


Applications of Deep Learning-Based Super-Resolution for Sea Surface Temperature Reconstruction

Bo Ping , Fenzhen Su, Xingxing Han, and Yunshan Meng

Abstract—Deep learning-based super-resolution (SR) methods have been widely used in natural images; however, their applications in satellite-derived sea surface temperature (SST) have not yet been fully discussed. Hence, it is necessary to analyze the validity of deep learning-based SR methods in SST reconstruction. In this study, an SR model, including multiscale feature extraction and multireceptive field mapping, was first proposed. Then, the proposed model and four other existing SR models were applied to SST reconstruction and analyzed. First, compared with the bicubic interpolation method, the SR models can improve the reconstruction accuracy. Compared with four other SR models, the proposed model can achieve the lowest mean squared error (MAE) in the East China Sea (ECS), in the northwest Pacific (NWP) and in the west Atlantic (WA), the second-lowest MAE in the southeast Pacific (SEP); the lowest root mean squared error (RMSE) in ECS and WA, the second-lowest RMSE in NWP and SEP. Additionally, ODRE model can acquire the highest or the second-highest peak single-to-noise ratio and structural similarity index in ECS, NWP, and SEP. Moreover, the number of missing pixels and SST variety are two essential factors in the SR performance. The proposed multiscale feature extraction process can enhance the SR performance, especially for small regions and stable SST regions. Finally, while a deeper network can be helpful in achieving SR performance, the approach of simply adding more dilation convolutions may not enhance the reconstruction accuracy.

Index Terms—Advanced microwave scanning radiometer 2 (AMSR2), deep learning, moderate-resolution imaging spectroradiometer (MODIS), sea surface temperature, super-resolution.

I. INTRODUCTION

SEA surface temperature (SST) is a significant parameter for analyzing the exchange of energy, momentum, and moisture between the oceans and the atmosphere [1], [2]. With the devel-

Manuscript received August 31, 2020; revised October 15, 2020 and November 8, 2020; accepted December 1, 2020. Date of publication December 3, 2020; date of current version January 6, 2021. This work was supported in part by the Natural Science Foundation of Tianjin under Grant 18JCQNJC08900 and in part by the State Key Laboratory of Resources and Environmental Information System. (Corresponding author: Bo Ping.)

Bo Ping and Xingxing Han are with the School of Earth System Science, Institute of Surface-Earth System Science, Tianjin University, Tianjin 300072, China (e-mail: pingbo@tju.edu.cn; hanxingxing@tju.edu.cn).

Fenzhen Su is with LREIS, Institute of Geographic Sciences and Natural Resources Research, University of the Chinese Academy of Sciences, Beijing 100101, China (e-mail: sufz@lreis.ac.cn).

Yunshan Meng is with National Marine Data and Information Service, Tianjin 300171, China (e-mail: mengys@lreis.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2020.3042242

opment of remote sensing techniques, satellite-derived SST data, including microwave-based SST data and infrared-based SST data, have become an effective way to monitor the global SST. While microwave-based SST data have a lower resolution than infrared-based SST data (25 km for microwave-based SST data compared to 1–4 km for infrared-based SST data), they have relatively complete coverage since microwaves can penetrate clouds. Consequently, taking advantages of these two types of SST data to synthesize SST with high spatial resolution and complete coverage is significant for the long-term monitoring of oceanic features in detail.

The missing data reconstruction methods [3]–[5] that mainly use the spatial information and spatio-temporal correlation are commonly applied for infrared-based SST reconstruction. Recently, some statistical methods have been used for fusing microwave-based and infrared-based SST data. Guan and Kawamura [6] blended multisource SST data including Advanced Very high resolution radiometer (AVHRR), stretched-visible infrared spin scan radiometer, tropical rainfall measuring mission, microwave imager, and visible and infrared scanner to generate 0.05° daily cloud-free SST products by using an objective analysis method, and Wang *et al.* [7] used a Kalman filter to combine infrared-based and microwave-based SST data, while Li *et al.* [8] combined the moderate-resolution imaging spectroradiometer (MODIS) and the advanced microwave scanning radiometer for EOS (AMSR-E) SST data to synthesize 8d average and spatially continuous SST products by using the Bayesian maximum entropy method. Reynolds and Smith [9] analyzed *in situ* and corrected satellite-derived SST data both weekly and daily by using the optimal interpolation (OI) method on a 1° latitude and longitude spatial grid. Then, Reynolds *et al.* [10] proposed an enhanced version of the OI analysis to reduce the satellite bias. Moreover, Reynolds *et al.* [11] developed two high-resolution SST analysis products based on AVHRR and AMSR data by using the OI method. However, some limits still exist in most of these methods. For example, the Kalman filter method needs a large number of calculations that can reduce its efficiency and the objective analysis method may generate singular values [12]. While the OI method has fewer calculations than the Kalman filter method and has been widely used in SST fusion, it also needs *a priori* knowledge, such as background field information, which can increase its calculation complexity. Currently, a number of SST products based on SST fusion methods have been developed, such as the US National Oceanic and Atmospheric

