Multimodal Deep Learning for Heterogeneous GNSS-R Data Fusion and Ocean Wind Speed Retrieval

Xiaohan Chu, Jie He, Hongqing Song, Yue Qi, Yueqiang Sun, Weihua Bai, Wei Li, and Qiwu Wu

***Abstract*—The comprehensiveness of the raw input data and the effectiveness of feature engineering are two key factors affecting the performance of machine learning. To improve the data comprehensiveness for Global Navigation Satellite System Reflectometry (GNSS-R) ocean wind speed retrieval, this article introduces a new input data structure, which is composed of Delay–Doppler maps (DDM) and all satellite receiver status (SRS) parameters. Then, to overcome the difficulty of handcrafted feature engineering and effectively fusion the information of DDM and SRS, we presented a heterogeneous multimodal deep learning (HMDL) method to retrieve the wind speed according to the heterogeneity of the input data. The proposed model is verified by the performance evaluation of realistic data sets obtained from TDS-1. The new input data structure improves the prediction accuracy at 13.5% to 30.7% on mean absolute error (MAE) at 10.6% to 29.5% on the root mean square error (RMSE). The HMDL improves the prediction accuracy at 7.7% on MAE and 7.1% on RMSE. The whole proposed solution improves the prediction accuracy at 36.3% on MAE and 36.8% on RMSE, comparing with the traditional neural network-based solution. The results clearly show that both the introduction of the new input data structure and HMDL effectively improve the accuracy and robustness of GNSS-R wind speed retrieval.**

***Index Terms*—Delay–Doppler map (DDM), GNSS-reflectometry (GNSS-R), multimodal deep learning, ocean surface wind retrieval.**

Manuscript received February 24, 2020; revised May 30, 2020 and July 2, 2020; accepted July 9, 2020. Date of publication August 11, 2020; date of current version October 14, 2020. This work was supported in part by the National Key R&D Program of China under Grant 2018YFB0704300 and Grant 2016YFC0901303, in part by the National Natural Science Foundation of China Project under Grant 61671056 and Grant 61971031, in part by Scientific and Technological Innovation Foundation of Shunde Graduate School, USTB under Grant BK19AF007, in part by Interdisciplinary Research Project of USTB under Grant FRF-IDRY-19-019, in part by the Natural Science Basic Research Plan in Shanxi Province of China under Grant 2020JM-361, the Young and Middle-Aged Scientific Research Backbone Projects of Engineering University of PAP under Grant KYGG201905, and in part by the Basic Research Foundation Project of Engineering University of PAP under Grant WJY201920. *(Xiaohan Chu and Jie He contributed equally to this work.) (Corresponding authors: Hongqing Song; Yue Qi.)*

Xiaohan Chu is with the School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: chuxiaohan95@163.com).

Jie He and Yue Qi are with the School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China, and also with the Shunde Graduate School, University of Science and Technology Beijing, Guangdong 528300, China (e-mail: hejie@ustb.edu.cn; qiyuee@ustb.edu.cn).

Hongqing Song is with the School of Civil and Resources Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: songhongqing@ustb.edu.cn).

Yueqiang Sun, Weihua Bai, and Wei Li are with the Beijing Key Laboratory of Space Environment Exploration, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China, and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: syq@nssc.ac.cn; baiweihua@nssc.ac.cn; liwei@nssc.ac.cn).

Qiwu Wu is with the Engineering University of PAP, Xi'an 710086, China (e-mail: wuqiwu700@163.com).

Digital Object Identifier 10.1109/JSTARS.2020.3010879

I. Introduction

MASTERING ocean surface wind has a great significance in climate studies, since sea wind is the main power for upper ocean movement, and it is the critical factor for wave generation, formulation of water masses, and ocean currents [1]. Besides, ocean surface wind has a significant impact on marine production activities, such as ship routing, marine fishing, and offshore resource development [2]. The traditional methods of ocean surface wind measurement include static meteorological station, buoy, and sounding balloon [3]. However, compared with the vast ocean area, the coverage of these methods is too small. Subsequently, the satellite remote sensing methods are developed, such as microwave scatterometer [4] and satellite altimeter [5]. Microwave scatterometer is relatively mature at present, but it does not work well under high wind speed. It is because the sea waves are broken under high wind speed, and the foam on the sea surface reduces the backscattering coefficient. Besides, this method also has the problem of wind direction ambiguity in the retrieval process. Satellite altimeter is a vertical measurement, and it only can detect in one direction. Furthermore, it accepts the expanded backscatter coefficient and has a limited measurement range between 2 and 15 m/s.

As a new remote sensing technology, Global Navigation Satellite System Reflection (GNSS-R) was proposed by European Space Agency scholar Maritin-neira in 1993. GNSS-R can be used to obtain earth-surface parameters such as ocean surface wind [6], ocean ice detection [7] and soil moisture retrieval [8], [9] through signal processing. Compared with the traditional remote sensing technologies, GNSS-R has advantages such as larger space coverage, longer time span, and less affection from the weather. Additionally, GNSS-R works by receiving passive signals and does not send signals by itself, so it has a low power consumption. In July 2014, U.K. TechDemoSate-1(TDS-1) was successfully launched and then provided a lot of measured data for the study of spaceborne GNSS-R [10].

Typically, delay–Doppler map (DDM), which is converted from the received GNSS signals when the incoherent scattering is dominant [11], is the main observable used for GNSS-R-based

retrieval [11]. The traditional methods of GNSS-R ocean wind speed retrieval are two theoretical model-based approaches. The first one is the waveform matching [12], [13]. The theoretical model is used to generate a large number of simulation DDMs by traversing the parameters and the typical cost function, Euclidean distance (ED) between the measured DDM and simulated DDMs are calculated to find the simulated DDM with minimum distance. Finally, the corresponding wind speed generating the simulated DDM with minimum ED would be considered as the retrieval value [14]. Another one is the retrieval based on empirical functions [15], [16]. In general, this method uses a large number of measured data to set up the functional relationship between ocean surface wind speed and one or two physical parameters extracted from DDM. Nevertheless, such a simple model is not robust enough to achieve high retrieval accuracy [17], [18].

At present, machine learning (ML) is becoming more and more popular in the studies on GNSS-R-based retrieval, such as using multilayer perception (MLP) to detect sea ice [19], using support vector regression (SVR) to detect sea ice [20], etc. These ML-based methods are more robust and accurate than the theoretical model-based approaches since it is data-based, and they can model the complicated relationship between the input and the output. DDM is derived from the receivers measurements on the frequency and strength of the received signal. The accuracy of these measurements are closely related to the status of the receivers radio frequency (RF) components, such as low noise amplifier (LNA), antennas, analog-digital (A/D) converters, which compose SRS. Consequently, SRS would affect the quality of the DDM, which is represented by signal to noise ratio (SNR), and thus the performance of the retrieval. Accordingly, SRS parameter should also be used as the input data of ML model to retrieve the ocean wind speed. From the perspective of model input data, ML-based GNSS-R ocean wind speed retrieval can be divided into two types. The first type of input data includes DDM and SRS parameters [21]. The second type of input data includes DDM feature (such as DDMA and LES) manually extracted from raw DDM data and SRS parameters [22], [23]. Though there are more than 30 SRS parameters, the previous research works, no matter the first or the second type, only use less than 10 SRS parameters to compose the input data. With the development of deep learning, the recent ML model architecture based on the automatic feature extraction algorithm and raw data has already achieved better performance than the traditional model architecture using manually extracted features. The recent study [21] on GNSS-R ocean wind speed retrieve also shows a similar result since the features automatically extracted from the raw DDM contains more useful information for retrieve than the manually extracted features. For the retrieval using the manfully extracted DDM features and SRS parameters, ML model only plays the role of regression [22], [23]. For the retrieval using the DDM and SRS parameters, the ML model has to fulfill feature extraction, data fusion, as well as regression. DDM is similar to the image, and SRS parameters are numerical values. Due to the difference in the data structure, specific automatic feature extraction, and data fusion should be considered during model architecture design. However, the typical solution in the previous studies, such as in [21], is to combine the raw DDM and SRS parameters and then put the combination to an NN model for feature extraction and regression.

