the best performance in each row.

TABLE V
REP RESULTS OF SENSOR-OBSERVED REFLECTANCE $r(\lambda)$ FOR ALL THE COMPARED METHODS ON REAL DATASETS

data set	REPs of the compared methods			
	SAA	LUT	EP	IOPE-Net
Sea	7.5850	7.3424	8.5708	**1.6668**
River	10.0389	9.3439	4.8768	**0.3996**
Lake	15.5488	14.4993	32.5897	**1.5124**
Average	11.0576	10.3952	15.3457	**1.1930**

The bold entries represent the best performance in each row.

Excepted for the absorption rate $a(\lambda)$ and backscatter rate $b_b(\lambda)$, our proposed method outperforms SAA, LUT and EP. However, the IOPE-Net can only achieve the suboptimum performance in terms of $a(\lambda)$ and $b_b(\lambda)$. After analyzing the estimation results minutely, two dominant reasons are presented to account for this phenomenon. On the one hand, the samples in simulated data are too simple to contain the deep sequence features. On the other hand, the distribution of training data is biased since the intraclass variability of simulated dataset is inadequate. Hence, the deep learning based methods can not achieve superior performance in this situation. But it is worth mentioning that only with a favorable initial point can the traditional methods attain an ideal result [43]. As for the IOPE-Net, it requires none of prior information but can acquire the substantially identical performance as the traditional methods. This justifies that the IOPE-Net can exhibit excellent estimation performance with better practicality for simulated dataset.

2) *Performance on Real Datasets:* For real datasets, the REPs of sensor-observed spectrum $r(\lambda)$ and the performances on high-level detection task are evaluated to exhibit the validity of IOPE-Net. First, the ground truths and estimation results of sensor-observed spectrum are illustrated in Fig. 10. According to visual analysis, we can draw the conclusion that estimation results derived by the IOPE-Net fit better with the corresponding reference spectrums compared with other traditional methods. Moreover, when the training data is complex (such as the sea scenario), traditional methods may fail to represent the variability of input spectrums and can merely product simple and smooth estimation results.

Owing to the hybrid sequence structure and HMS loss, IOPE-Net can preserve the features of input spectrums regardless of the complexity of dataset. Therefore, the performance gaps between IOPE-Net and traditional methods are magnified as the water environment becomes more complicated in real data sets. This phenomenon can convincingly justify that our proposed method possesses better estimation and generalization performances in realistic scenes.

To further explore the performance gaps among different methods, a detailed numerical analysis of retrieving results is demonstrated in Table V. Similarly, in terms of REPs, our proposed method achieves the best performance among all the compared methods. And the performance difference reflected in quantitative results is more distinct that the REP of IOPE-Net is only one-tenth of the suboptimum method. Not surprisingly,

the LUT and SAA achieve similar performance for sharing the identical initial points. Although the REPs results indicate that LUT outperforms SAA on the average performance, it is hard to determine which algorithm is better. Obviously, without employing the results of SAA as initial points, LUT may spend plenty of time (a day even longer) to seek out the optimal solution. At the worst, LUT may fail to find out a rational result after costing lots of time.

Unfortunately, we can not simply pick out the optimal method based on criterion REPs for lacking the ground truths of IOPs. The smaller REPs of sensor-observed spectrum $r(\lambda)$ merely prove that IOPs derived from algorithms are one set of parameters, which can reconstruct the target curves well. That is to say, the results figured out by traditional algorithms or IOPE-Net are feasible solutions instead of unique Solutions. Intuitively, if the estimation results can contribute to the performance of subsequent detection tasks, the corresponding algorithm is of great usage in practical scenes and the estimation result is close to the unique solution. Consequently, two classical land-based detection algorithms are modified by the estimation results of compared methods. The perspective to coalesce detection algorithms and compared methods is predicting what detection target would look like in the given body of water with (1) and estimation IOPs. Then, the influence of water body is removed and the detection method can achieve a better performance.

For all the real datasets, the corresponding maps such as ground truths and detection results are illustrated in Figs. 11 and 12. As expected, the detection results based on IOPE-Net product the slightest visual difference with reference maps. To put this in another way, the IOPs stemming from IOPE-Net make the greatest contribution to detection task in vision. Furthermore, in comparison with traditional methods, our method has significantly highlighted the desired targets and suppressed the background (water body pixels) in detection maps. This phenomenon indicates the IOPs estimation result of our proposed method has effectively remove the influence of water body.