Administration (NOAA) OISST products [11], NOAA RTF-HR products [13], the Japan Meteorological Agency merged SST dataset [14] and the Met Office's Operational Sea Surface Temperature and Sea Ice Analysis products [15]. Generally, the spatial resolution of these SST products ranges from 0.05° to 0.25° and is coarser than that of original infrared-based SST data. Therefore, synthesizing high resolution SST products with complete coverage at low complexity based on infrared-based and microwave-based SST data is necessary.

Due to the similar spatial distributions of SST in low- and high-resolution maps, the super-resolution (SR) may be a new way to generate high spatial resolution SST maps with complete coverage based on infrared-based and microwave-based SST datasets. In the computer vision community, many SR methods have been proposed, such as neighbor embedding [16], [17], sparse coding [18], [19], and random forests [20]. With the renaissance of the neural network, deep learning methods have exhibited their superior performance in SR and are more recognized. Dong *et al.* [21] first introduced a deep learning-based SR method called SRCNN containing only three convolution layers. The relationship between SRCNN and sparse coding SR methods was also discussed. Because the input of the SRCNN is a bicubic-interpolated image that enhances the computational cost and time, Dong *et al.* [22] proposed the FSRCNN method that uses the deconvolution layer to reduce the computational cost. Kim *et al.* [23] proposed the very deep SR network (VDSR) method that contains 20 weight layers to improve the SR performance. Two conclusions can be drawn from the VDSR method: 1) the capacity of deep learning-based SR can benefit from deeper networks; and 2) the residual learning framework [24] can be useful in solving the gradient explosion/vanishing problem. Compared to the shallow CNN-based SR methods, such as SRCNN, a deeper architecture can enhance the receptive field, which means more information can be used for SR. However, a deeper architecture will boost the number of learnable parameters and may cause an overfitting problem. To solve this problem, Kim *et al.* [25] proposed the deeply recursive convolutional network method, which employs the recursive network and weight sharing strategy to enhance the receptive field and subsequently decrease the number of learnable parameters. Tai *et al.* [26] proposed the deep recursive residual network (DRRN) method that employs global and local residual learning to alleviate the difficulty of training the deep architecture and explores recursive learning to increase the receptive field while limiting the model parameters. Chang and Luo [27] incorporated the bidirectional long short term memory (LSTM) into a recursive network for remote sensing image SR. In addition to the recursive network, dilation convolution is another effective way to enhance the receptive field. Zhang *et al.* [28] presented symmetrical dilated residual convolution network (FDSR) that uses the dilation convolution to enhance the receptive field while employing the symmetrical skip connection to alleviate the gradient vanishing problem. The dilation convolution and symmetrical skip connection are also verified in stripe and cloud recovery for remote sensing images [29]. Additionally, the convolutional encoder-decoder network with symmetric skip connections (RED-Net) proposed by Mao *et al.* [30] uses the convolutional encoder layers to extract the features and eliminate

the noises and explores the corresponding decoder layers to recover the details. The encoder layers and their corresponding decoder layers are linked using symmetric skip connections. A similar architecture is also used for skin cloud elimination for remote sensing images [31].

Despite that many SR networks have been proposed and applied to natural images and remote sensing images, they are not fully analyzed for satellite-derived SST. Moreover, in contrast to the training process in the traditional SR networks, the low- and high-resolution data used in SST SR are obtained from microwave-based and infrared-based SST datasets instead of downscaling the high-resolution images to acquire the corresponding low-resolution images. Hence, the correlation between low- and high-resolution data is relatively low. To verify the validity of SR models in SST reconstruction and to analyze the main factors in SR performance for SST reconstruction, in this study, an SR network is proposed for SST data, called the oceanic data reconstruction (ODRE) network, which takes advantage of the existing SR networks. Then, the validity of the proposed network and four other remarkable SR networks including FDSR, DRRN, SRCNN, and VDSR is discussed. Finally, some details about the proposed network are also analyzed. The proposed SR model may become a new way for SST reconstruction.

In the remainder of this article, the experimental data will be introduced in Section II; then the proposed ODRE SR model will be described in Section III; the experimental results will be analyzed in Section IV; the details of ODRE model will be discussed in Section V; finally Section VI concludes this article.