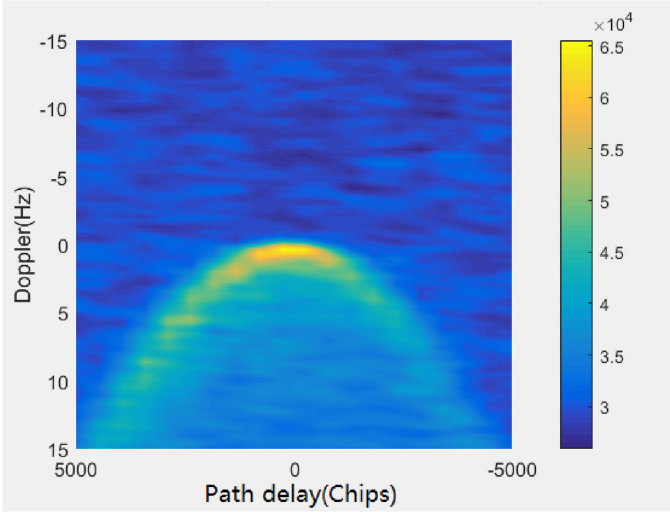


Fig. 1. Typical DDM obtained from TDS-1.

In this article, we first introduce a new input data structure, which is composed of DDM and all SRS parameters, to make the input data more comprehensive. Then, considering the difference in data structure between raw DDM and DDM-related parameters, we designed a heterogeneous multimodal deep learning (HMDL) architecture for GNSS-R ocean wind speed retrieval. HDML is composed of a convolutional neural network (CNN) and two MLPs. CNN is used to extract the feature of DDM. One MLP is for the feature extraction of SRS parameters. Anther MLP is for the regression of the wind speed. Data fusion is fulfilled by concatenating the extracted features between the process of feature extraction and regression. Based on improvements on input data structure and NN model structure, this article achieves state-of-the-art accuracy on GNSS-R ocean wind speed retrieval. The novelty is summarized as follows. First, a new input data structure, which is composed of all DDM-related parameters and raw DDM, is proposed for ML-based GNSS-R ocean wind speed retrieval. The new structure contains more useful information for retrieval than the traditional structure, which is composed of parts of DDM-related parameters and raw DDM or manually extracted DDM features. Second, a new model architecture named HMDL is designed according to the difference in data structure of DDM and DDM-related parameters. HMDL is more effective on automatic feature extractions and data fusion than the traditional NN structure.

II. Problem Description

A. Heterogeneous GNSS-R Data

1) DDM: DDM is a measure of the scattered signal power as a function of path delay and Doppler frequency [21]. A typical example of DDM is illustrated in Fig. 1, and the horizontal represents delay, the ordinate represents Doppler frequency. Each DDM obtained from TDS-1 is the result of noncoherent

TABLE I
ELEMENTS OF SRS

No	1	2	3	4	5	6
name	SPIncidenceAngle	SPElevationORF	SPAzimuthORF	SPElevationARF	SPAzimuthARF	specularPointLat
No	7	8	9	10	11	12
name	specularPointLon	antennaGainTowards-SpecularPoint	DDMSNRAtPeak-SingleDDM	transmitter-VelocityZ	specularPoint-PositionX	specularPath-RangeOffset
No	13	14	15	16	17	18
name	transmitter-PositionX	transmitter-PositionY	Transmitter-PositionZ	transmitter-VelocityX	transmitter-VelocityY	directSignalInDDM
No	19	20	21	22	23	24
name	LNATemperature	ADCPercentage-Positive	ADCOffset	noiseBoxRows	specularPoint-PositionZ	ADCPercentage-Magnitude
No	25	26	27	28	29	30
name	specularPoint-PositionY	antennaGain-RangeMin	antennaGain-RangeMax	antennaTemperature-ExtRef	Antenna-Temperature	kurtosisNoiseBox
No	31	32	33	34		
name	meanNoiseHigh-Doppler	meanNoiseBox	DDMOutput-NumericalScaling	DDMPixelValueNoise		

integration of 1000 consecutive 1-ms. The size of DDM is 20×128, where 20 represents the number of Doppler pixels, and 128 represents the number of delay pixels. A GNSS-R DDM depicts the scattered power of the observed surface [19]. Each pixel in DDM corresponds to scattering intensity in a certain area [19]. There is a horseshoe-shaped waveform in the DDM, and it shows the distribution of scattering signal power in different areas of the sea surface. The wind causes the roughness of the ocean surface. It then results in a reduction in the reflected signal power near the specular point and an increase in the scattered signal away from the specular reflection point. Therefore, DDM implies the affection of the wind speed on the GNSS-R signal and can be used in wind speed retrieval.

2) SRS: The SRS parameters are listed in Table I [11]. The first 30 elements excluding the ninth are related to the parameters of the RF and the position of the receiver. The first nine elements are related to the calculation formula of the scattering coefficient. The last four and the ninth elements are related to DDM directly. The formulation of DDM is associated with the receivers measurement on the received signal, while RF and other parameters influence the accuracy of the measure. Therefore, these parameters should be added to the input data of retrieval.

3) Heterogeneity: The heterogeneity of DDM and SRS is apparent, as DDM is 2-D and similar to an image while SRS is 1-D discrete numerical data. Due to the heterogeneity, data fusion cannot be performed directly on the raw data. Besides, owing to the complexity of DDM, handcrafted feature extraction is hard to extract all valid information, so it will eventually affect the retrieval accuracy. To solve this problem, two automatic feature extractors, which can output the features with the same data structure, are designed for DDM and SRS, respectively. Then, their features are fused to predict the wind speed.

B. Solving Procedure

Wind speed retrieval can be regarded as a regression problem. The input data denoted by $X = X^1 U X^2$,where $X^1 \in R^{20\times128}$ denotes the DDM, $X^2 \in R^n, n = \{8, 9, 29, 34\}$ denotes the SRS, the n represents the number of SRS parameters. The real sea surface wind speed denoted by $Y \in R^m$, where m is the number of training samples. The objective of the proposed method is to drive an optimal function $F(X) = W \cdot X + b$, as shown in 1, where W is the weight vector of the proposed method and b is the bias. Mean square error is selected as our loss function according to the character of the task. L2 norm regularization and dropout were adopted to control overfitting and underfitting. So, W and b can be obtained by optimizing the following formula:

$$(W^*, b^*) = \text{argmin} \sum_{i=1}^{n} (F(x_i - y_i))^2 + \alpha ||W||^2 \qquad (1)$$

where n is the batch size, $||\cdot||$ is Frobenius form, α is regularization parameter, $||W||^2$ is regularization term which can reduce the risk of overfitting.

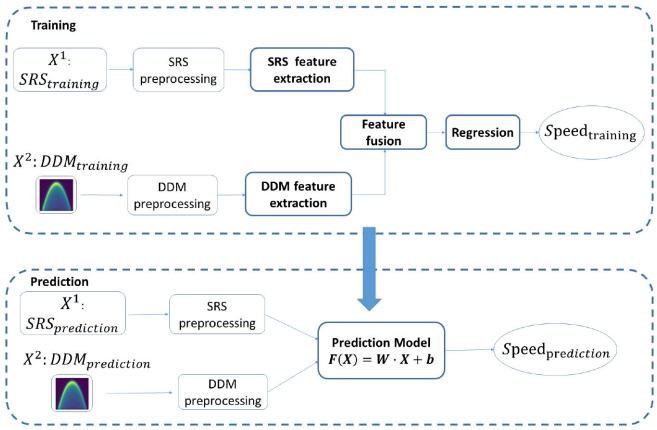


Fig. 2. Framework for heterogeneous GNSS-R data fusion and multimodal learning. Firstly, data preprocessing and feature extraction. Subsequently, fusion of extracted feature. Finally, regression prediction.