In order to analyze the detection performance qualitatively, the ROC curves of false positive rate and true positive rate are plotted in Figs. 13 and 14. It is obvious that our method remains over the existing methods SAA, LUT, and EP for all the datasets. This confirms the IOPE-Net generates more positive influence on detection algorithms with the estimated IOPs result. At the same time, compared with using CEM or MF to detect underwater targets directly, the combination of IOPE-Net and land-based
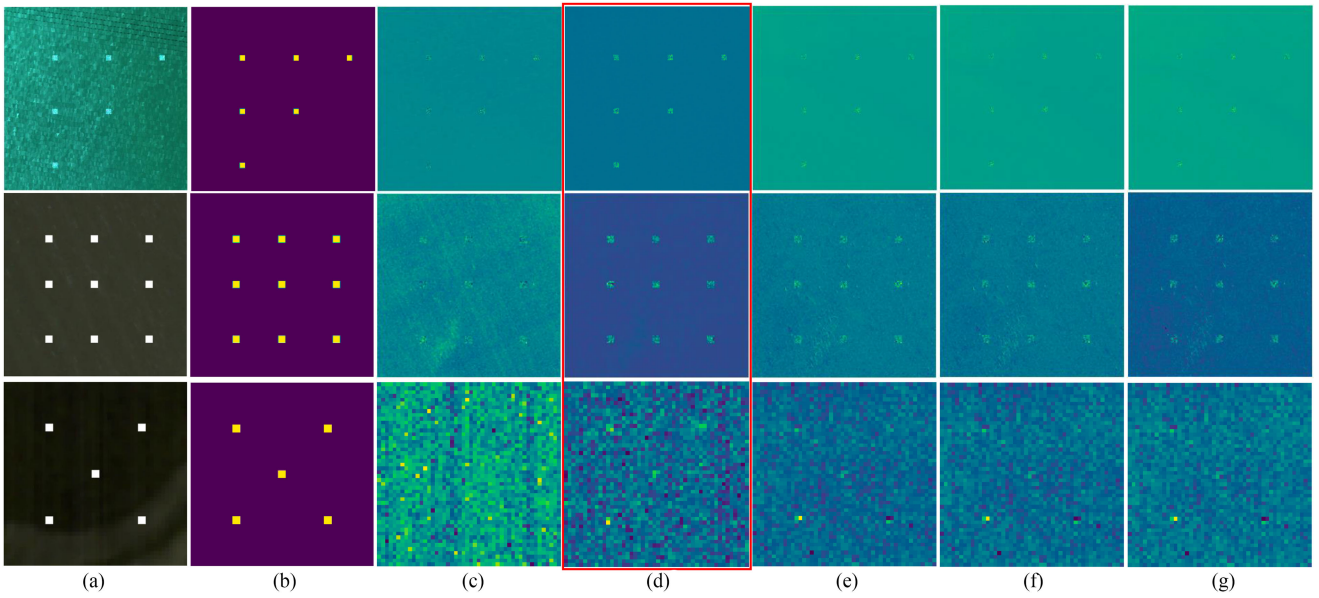
Fig. 12. Underwater target detection results of compared methods for the real data sets. (a) Color composites of Sea, River and Lake scenarios. (b) Reference map. (c) MF. (d) IOPE-Net+MF. (e) SAA+MF. (f) LUT+MF. (g) EP+MF. The best performances are highlighted with red.
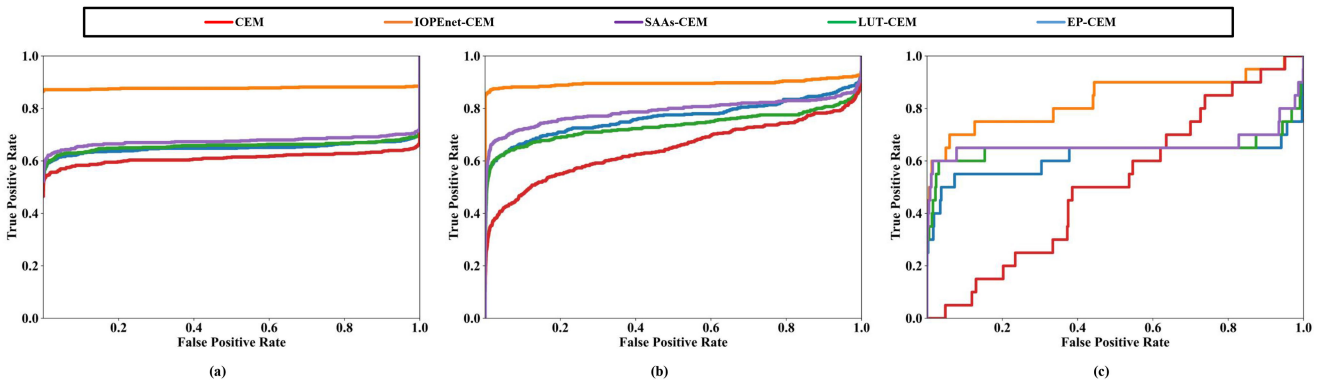