The solution we proposed to retrieve wind speed from the heterogeneous GNSS-R raw data is shown in Fig. 2. The proposed solution consists of two parts, training and prediction. The input of the framework contains DDM and SRS parameters. First, two parts go through their respective preprocessing, which makes the input data more suitable to the neural network (NN). The SRS preprocessing achieves normalization. The DDM preprocessing

performs normalization and noise reduction. Next, the preprocessed data passes through their feature extraction networks in parallel and then get their respective extracted features. A specific algorithm will be selected for feature extraction according to the structural characteristic of the input data. After that, the features extracted from DDM and SRS will be fused. Finally, the regression is performed based on the fused features. After completing these operations, we will get a predictive model. In the prediction section, the input data needs to preprocess first, then send preprocessed data into the predictive model and finally get the predicted wind speed.

III. Multimodal Methodology

A. Preprocessing

1) DDM Preprocessing:

a) Data noise reduction: To reduce the noise effect caused by antenna gain and system noise, each DDM will subtract noise floor [19], which can be calculated using the following expression:

$$M_{\text{nosie}} = \frac{1}{N}\sum_{\tau_1}^{\tau_i}\sum_{f_1}^{f_i} M(\tau, f) \tag{2}$$

$$DDM_{\text{denoise}} = DDM_i - M_{\text{noise}} \tag{3}$$

where M_{nosie} represents the noise, τ_1 and τ_i are the limits of delay chips in noise box, f_1 and f_i are the limits of frequency in noise box, N is pixel sum of noise box, DDM_i represents each pixel in DDM.

b) DDM normalization: NNs are more sensitive to data between 0 and 1. So, the DDM is normalized to eliminate the order of magnitudes differences among dimensions and make data have the same statistical distribution. Besides, the normalization usually can improve NN prediction accuracy and accelerate the convergence of the training network. The normalization of DDM can be expressed as follows:

$$M_{\text{normal}}(\tau, f) = M(\tau, f)/M_{\max} \tag{4}$$

where $M_{\text{normal}}(\tau, f)$ represents each normalized pixel value, $M(\tau, f)$ represents the original pixel value, $M_{\max}$ represents the max pixel value in original DDM.

2) SRS Preprocessing:

a) SRS normalization:

SRS is also normalized between 0 and 1. The normalization of SRS can be expressed as follows:

$$M(i)_{\text{normal}} = M(i)/M(i)_{\max} \tag{5}$$

where M_i represents the ith parameter, $M(i)_{\max}$ represents the maximum value of the ith parameter, $M(i)_{\text{normal}}$ represents the parameter that after normalization.

B. NN Architecture

DDM can be regarded as image and SRS are discrete numerical data. Therefore, according to the heterogeneity of the input data, MLP is used to extract the features of SRS, and a CNN is used to extract the features of DDM. The extracted features of SRS and DDM are fused before regression. Finally, another MLP is used as the regression model with the fused features as the input.

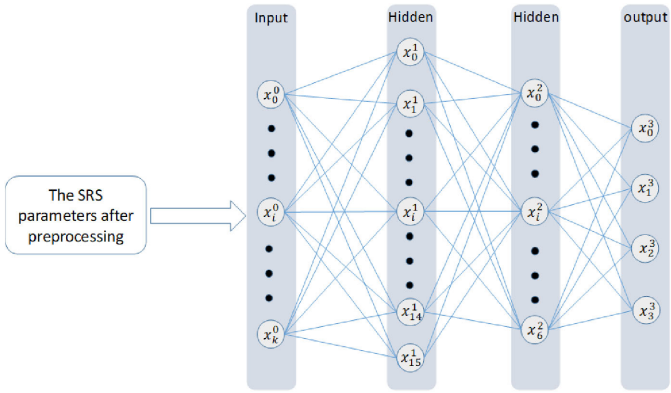


Fig. 3. Framework for SRS feature extraction, the framework consists of three MLPs.

1) MLP-Based SRS Feature Extraction: The SRS features extraction stacks MLP, whose specific network structure is shown in Fig. 3. In the input layer, SRS after preprocessing directly input to the network. Assuming that MLP has K_1 hidden layer. $H_i \in R^{n_i}, i = 1, 2, \ldots k_1$ h_i is the output of each hidden layer, n_i is the neuron number of every hidden layer. In the first hidden layer, parameters $w_1 \in R^{n_1 \times n}$ and $b_1 \in R^{n_1}$ are randomly initialized. The output result of first hidden layer is expressed as

$$h_1 = \sigma(W_1 x + b_1) \tag{6}$$

where σ is the nonlinear activation function. In this article, Relu is adopted as the activation function. The mathematical expression of Relu can be expressed as follows:

$$\text{relu}(x) = \max(0, x). \tag{7}$$

The subsequent hidden layer $i = 2, 3 \ldots k_1$ assuming that their initialization parameters are $w_i \in R^{n_i \times n_{i-1}}$ and $b_1 \in R^{n_i}$. The output result of hidden layer is as follows:

$$h_i = \sigma(w_i h_{i-1} + b_i). \tag{8}$$

2) CNN-Based DDM Feature Extraction: The network structure of DDM feature extraction is shown in Fig. 4. It has $k_2 = 8$ convolution layer and the filter size is 3×3. We use the symbol H_i represent feature map. The preprocessed DDM is the first feature map, represented as

$$H_0 = X \qquad H_0 \in R^{128 \times 20} \tag{9}$$

where x denotes the preprocessed DDM. For subsequent convolution layer $i = 2, 3, \ldots, k2$, assume that their parameters are $W_i \in R^{3\times 3}$, $b_i \in R^3$. The convolution operation can be described as

$$H_i = \sigma(H_{i-1} \otimes W_i + b_i). \tag{10}$$

The symbol $\otimes$ denotes the convolution between filter and input data or the H_{i-1}. The output of the filter is added to the offset b_i

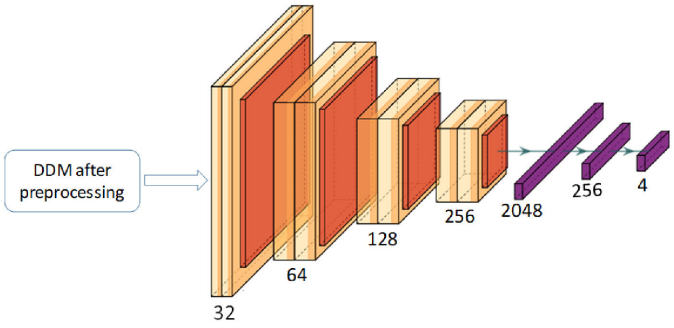


Fig. 4. Framework for DDM feature extraction, the framework consists of eight convolution layer and three MLPs.

of the ith layer. σ represents an activation operation. Common activation functions of CNN are

$$\sigma(x) = \begin{cases} \frac{1}{1+e^{-x}} & \text{sigmoid} \\ \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} & \tanh \\ \max(0,x) & \text{Relu.} \end{cases} \tag{11}$$

In this work, Relu is used as an activation function. Pooling operation follows after convolution operation. Here, we adopt max pooling, which can be expressed as

$$z = \max(x_{i,j}, x_{i,j+1}, x_{i+1,j}, x_{i+1,j+1}). \tag{12}$$

After finishing all the operations mentioned above, the 3-D feature maps will be transformed into 1-D feature maps through the full connection layer, making it easier to join with the features extracted from SRS.

3) MLP-Based Feature Fusion and Regression: In the feature fusion section, the main task is to merge the feature extraction results. Fig. 5 shows the process. After the feature extraction parts, the heterogeneous data is transformed into features with the same data structure. Suppose the result of SRS extraction is O_1, and the result of DDM extraction is O_2. Then, the fused features can be obtained by aggregating these two 1-D feature maps, expressed as

$$O = O_1 \cup O_2 \tag{13}$$

where O is the merged features, and it will input to the final regression network.

The regression network consists of two fully-connected (FC) layers and linear output. O was directly inputted to the input layer. The input layer to the hidden layer can be expressed as

$$z_1 = \sigma(W_1 x + b_1) \tag{14}$$

where $W_1 \in R^{4\times 8}$, $b \in R^4$ O is fusion features and σ is Relu activation function. The final output is a linear combination of the hidden layer and it can be expressed as

$$\hat{y} = W_2 z_1 + b_2 \tag{15}$$

where $W_2 \in R^{14}, b_2 \in R^1$. The $\hat{y}$ can be considered as the predicted wind speed. After obtaining the $\hat{y}$, the loss function measures the deviation between the target wind speed and the predicted wind speed in the training process. According to the error, the back propagation algorithm is used to adjust the network parameters.

Algorithm 1: Pseudo Code of Training Process in HMDL.

Input: training set $D = (X_k, Y)_{k=1}^m$ where m is sample size, $X_k = X_k^1 X_k^2 2$
Output: HMDL with confirmed weight vector and threshold value
1: Initialize weight vector and threshold value of HMDL in the range between 0 and 1
2: **repeat**
3: **for** all $(X_k, Y_k) \in D$ **do**
4: Forward propagate X_k^1 into DDM extraction and get output O_1
5: Forward propagate X_k^2 into SRS extraction and get output O_2
6: Concatenate O_1 and O_2 then get O
7: Forward propagate O into regression network and get final output
8: Calculation preparation: $g_t = \nabla J(\theta)$, $v_t = \beta_1 v_{t-1} + (1-\beta_1) g_t, s_t = \beta_2 s_{t-1} + (1-\beta_2) g_t^2, \hat{V}_t = \frac{v_t}{1-\beta_1^t}, \hat{S}_t = \frac{s_t}{1-\beta_2^t}$
9: Calculate the gradient descent $g_t^{'}$ using $g_t^{'} = \frac{\eta \hat{v}_t}{\sqrt{\hat{s}_t + \epsilon}}$
10: Back propagate parameters of regression network using $\eta_i = \eta_{i-1} - g_t^{'}$
11: Back propagate from feature fusion to each branch and update parameters in each branch using $\eta_i = \eta_{i-1} - g_t^{'}$
12: **until** reaching stop condition

IV. Model Implementation and Experiment Setups

A. Model Implementation

The configurations of HDML are listed in Table II. The number of layers is set referring to AlexNet [27], and then optimized by experiments. Algorithm 1 describes the process of iterative training, in which we use back propagation and Adam optimization algorithm to obtain the best parameters of the network. The learning of the network mainly includes a forward propagation and a back propagation. In the forward propagation, the data is processed layer by layer from the input layer to the output layer. During the back propagation, the parameters in the regression network are first updated, and then the gradient is transferred from feature fusion to two feature extraction branches. Keras and Tensorflow are used to implement our proposed method.

B. Experiment Setups

1) Data Set: In this study, both DDM and SRS used in training and testing are obtained from TDS-1[1]. The reference wind speed employed in this article is derived from the European center for medium weather forecasting (ECMWF)[2]. ECMWF

[1][Online]. Available: www.merrbys.co.uk
[2][Online]. Available: https://apps.ecmwf.int/datasets/

TABLE II
DETAILED IMPLEMENTATION OF HDML

	Architecture	Number oflayer	Kernel/unit size	channels
DDM feature extraction	CONV+CONV	4	(3,3)	64,64,128,128
	pool	4	(2,2)	64,64,128,128
	FC	3	1024,512,4	-
SRS feature extraction	FC	3	8/9/30/34,16,4	-
Feature fusion and regression	FC	3	8,4,1	-

leads the forecasting in the numerical weather prediction, and it provides high precision numerical weather prediction. The data for a month is gathered, and the total amount is 252 926. The ratio of training data to testing data is 7:3. The training data and testing data were selected randomly. Each data was bound a random number between 0 and 1, which is produced by a random function. Then, the data with a random number between 0 and 0.7 was used as training data. The data with the random number between 0.7 and 1 is used as prediction data. The spatial resolution of the reference data is $0.125^\circ \times 0.125^\circ$. The TDS-1 data includes the latitude and longitude, where the data is obtained. The latitude and longitude are used to find its closest location in ECMWF data to fulfill the spatial match. The time resolution of the reference data was 6 h. The wind speed was measured at 00:00 A.M., 06:00 A.M., 12:00 P.M., and 18:00 P.M. every day. The time resolution of TDS-1 data is 1 s. The TDS-1 data located in the following time spans are selected to compose the data set: 23:30 P.M. to 00:30 A.M., 05:30 A.M. to 06:30 A.M., 11:30 A.M. to 12:30 P.M., and 17:30 P.M. to 18:30 P.M.. The TDS-1 data between 23:30 P.M. to 00:30 A.M. was matched to the reference data measured at 00:00 A.M.. The TDS-1 data between 23:30 P.M. to 00:30 A.M. was matched to the reference data measured at 00:00 A.M.. The TDS-1 data between 05:30 A.M. to 06:30 A.M. was matched to the reference data measured at 06:00 A.M.. The TDS-1 data between 11:30 A.M. to 12:30 A.M. was matched to the reference data measured at 12:00 P.M.. The TDS-1 data between 17:30 A.M. to 18:30 A.M. was matched to the reference data measured at 18:00 P.M..

2) Performances: Mean absolute error (MAE) and root mean square error (RMSE) are used to evaluate the statistical accuracy and robustness of the model. The cumulative distribution function (CDF) curve would be the plot to show the overall distribution of the retrieval error. MAE and RMSE are defined as follows:

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}|\hat{y}_i - y_i| \tag{16}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum\nolimits_{i=1}^{m}(\hat{y}_i - y_i)^2} \tag{17}$$

where $\hat{y}_i$ means the wind speed predicted by the model, and y_i means the real ocean wind speed.

SNR is the key index of the quality of DDM. Most previous studies only achieve excellent performance in high-SNR conditions. In order to evaluate the accuracy of the proposed method under different data quality and its advantage on resisting noise, the performance of two partial testing data selected by SNR thresholds, as well as the overall testing data set, would be counted and illustrated. The selection criteria are set at SNR > 3 and SNR > 0, according to the distribution of SNR. The number of data in the overall testing data set is 75 878. The numbers of data in two subsets are 7316 (SNR > 3) and 15801 (SNR > 0), respectively. Thus, there are about 10% SNRs larger than 3 dB, and 20% SNRs larger than 0 dB. The distribution of the SNR of the overall testing data set and two subsets are shown in Fig. 6. As shown in figure, there are about 80% SNR is lower than 0 in the overall testing data set. Therefore, most of the testing data is

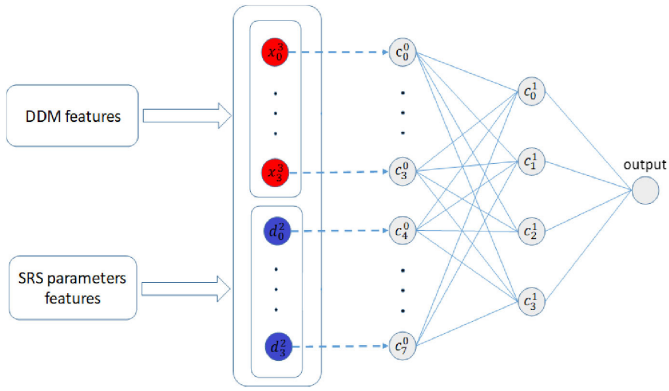


Fig. 5. Framework for feature fusion and regression network, the middle part is feature fusion and the latter part is regression network which bases on MLP.