Fig. 13. ROC curves for CEM target detection method on (a) Sea dataset. (b) River dataset. (c) Lake dataset.
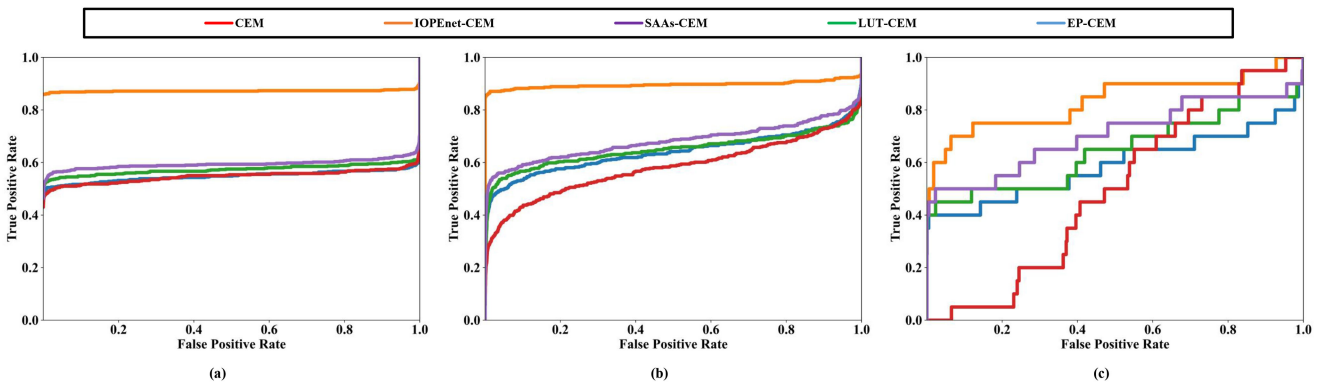

Fig. 14. ROC curves for MF target detection method on (a) Sea dataset. (b) River dataset. (c) Lake dataset.

TABLE VI
AUC VALUES OF UNDERWATER TARGET DETECTION RESULTS FOR ALL THE COMPARED METHODS WITH CEM ON REAL DATA SETS

data set	Area Under ROC Curves				
	CEM	SAA+CEM	LUT+CEM	EP+CEM	IOPE-Net+CEM
Sea	0.5908	0.6650	0.6368	0.6294	**0.8778**
River	0.5407	0.7309	0.6683	0.6823	**0.8363**
Lake	0.5137	0.6287	0.6124	0.5839	**0.8358**
Average	0.5417	0.6782	0.6458	0.6385	**0.8500**

The bold entries represent the best performance in each row.

TABLE VII
AUC VALUES OF UNDERWATER TARGET DETECTION RESULTS FOR ALL THE COMPARED METHODS WITH MF ON REAL DATA SETS

data set	Area Under ROC Curves				
	MF	SAA+MF	LUT+MF	EP+MF	IOPE-Net+MF
Sea	0.5031	0.5940	0.5721	0.5179	**0.8723**
River	0.5209	0.6404	0.6292	0.6151	**0.8265**
Lake	0.49534	0.6652	0.6150	0.5901	**0.8344**
Average	0.5065	0.6332	0.6054	0.5744	**0.8444**

The bold entries represent the best performance in each row.