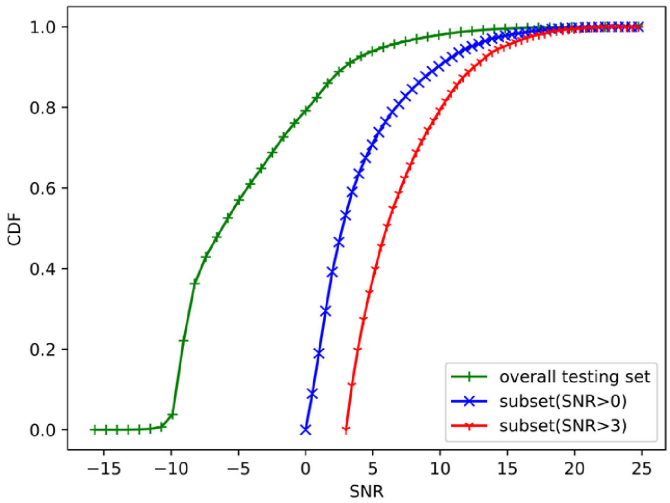


Fig. 6. SNR distribution of the overall testing set, subset (SNR > 0) and subset (SNR > 3).

TABLE III
IMPROVEMENT OF DDM+SRS_{ALL}

Metrics(m/s)	Data	DDM	DDM+ $SRS_{1\sim8}$	DDM+ $SRS_{1\sim9}$	DDM+ SRS_{ALL-5}	DDM+ SRS_{ALL}	Improvement(%)
MAE	Subset(SNR>3)	1.78	1.57	1.60	1.43	1.36	23.6
	Subset(SNR>0)	1.97	1.78	1.75	1.57	1.50	23.9
	Overall testing set	2.39	1.95	1.93	1.90	1.67	30.1
RMSE	Subset(SNR>3)	2.45	2.05	2.15	1.90	1.78	27.3
	Subset(SNR>0)	2.37	2.31	2.28	2.07	1.95	17.7
	Overall testing set	3.10	2.56	2.57	2.55	2.20	29.0

with low SNR, and the prediction accuracy of the overall testing data set represents the ability of noise resistance.

3) Experiments: Verification experiments and comparative experiments are conducted to verify the proposed method. The verification experiments are to show the effectiveness of SRS and HMDL on the improvement of wind speed retrieval. The comparative experiment is to display the improvement of the overall proposed method by comparison with the traditional ML approaches.

a) Verification of DDM + SRS_{ALL}: The baseline of this experiment is the performance of wind retrieval only based on DDM. The ML method of baseline performance is CNN+MLP, in which CNN is for feature extraction, and MLP is for regression. The proposed input data structure can be expressed as DDM + SRS_{ALL}, where SRS_{ALL} means all SRS parameters. It is also compared with the following combinations:

1) DDM + $\text{SRS}_{\text{ALL}-5}$, where $\text{SRS}_{\text{ALL}-5}$ includes all the elements except the 9th, 31st, 32nd, 33rd, 34th, which is related to DDM.
2) DDM + $\text{SRS}_{1\sim9}$, where $\text{SRS}_{1\sim9}$ includes the $1\sim9$ elements, which are used to calculate the DDM from the received signal.
3) DDM + $\text{SRS}_{1\sim8}$, where $\text{SRS}_{1\sim8}$ includes the $1\sim8$ elements.

b) Verification of HMDL: The baseline of this experiment is the performance of a homogeneous multimodal approach, which is named as MLP+MLP+MLP. The first MLP is used to extract the feature of SRS parameters. The second MLP is for the feature extraction of DDM. The last MLP is for wind speed regression. The difference between the homogeneous multimodal architecture and HDML is a feature extraction algorithm of DDM. The input and output of the homogeneous approach are the same as the heterogeneous approach.

c) Comparative experiments: Two experiments are carried out for comparison. Comparison1: The baseline of the experiments is the retrieval based on three typical ML algorithms and DDM. The baseline ML methods include MLP, support vector machine (SVR), and random forest. The input data of these three methods is DDM. The MLP have been used in other GNSS-R related articles [19], [21], [24]. SVR is a modified version of the support vector machine (SVM), which is specifically designed for regression problems [25]. Random forest is a kind of integrated learning, which can improve the accuracy of single prediction [26]. Comparison2: The baseline of this experiment is the typical traditional NN-based GNSS-R-based ocean wind speed retrieval proposed in [21]. We applied the complete solution of [21], including the input data structure and model architecture, on our data set to obtain the baseline performance. The performance of Liu *et al.*'s [21] model architecture with DDM + SRS_{ALL} as the input data is also obtained for comparison.

V. RESULTS AND DISCUSSION

A. Verification Experiments

1) $DDM + SRS_{ALL}$: Table III lists MAE and RMSE of retrieved wind speed using four combinations of DDM+SRS and only DDM. The results indicate that all four combinations of DDM+SRS is better than only DDM in any situation. Besides, Fig. 7 shows the error between prediction value and real value. The curves in the figure also indicate that the error of SRS+DDM is less than only DDM. These experimental results show that SRS provides useful information for wind speed prediction and can improve prediction accuracy and robustness.

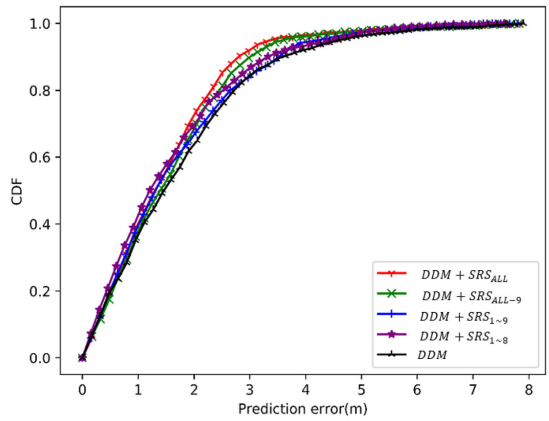
Both Table III and Fig. 7 show that the model achieves better results with more SRS parameters, and the performance of DDM+SRS_{ALL} is the best among the four combinations. DDM+$\text{SRS}_{1\sim9}$ and DDM+$\text{SRS}_{1\sim8}$ were differed by one parameter. The performance of DDM+$\text{SRS}_{1\sim9}$ is a little better than DDM+$\text{SRS}_{1\sim8}$. DDM+SRS_{ALL} and DDM+$\text{SRS}_{\text{ALL}-5}$ were differed by the ninth parameter. DDM+SRS_{ALL} is better than DDM+$\text{SRS}_{\text{ALL}-5}$. These results indicate that each parameter of SRS is valuable during the wind speed prediction.

With the decrease of SNR, the predictive power of all methods has declined. However, the downward trend of the proposed method that uses DDM+SRS is slower than the method only using DDM. From the figure, we can get that more SRS elements are involved, the slower the accuracy declines. This demonstrates that SRS enhances the resistance to noise.

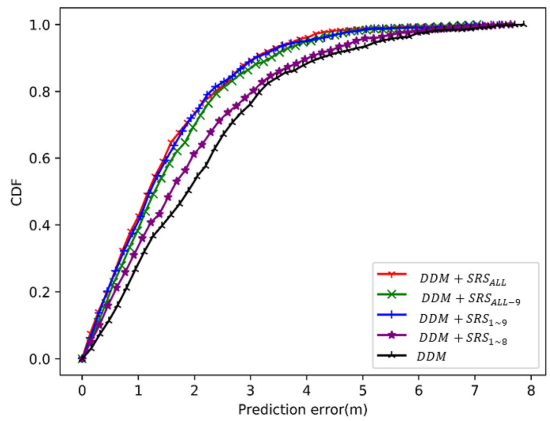
2) Heterogeneous Multimodal Versus Homogeneous Multimodal: Table IV lists the performance of heterogeneous multimodal and homogeneous multimodal. The results indicate that heterogeneous multimodal is better than homogeneous multimodal in any SNR cases. Fig. 8 shows the cumulative probability distribution of the prediction error of two methods. There is always a gap between two curves, which shows that the prediction effect of heterogeneous multimodal is always better than homogeneous multimodal. The curves change in the figure conforms to the data changes in the table. The difference between heterogeneous multimodal and homogeneous multimodal in network structure is that homogeneous multimodal uses MLP in DDM feature extraction while heterogeneous multimodal use CNN. Heterogeneous multimodal uses original, while homogeneous multimodal flattens the DDM input into a vector without

TABLE IV
MAE AND RMSE COMPARISON: HETEROGENEOUS MULTIMODAL VERSUS HOMOGENEOUS MULTIMODAL

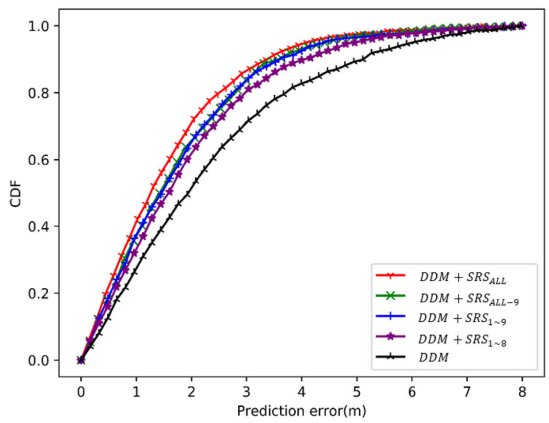
Metrics(m/s)	Data	Homogeneous multimodal (MLP+MLP+MLP)	Heterogeneous multimodal (MLP+CNN+MLP)	Improvement(%)
MAE	Subset(SNR>3)	1.44	1.36	5.6
	Subset(SNR>0)	1.55	1.50	3.2
	Overall testing set	1.81	1.67	7.7
RMSE	Subset(SNR>3)	1.85	1.78	3.8
	Subset(SNR>0)	2.00	1.95	2.5
	Overall testing set	2.37	2.20	7.1



(a)

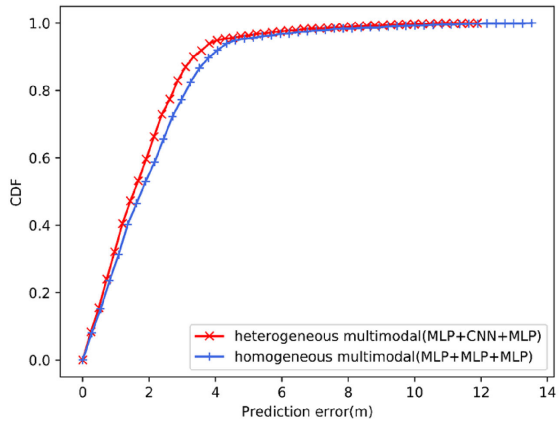


(b)

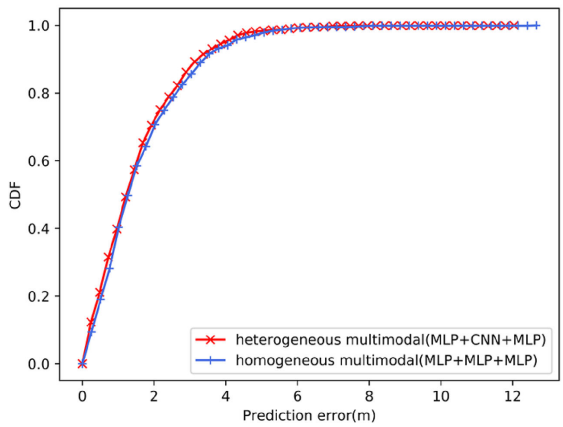


(c)

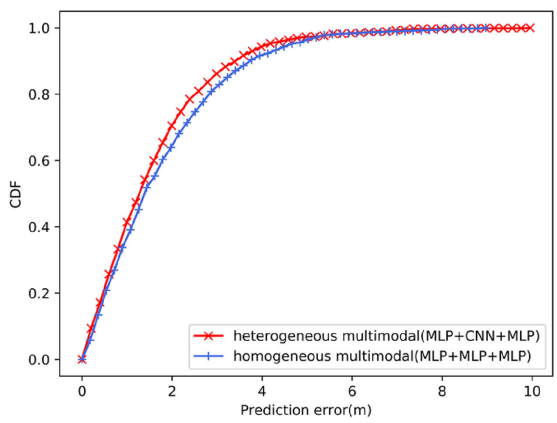
Fig. 7. CDF comparison: DDM+SRS versus DDM. (a) Subset (SNR $>$ 3). (b) Subset (SNR $>$ 0). (c) Overall testing set.



(a)



(b)



(c)

Fig. 8. CDF comparison: Heterogeneous multimodal versus homogeneous multimodal. (a) Subset (SNR $>$ 3). (b) Subset (SNR $>$ 0). (c) Overall testing set.

TABLE V
MAE AND RMSE COMPARISON: HMDL VERSUS ML METHODS

Metrics(m/s)	Data	MLP with DDM	Random Forest with DDM	SVR with DDM	HMDL with DDM+SRS_{ALL}	Improvement(%)
MAE	Subset(SNR>3)	1.70	1.47	1.71	1.36	7.0∼20.5
	Subset(SNR>0)	1.78	1.79	1.80	1.50	15.7∼16.7
	Overall testing set	1.93	2.41	2.33	1.67	13.5∼30.7
RMSE	Subset(SNR>3)	2.25	1.97	2.29	1.78	9.6∼22.3
	Subset(SNR>0)	2.37	2.35	2.35	1.95	17.0∼17.7
	Overall testing set	2.46	3.12	3.00	2.20	10.6∼29.5

the structural information. Besides, CNN used in heterogeneous multimodal has stronger data mining ability and extracted more useful features from DDM than MLP used in homogeneous multimodal [28]. Therefore, the homogeneous multimodal is more resistant to noise and improve the prediction ability. The structure of heterogeneous multimodal is good at resisting noise and mining hidden information.

B. Comparative Experiments

Table V lists the results of the proposed method and the other three ML methods on MAE and RMSE. The results indicate that the proposed method yields the smallest MAE and RMSE. Fig. 9 shows the prediction error of the four methods. As we can see that the proposed method gets the best prediction error all the time and less affected by SNR, while the other three ML methods are greatly affected by SNR. In the overall testing set, the proposed method improves the prediction accuracy by 30.7% on MAE and 29.5% on RMSE comparing with random forest that has the worst performance among the three algorithms. The proposed method improves the prediction accuracy by 13.5% on MAE and 10.6% on RMSE comparing with MLP that has the best performance among the three algorithms. The comparison experiments show that the heterogeneous model and heterogeneous data help model achieve better prediction error and are superior to the other three ML methods in resistance to noise.

With the decrease of SNR, the predictive accuracy of HMDL and other typical methods are declined, but HMDL achieves the least affection. It indicates that heterogeneous deep learning model is more resistant to noise than the traditional ML approaches. Table VI lists the performance comparison between our solution and Liu *et al.*'s [21] solution. The new input data structure improves the prediction accuracy at 21.4% on MAE and 23.6% on RMSE. The proposed whole solution improves the prediction accuracy at 36.3% on MAE and 36.8% on RMSE.

C. Effect of Data Distribution and Wind Speed on Retrieval Accuracy

To analyze the effect of data distribution and wind speed on retrieval accuracy, we counted the number of the samples and computed the average absolute error at different wind speeds regimes, as shown in Fig. 10. It shows that the accuracy is closely related to the number of the samples, but not related to the wind speed. Therefore, more data will be downloaded to compose a balanced data set. Further analysis on a larger data set will be conducted, as a subject of future research, to consolidate the obtained conclusions and the adequacy of the model.