TABLE VIII
AVERAGE EXECUTION TIME (IN SECONDS) OF THE COMPARED METHODS

Data set	Execution time (in seconds)			
	SAA	LUT	EP	IOPE-Net
Simulated	**27.841**	97.547	74.427	44.877
Sea	**59.153**	252.483	155.128	94.758
River	**35.718**	174.392	106.714	72.297
Lake	**15.578**	53.419	32.741	24.485
Average	**34.573**	144.460	92.253	59.104

The bold entries represent the best performance in each row.

detection can achieve a prominent improvement on detection performance. This exciting result indicates that IOPE-Net can be applied to tackle the issues existing in hyperspectral underwater target detection research work.

Furthermore, for depicting the detection results in a quantitative way, the AUC scores of plotted ROC curves are calculated in Tables VI and VII. Notably, the IOPE-Net demonstrates the best performance among all the testing methods once again. With the assistance of IOPE-Net, the average AUC of the MF and CEM are 0.8444 and 0.85, respectively, while the baselines are only 0.5065 and 0.5417. In addition, the execution time of each method is also evaluated as another metric to provide quantitative analysis which are shown in Table VIII. All the compared methods have been conducted on the experimental conditions mentioned in the section *Experiment Settings*. Obviously, for a learning-based method, it will spend plenty of time on the training process. Therefore, we only record the running time of testing stage after the developed model has been well fitted. From Table VIII, we can find that the SAA method can achieve best performance in efficient evaluation, while IOPE-Net surpasses

TABLE IX
STATISTICAL INFORMATION OF THE REP RESULTS FOR METHOD ANALYSIS

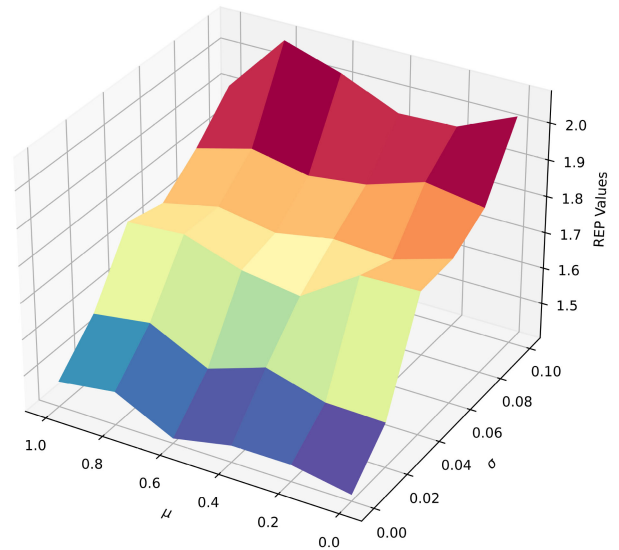| Data set	Statistical information			
	Minimal Value	Maximal Value	Median Value	Avarage Value
Simulated	1.414	2.072	1.723	1.756
Sea	1.632	2.198	1.985	1.959
River	0.411	1.288	0.816	0.843
Lake	1.514	2.127	1.803	1.801
Average	1.243	1.921	1.582	1.590

Fig. 15. REP results of sensor-observed reflectance $r(\lambda)$ for IOPE-Net on simulated data under different noisy conditions.

the EP method and LUT method. The reason accounting for this result is that SAA has simplified the bathymetric model during the water IOPs retrieving process, which can contribute to the computing speed but undermine the estimation accuracy. As for our proposed method, IOPE-Net is capable of achieving an excellent retrieving result with a tolerable time consuming. Considering every aspect of the experiment result, we can finally come to the conclusion that IOPE-Net can achieve superior performance in promoting the detection capability of land-based algorithms for all the real datasets.