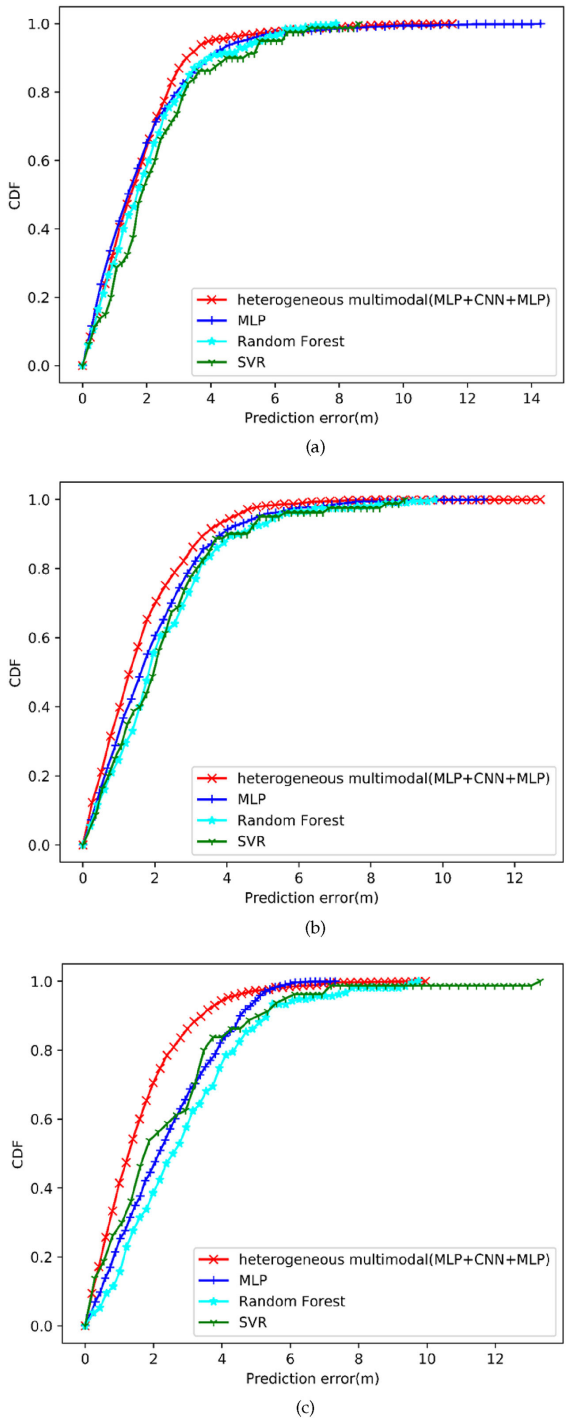


Fig. 9. CDF comparison: Heterogeneous multimodal versus ML methods. (a) Subset (SNR $>$ 3). (b) Subset (SNR $>$ 0). (c) Overall testing set.

TABLE VI
MAE AND RMSE COMPARISON: HMDL WITH DDM+SRS_{ALL} VERSUS TRADITIONAL NN-BASED SOLUTION

Metrics	Data	Traditional NN based solution[19]	Traditional NN[19] with DDM + SRS_{ALL}	DDM+SRS_{ALL} Improvement(%)	HMDL with DDM+SRS_{ALL}	Total Improvement(%)
MAE	Subset(SNR>3)	1.89	1.48	21.7%	1.36	28.0%
	Subset(SNR>0)	2.05	1.83	10.7%	1.5	26.8%
	Overall testing set	2.62	2.06	21.4%	1.67	36.3%
RMSE	Subset(SNR>3)	2.47	1.95	21.1%	1.78	27.9%
	Subset(SNR>0)	2.71	2.34	13.7%	1.95	28.0%
	Overall testing set	3.48	2.66	23.6%	2.2	36.8%

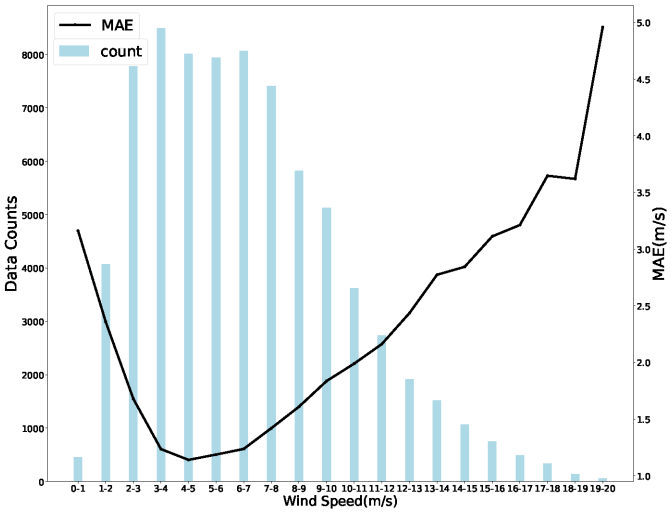


Fig. 10. MAE of HMDL versus wind speed (right vertical axis), along with the data histogram in the background (left vertical axis).

VI. CONCLUSION AND FUTURE WORK

In this article, we first introduced all SRS parameters to retrieve ocean wind speed. Then, we present a data-driven retrieve method HMDL, which is composed of MLP-based SRS feature extraction, CNN-based DDM feature extraction, MLP-based feature fusion, and regression. The new input data structure improves the prediction accuracy at 13.5% to 30.7% on MAE and at 10.6% to 29.5% on the RMSE. The HMDL improves the prediction accuracy by at 7.7% on MAE and 7.1% on RMSE. The whole proposed solution improves the prediction accuracy at 36.3% on MAE and 36.8% on RMSE, compared with the traditional NN-based solution. Notably, the comparison among the experimental results of three different SNR conditions shows that the accuracy improvement of the proposed method increase as the SNR decreases. Therefore, the results of verification experiments indicate that the proposed method effectively enhances the ability of noise resistance and then achieves the state-of-the-art accuracy on GNSS-R-based ocean wind speed retrieval.

Further analysis on a larger data set will be conducted, as a subject of future research, to consolidate the obtained conclusions and the adequacy of the model. The research on the data balancing method for GNSS-R retrieval will be conducted as another subject of future research. As illustrated in this article, the data comprehensiveness and retrieve method are both critical to the prediction performance. We also plan to introduce the historical data and propose a corresponding deep learning model to improve the prediction accuracy. Moreover, we will apply the proposed method to the other GNSS-R-based retrievals, such as ocean ice retrieval. Another important work is to improve the accuracy of the wind speeds regimes with a small number of samples.

REFERENCES

[1] E. J. Metzger, "Upper ocean sensitivity to wind forcing in the South China Sea," *J. Oceanogr.*, vol. 59, no. 6, pp. 783–798, 2003.
[2] M. P. Clarizia, C. S. Ruf, P. Jales, and C. Gommenginger, "Spaceborne GNSS-R minimum variance wind speed estimator," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6829–6843, Nov. 2014.
[3] C. Ruf *et al.*, "CYGNSS: Enabling the future of hurricane prediction [remote sensing satellites]," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 52–67, Jun. 2013.
[4] F. T. Ulaby, R. K. Moore, and A. K. Fung, "Microwave remote sensing: Active and passive, Volume 3," in *Volume 3-From Theory to Applications*. Artech House: Norwood, MA, USA, 1986.
[5] L. C. Smith, "Satellite remote sensing of river inundation area, stage, and discharge: A review," *Hydrological Processes*, vol. 11, no. 10, pp. 1427–1439, 1997.
[6] G. Foti *et al.*, "Spaceborne GNSS reflectometry for ocean winds: First results from the UK TechDemoSat–1 mission," *Geophysical Res. Lett.*, vol. 42, no. 13, pp. 5435–5441, 2015.
[7] Q. Yan and W. Huang, "Spaceborne GNSS-R sea ice detection using delay-Doppler maps: First results from the UK TechDemoSat-1 mission," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 10, pp. 4795–4801, Oct. 2016.
[8] N. Rodriguez-Alvarez *et al.*, "Soil moisture retrieval using GNSS-R techniques: Experimental results over a bare soil field," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3616–3624, Nov. 2009.
[9] H. Carreno-Luengo, S. T. Lowe, C. Zuffada, S. Esterhuizen, and S. Oveisgharan, "Spaceborne GNSS-R from the SMAP mission: First assessment of polarimetric scatterometry over land and cryosphere," *MDPI Remote Sens.*, vol. 9no. 4, 2017, Art. no. 362.
[10] P. Jales and M. Unwin, "Mission description-GNSS reflectometry on TDS-1 with the SGR-ReSI," Surrey Satellite Technol., Ltd., Guildford, UK, Tech. Rep. SSTL Rep, 2015, Art. no. 248367.
[11] P. Jales and M. Unwin, "MERRByS product manual: GNSS reflectometry on TDS-1 with the SGR-ReSI," Surrey Satellite Technol., Ltd., Guildford, UK, 2015.
[12] V. U. Zavorotny and A. G. Voronovich, "Scattering of GPS signals from the ocean with wind remote sensing application[J]," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 951–964, Mar. 2000.
[13] J. Garrison, A. Komjathy, V. Zavorotny, and S. J. Katzberg, "Wind speed measurement using forward scattered GPS signals," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 1, pp. 40–65, Jan. 2002.
[14] N. Rodriguez-Alvarez and J. L. Garrison, "Generalized linear observables for ocean wind retrieval from calibrated GNSS-R delay–Doppler maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 1142–1155, Feb. 2016.
[15] M. Unwin, P. Jales, J. Tye, C. Gommenginger, G. Foti, and J. Rosello, "Spaceborne GNSS-reflectometry on TechDemoSat-1: Early mission operations and exploitation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 10, pp. 4525–4539, Oct. 2016.

[16] N. Rodriguez-Alvarez, D. M. Akos, V. U. Zavorotny, J. A. Smith, A. Camps, and C. W. Fairall, "Airborne GNSS-R wind retrievals using delay–Doppler maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 626–641, Jan. 2013.
[17] M. P. Clarizia and C. S. Ruf, "Wind speed retrieval algorithm for the Cyclone Global Navigation Satellite System (CYGNSS) mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4419–4432, Aug. 2016.
[18] S. Gleason and C. Ruf, "Overview of the delay Doppler mapping instrument (DDMI) for the cyclone global navigation satellite systems mission (CYGNSS)," in *Proc. IEEE MTT-S Int. Microw. Symp.*, 2015, pp. 1–4.
[19] Q. Yan, W. Huang, and C. Moloney, "Neural networks based sea ice detection and concentration retrieval from GNSS-R delay-Doppler maps," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3789–3798, Aug. 2017.
[20] Q. Yan and W. Huang, "Detecting sea ice from techdemosat-1 data using support vector machines with feature selection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 5, pp. 1409–1416, May 2019.
[21] Y. Liu, I. Collett, and Y. J. Morton, "Application of neural network to GNSS-R wind speed retrieval," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9756–9766, Dec. 2019.
[22] J. Reynolds, M. P. Clarizia, and E. Santi, "Wind speed estimation from CYGNSS using artificial neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 708–716, 2020.
[23] M. Asgarimehr, I. Zhelavskaya, G. Foti, S. Reich, and J. Wickert, "A GNSS-R geophysical model function: Machine learning for wind speed retrievals," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 8, pp. 1333–1337, Aug. 2020.
[24] Q. Yan and W. Huang, "Sea ice sensing from GNSS-R data using convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 10, pp. 1510–1514, Oct. 2018.
[25] H. Drucker *et al.*, "Support vector regression machines," in *Proc. Adv. Neural Inf. Process. Syst.*, 1997, pp. 155–161.
[26] A. Liaw and M. Wiener, "Classification and regression by random forest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
[27] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," NIPS. Curran Associates, Inc., 2012.
[28] S. B. Driss *et al.* "A comparison study between MLP and convolutional neural network models for character recognition," *Real-Time Image Video Process. Int. Soc. Opt. Photon.*, 2017, Art. no. 1022306.

Xiaohan Chu is currently working toward the Graduate degree with the Beijing University of Science and Technology, Beijing, China.

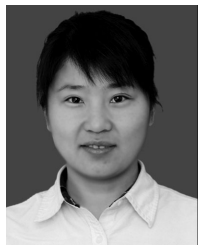
His current research interests include machine learning.

Jie He received a B.E. and Ph.D. degrees in computer science from the University of Science and Technology Beijing (USTB), Beijing, China, in 2005 and 2012, respectively.

He is currently a Professor with the School of Computer and Communication Engineering, USTB, where he was an Associate Professor from July 2017 to June 2020, an Assistant Professor from January 2015 to June 2017, and a Postdoctoral Researcher from January 2013 to December 2014. From April 2011 to April 2012, he was a Visiting Ph.D. Student with the Center for Wireless Information Network Studies, Department of Electrical and Computer Engineering, Worcester Polytechnic Institute. From September 2019 to August 2020, he was a Visiting Researcher with the School of Electrical and Electronic Engineering, Nanyang Technological University. His research interests include machine learning for the internet of things, indoor positioning, and big data.

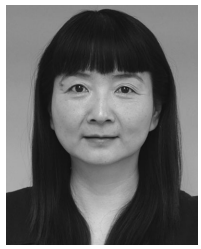
Hongqing Song received the B.S. degree in chemical machinery from Dalian University of Technology, Dalian, China, in 2004, and the Ph.D. degree from the University of Science and Technology Beijing, Beijing, China, majoring in fluid mechanics (master and doctoral program).

He is currently a Professor with the University of Science and Technology Beijing and the Deputy Director of National and Local Joint Engineering Lab for Big Data Analysis and Computing Technology. His research interests include big data analysis and application in area of energy and environment.



Yue Qi received the B.S. degree in computer science from the University of Science and Technology Beijing, Beijing, China, in 1998, and the Ph.D. degree in computer application from the University of Science and Technology Beijing, Beijing, China, in 2007.

His past and present interest involves computer architecture, artificial intelligence, and IC design.

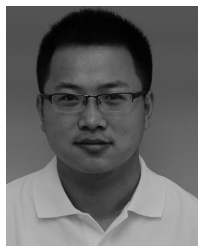


Yueqiang Sun is currently a Professor with the National Space Science Center, Chinese Academy of Sciences, Beijing, China.

Her research interest includes GNSS radio occultation, GNSS-R remote sensing techniques, atmospheric physics, ionospheric physics, magnetic physics, space physics, etc.

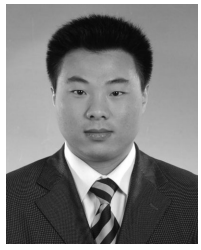
Weihua Bai is currently a Professor with the National Space Science Center, Chinese Academy of Sciences, Beijing, China.

His current research interests concentrate on edge computing, GNSS radio occultation, and GNSS-R remote sensing techniques.



Wei Li is currently a Senior Engineer with the National Space Science Center, Chinese Academy of Sciences, Beijing, China.

His past and present interest involves on GNSS radio occultation and GNSS-R remote sensing techniques.



Qiwu Wu received the B.E. degree in Computer Science and Technology from Hunan Normal University, Hunan, China, in 2005, and the Ph.D. degree in communication and information systems from the University of Science and Technology Beijing, Beijing, China, in 2010.

He is currently an Associate Professor with Engineering University of PAP, Xi'an, China. His research interests include the Internet of Things, secure communications, and optical networks.