E. Method Analysis

In this section, some analyses about the characteristics of IOPE-Net are demonstrated. Apparently, robustness and stability are the essential qualities of an excellent method. Therefore, we tend to evaluate whether the proposed method can achieve a satisfied performance under a certain degree of the environmental interference. Noise, the most common environmental interference, is convenient to be imposed on the training datasets. In our experiments, we employ the multivariate white Gaussian noise (MWGN) to test the robustness and stability of IOPE-Net,
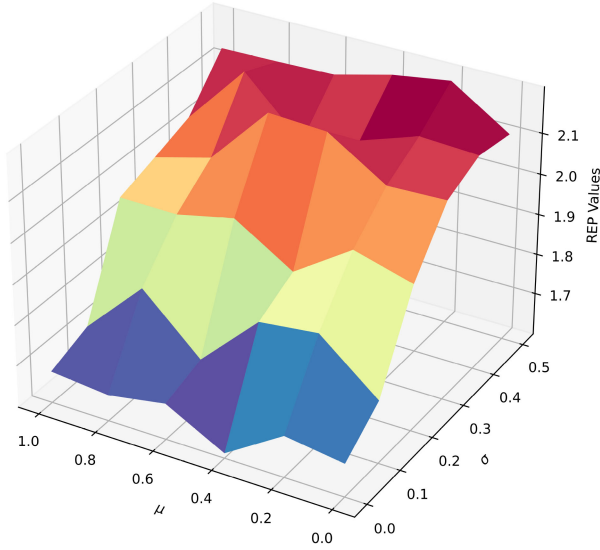
Fig. 16. REP results of sensor-observed reflectance $r(\lambda)$ for IOPE-Net on sea data under different noisy conditions.
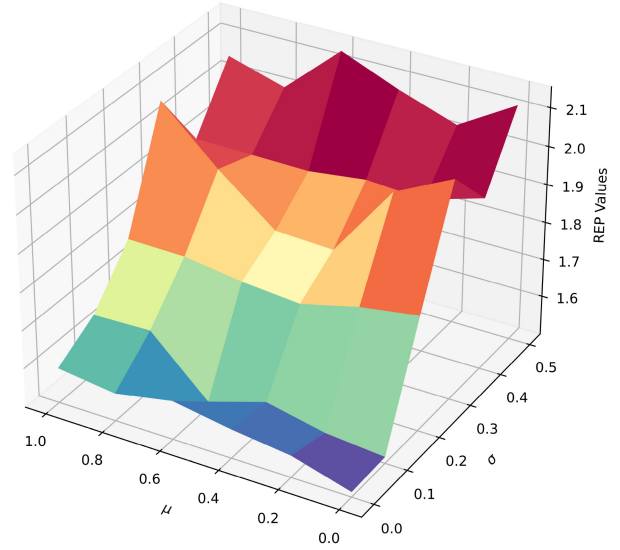

Fig. 18. REP results of sensor-observed reflectance $r(\lambda)$ for IOPE-Net on lake data under different noisy conditions.
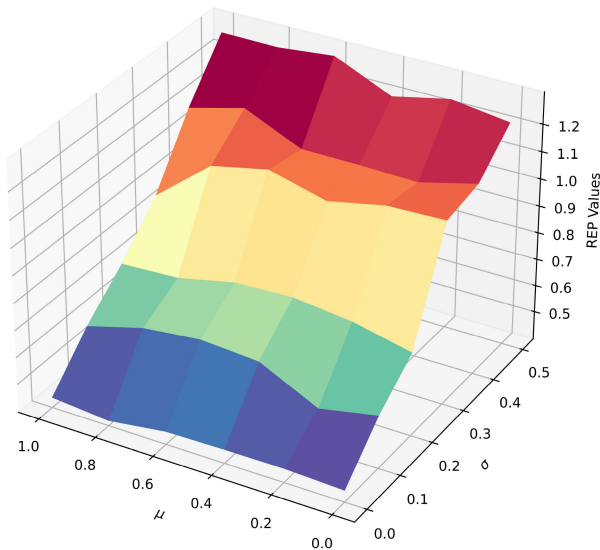

Fig. 17. REP results of sensor-observed reflectance $r(\lambda)$ for IOPE-Net on river data under different noisy conditions.

all the datasets and the associated results have been illustrated in Figs. 15–18.

It is effortless to find that the parameter covariance value seems to impose more effect on the IOPs estimation result compared with mean value. According to the visual judgement, the MGWN would not bring dramatic change to the retrieving performance of the IOPE-Net, which indicates that our method can conquer the interference of environmental noises. So, as to acquire a quantitative analysis, some statistical information of the REP results is listed in Table IX. Integrating the results demonstrated in Table IV, V, and IX, we can draw the conclusion that IOPE-Net still outperforms other compared methods even if the datasets have been interfered by MWGN. In summary, the IOPE-Net is a stable and robust water IOPs estimation method, which may have preferable generalization performance and adaptive capacity in practical application.

V. CONCLUSION

In this article, a novel neural network named IOPE-Net based on deep learning is proposed to retrieve IOPs of water body for HSIs. It contains three meaningful contributions: hybrid sequence structure, spectrum reconstruction module, and multi-criterion loss. In hybrid sequence structure, we employ 1-D CNN and RNN blocks to construct a specific network structure as predictor to estimate the IOPs of a given hyperspectral water image. The corresponding supplementary tests have demonstrated this structure can achieve better performance. Then, to tackle the issue of lacking labeled training samples, we design a spectrum reconstruction module with bathymetric model to guarantee the IOPE-Net can be trained in an unsupervised methodology. With this spectrum reconstruction module, our proposed method attains better generalization capacity and follows the identical physical background as the existing methods.

which can be depicted as follows:

$$n(\lambda) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\lambda - \mu)^2}{2\sigma^2}\right) \quad (21)$$

where μ and σ represent the mean vector and covariance matrix, respectively. For simplicity, the statistical parameters of $n(\lambda)$ are set to be identical in different bands. Then, the mean value μ and covariance value σ are the main variables for this MWGN, which are sampled from the range 0–1 with step 0.2 and the range 0–0.5 with step 0.1. After that, corresponding noises are added to hyperspectral datasets to generate new training and testing samples. We employ the REPs of sensor-observed spectrum $r(\lambda)$ as metric to validate the performances of IOPE-Net for disposing

As for multicriterion, it is a combination of different loss terms, which aims at constructing a more comprehensive objective function for IOPE-Net. In addition, HMS loss is the most significant component of multicriterion among different loss terms. A pretrained Siamese network is used in HMS loss to calculate the difference between reference curve and reconstruction curve. Taking the discrepancy of different spectral feature scales into consideration, we compute the sum of spectral feature loss in different layers as the final HMS loss. Experiments accomplished on both simulated and real datasets all indicate that IOPE-Net is superior to the compared methods. The outperformance of our method validates the significance of overcoming the challenges with deep learning methods in hyperspectral underwater research field.

ACKNOWLEDGMENT

REFERENCES

[1] S. Mei, G. Zhang, J. Li, Y. Zhang, and Q. Du, "Improving spectral-based endmember finding by exploring spatial context for hyperspectral unmixing," *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.*, vol. 13, pp. 3336–3349, Jun. 2020.

[2] X. Xu, J. Li, S. Li, and A. Plaza, "Curvelet transform domain-based sparse nonnegative matrix factorization for hyperspectral unmixing," *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.*, vol. 13, pp. 4908–4924, Aug. 2020.

[3] L. He, J. Zhu, J. Li, D. Meng, J. Chanussot, and A. Plaza, "Spectral-fidelity convolutional neural networks for hyperspectral pansharpening," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5898–5914, Sep. 2020.

[4] D. Zhu, B. Du, and L. Zhang, "Target dictionary construction-based sparse representation hyperspectral target detection methods," *IEEE J. Sel. Top. Appl. Earth Observ. Remote. Sens.*, vol. 12, no. 4, pp. 1254–1264, Apr. 2019.

[5] W. Xie, X. Zhang, Y. Li, K. Wang, and Q. Du, "Background learning based on target suppression constraint for hyperspectral target detection," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5887–5897, Sep. 2020.

[6] M. Borengasser, W. Hungate, and R. Watkins, *Hyperspectral Remote Sensing: Principles and Applications*. Boca Raton, FL, USA: CRC Press, 2008.

[7] S. Wan, C. Gong, P. Zhong, S. Pan, G. Li, and J. Yang, "Hyperspectral image classification with context-aware dynamic graph convolutional network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 597–612, May 2021.

[8] S. Wan, C. Gong, P. Zhong, B. Du, L. Zhang and J. Yang, "Multiscale dynamic graph convolutional network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3162–3177, May 2020.

[9] C. Gong, H. Shi, T. Liu, C. Zhang, J. Yang, and D. Tao, "Loss decomposition and centroid estimation for positive and unlabeled learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 3, pp. 918–932, Sep. 2021.

[10] C. Gong, D. Tao, W. Liu, L. Liu and J. Yang, "Label propagation via teaching-to-learn and learning-to-teach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 6, pp. 1452–1465, Jun. 2017.

[11] C. Gong, D. Tao, S. J. Maybank, W. Liu, G. Kang and J. Yang, "Multimodal curriculum learning for semi-supervised image classification," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3249–3260, Jul. 2016.

[12] J. G. P. W. Clevers, and L. Kooistra, "Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 574–583, Apr. 2012.

[13] S. Rashwan, N. Dobigeon, W. Sheta, and H. Hassan, "Non-linear unmixing of hyperspectral images using multiple-kernel self-organising maps," *IET Image Process.*, vol. 13, no. 12, pp. 2190–2195, 2019.

[14] Z. Gong, P. Zhong, Y. Yu, and W. Hu, "Diversity-promoting deep structural metric learning for remote sensing scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 371–390, Jan. 2018.

[15] Z. He, J. Li, K. Liu, L. Liu, and H. Tao, "Kernel low-rank multitask learning in variational mode decomposition domain for multi-/hyperspectral classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 4193–4208, Jul. 2018.

[16] P. Zhong, Z. Gong, S. Li, and C. Schonlieb, "Learning to diversify deep belief networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3516–3530, Jun. 2017.

[17] W. Xie, J. Yang, J. Lei, Y. Li, Q. Du, and G. He, "SRUN: Spectral regularized unsupervised networks for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1463–1474, Feb. 2020.

[18] K. He, X. Zhang, S. Ren and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1026–1034.

[19] T. Jiang, Y. Li, W. Xie, and Q. Du, "Discriminative reconstruction constrained generative adversarial network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4666–4679, Jul. 2020.

[20] X. Xu, J. Li, S. Li, and A. Plaza, "Generalized morphological component analysis for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2817–2832, Apr. 2020.

[21] Y. Su, J. Li, A. Plaza, A. Marinoni, P. Gamba, and S. Chakravortty, "DAEN: Deep autoencoder networks for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4309–4321, Jul. 2019.

[22] C. D. Mobley, "Comparison of numerical models for computing underwater light fields," *Appl. Opt.*, vol. 32, no. 36, pp. 7484–7504, 1993.

[23] P. J. Werdell *et al.*, "An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing," *Prog. Oceanogr.*, vol. 160, pp. 186–212, 2018.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[25] H. Wu and S. Prasad, "Convolutional recurrent neural networks for hyperspectral data classification," *Remote Sens.*, vol. 9, no. 3, pp. 298–317, 2017.

[26] A. Albert and C. D. Mobley, "An analytical model for subsurface irradiance and remote sensing reflectance in deep and shallow case-2 waters," *Opt. Exp.*, vol. 11, no. 22, pp. 2873–2890, 2003.

[27] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6690–6709, Apr. 2019.

[28] R. J. W. Brewin, E. Devred, S. Sathyendranath, S. J. Lavender, and N. J. Hardman-Mountford, "Model of phytoplankton absorption based on three size classes," *Appl. Opt.*, vol. 50, no. 22, pp. 4535–4549, 2011.

[29] Z. Lee, K. L. Carder, C. D. Mobley, R. G. Steward, and J. S. Patch, "Hyperspectral remote sensing for shallow waters. A semianalytical model," *Appl. Opt.*, vol. 37, no. 27, pp. 6329–6338, 1998.

[30] Z. Lee, K. L. Carder, C. D. Mobley, R. G. Steward, and J. S. Patch, "Hyperspectral remote sensing for shallow waters: 2. deriving bottom depths and water properties by optimization," *Appl. Opt.*, vol. 38, no. 18, pp. 3831–3843, 1999.

[31] H. R. Gordon, "Simple calculation of the diffuse reflectance of the ocean," *Appl. Opt.*, vol. 12, no. 12, pp. 2803–2804, 1973.

[32] Q. Dong, S. Shang, and Z. Lee, "An algorithm to retrieve absorption coefficient of chromophoric dissolved organic matter from ocean color," *Remote Sens. Environ.*, vol. 128, pp. 259–267, 2013.

[33] R. A. Garcia, L. I. W. McKinna, J. D. Hedley, and P. R. C. S. Fearns, "Improving the optimization solution for a semi-analytical shallow water inversion model in the presence of spectrally correlated noise," *Limnology Oceanogr. Methods*, vol. 12, no. 10, pp. 651–669, 2014.

[34] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distrib. Process., Explorations Microstruct. Cogn.: Found.*, MIT Press, 1987, pp. 318–362.

[35] B. T. Polyak, "Newton's method and its use in optimization," *Eur. J. Oper. Res.*, vol. 181, no. 3, pp. 1086–1096, 2007.

[36] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Math. Prog.*, vol. 45, no. 1, pp. 503–528, 1989.

[37] V. E. Brando, A. G. Dekker, Y. J. Park, and T. Schroeder, "Adaptive semianalytical inversion of ocean color radiometry in optically complex waters," *Appl. Opt.*, vol. 51, no. 15, pp. 2808–2833, 2012.

[38] P. J. Werdell, B. A. Franz, S. W. Bailey, G. C. Feldman, and E. Boss, "Generalized ocean color inversion model for retrieving marine inherent optical properties," *Appl. Opt.*, vol. 52, no. 10, pp. 2019–2037, 2013.

[39] R. J. W. Brewin *et al.*, "Regional ocean-colour chlorophyll algorithms for the Red Sea," *Remote Sens. Environ.*, vol. 165, pp. 64–85, 2015.

[40] C.-C. Liu and R. L. Miller, "Spectrum matching method for estimating the chlorophyll-a concentration, CDOM ratio, and backscatter fraction from remote sensing of ocean color," *Can. J. Remote Sens.*, vol. 34, no. 4, pp. 343–355, 2008.

[41] J. Hedley, C. Roelfsema, and S. R. Phinn, "Efficient radiative transfer model inversion for remote sensing applications," *Remote Sens. Environ.*, vol. 113, no. 11, pp. 2527–2532, 2009.

[42] C. D. Mobley, L. K. Sundman, C. O. Davis, J. H. Bowles, and T. V. Downes, "Interpretation of hyperspectral remote-sensing imagery by spectrum matching and look-up tables," *Appl. Opt.*, vol. 44, no. 17, pp. 3576–3592, 2005.

[43] D. B. Gillis, J. H. Bowles, M. J. Montes, and W. D. Miller, "Deriving bathymetry and water properties from hyperspectral imagery by spectral matching using a full radiative transfer model," *Remote Sens. Lett.*, vol. 11, no. 10, pp. 903–912, 2020.

[44] F. Cao *et al.*, "Remote sensing retrievals of colored dissolved organic matter and dissolved organic carbon dynamics in north american estuaries and their margins," *Remote Sens. Environ.*, vol. 205, pp. 151–165, 2018.

[45] S. E. Craig *et al.*, "Deriving optical metrics of coastal phytoplankton biomass from ocean colour," *Remote Sens. Environ.*, vol. 119, pp. 72–83, 2012.

[46] A. Mannino, M. G. Novak, S. B. Hooker, K. Hyde, and D. Aurin, "Algorithm development and validation of CDOM properties for estuarine and continental shelf waters along the northeastern US coast," *Remote Sens. Environ.*, vol. 152, pp. 576–602, 2014.

[47] M. Soja-Woźniak, M. Darecki, B. Wojtasiewicz, and K. Bradtke, "Laboratory measurements of remote sensing reflectance of selected phytoplankton species from the Baltic sea," *Oceanologia*, vol. 60, no. 1, pp. 86–96, 2018.

[48] H. Sun, X. Zheng, X. Lu, and S. Wu, "Spectral-spatial attention network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3232–3245, May 2020.

[49] J. Chen, W. Quan, T. Cui, Q. Song, and C. Lin, "Remote sensing of absorption and scattering coefficient using neural network model: Development, validation, and application," *Remote Sens. Environ.*, vol. 149, pp. 213–226, 2014.

[50] S. Jay, M. Guillaume, and J. Blanc-Talon, "Underwater target detection with hyperspectral data: Solutions for both known and unknown water quality," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1213–1221, Aug. 2012.

[51] Z. Xu, and J. Sun, "Model-driven deep-learning," *Nat. Sci. Rev.*, vol. 5, no. 1, pp. 22–24, 2017.

[52] D. B. Gillis, "An underwater target detection framework for hyperspectral imagery," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1798–1810, Apr. 2020.

[53] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2015, pp. 234–241.

[54] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3194–3203.

[55] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 539–546.

[56] J. Wang, X. Song, L. Sun, W. Huang, and J. Wang, "A novel cubic convolutional neural network for hyperspectral image classification," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4133–4148, Jul. 2020.

[57] G. E. Hinton, and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[58] V. E. Brando, J. M. Anstee, M. Wettle, A. G. Dekker, S. R. Phinn, and C. Roelfsema, "A physics based retrieval and quality assessment of bathymetry from suboptimal hyperspectral data," *Remote Sens. Environ.*, vol. 113, no. 4, pp. 755–770, 2009.

[59] R. Richter and D. Schläpfer, "Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction," *Int. J. Remote Sens.*, vol. 23, no. 13, pp. 2631–2649, 2002.