II. DATA

The 3-day averaged $0.25^\circ \times 0.25^\circ$ grid advanced microwave scanning radiometer 2 (AMSR2) and daily 4-km L3 mapped MODIS Terra SST data from 01/01/2013 to 12/31/2019 were selected as low- and high-resolution maps. These two datasets can be downloaded from remote sensing systems¹ and NASA OceanColor Web.² The averaged AMSR2 data were selected to reduce the influence of missing data caused by rain, bad observations and navigation gaps. To keep temporal and spatial consistency, the daily MODIS data were first transformed into 3-day averaged data using data two days before target data and target data itself; and then a median filtering was used to reduce the influence of possible abnormal SST values. In this study, the SST data from 2013 to 2018 were used to train the proposed and other four SR models and the data in 2019 were explored to assess the trained models. The AMSR2 and MODIS data on the same day can be deemed as a low- and high-resolution image pair. The total number of image pairs for training the models is 2 181 due to the nonexistence of AMSR2 data from 05/11/2013 to 05/13/2013 and MODIS data from 02/19/2016 to 02/24/2016 and 12/15/2016. The number of image pairs for testing the models is 365.

The AMSR2 data values between 0 and 250 can be transferred into meaningful SST values and the MODIS data can acquire the SST values directly. Hence, the AMSR2 data values larger than

¹[Online]. Available: <http://www.remss.com/>

²[Online]. Available: <https://oceancolor.gsfc.nasa.gov/>

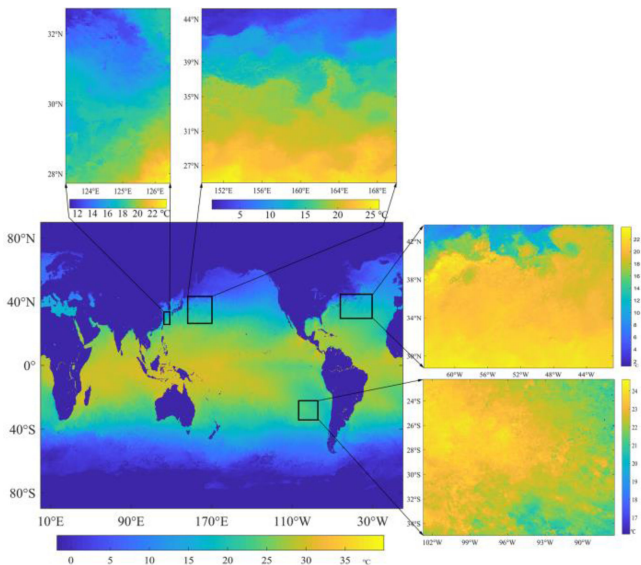


Fig. 1. Four selected study areas including the ECS, NWP, WA, and SEP. The sea surface temperature map is monthly MODIS Terra SST product in January 2020.

250 and MODIS data with Nan values were deemed as missing data or land pixels in this study. There are two preprocesses including: 1) the scale factor between AMSR2 and MODIS data is 6, so the AMSR2 data were upscaled to the same size of MODIS data using the bicubic interpolation method; and 2) there is a spatial offset of 180° in longitude between AMSR2 and MODIS data, so the MODIS data were shifted 180° in longitude. Training image pairs were split into 40 by 40 patches with a stride of 40 and the patches with missing data were excluded, so the total number of training patches is 408536. All the training patches and testing images were normalized into [0 1] to follow the requirements of Caffe. The total training patches were divided by 9:1 into training set and validation set, so the number of patches in the training set and validation set is 367680 and 40856, respectively.

At the test stage, as shown in Fig. 1, four study areas were selected to assess the proposed and other four trained models. The first one is in the East China Sea, ranging from 123.27° – 126.48° E, 27.77° – 32.73° N, which is a marginal sea east of China; the second one is in the northwest Pacific (NWP), ranging from 149.77° – 169.98° E, 25.02° – 45.23° N; the third one is in the west Atlantic (WA), ranging from 40.77° – 63.73° W, 28.77° – 43.73° N; the last one is in the southeast Pacific (SEP), ranging from 87.52° – 102.73° W, 22.27° – 34.98° S. The NWP and WA were selected because of their relatively large SST varieties, and moreover the marginal sea area (ECS) and area covered by stable SST (SEP) were also selected to evaluate the trained SR models. Due to the probable missing pixels in the AMSR2 data, the linear interpolation method was used first to complete the AMSR2 SST.

Fig. 2 shows the distributions of percentages of missing data in the four selected study areas. Overall, the average percentages of missing data of AMSR2 and MODIS datasets are approximately 1.18% and 47.40%, respectively. In ECS, the average

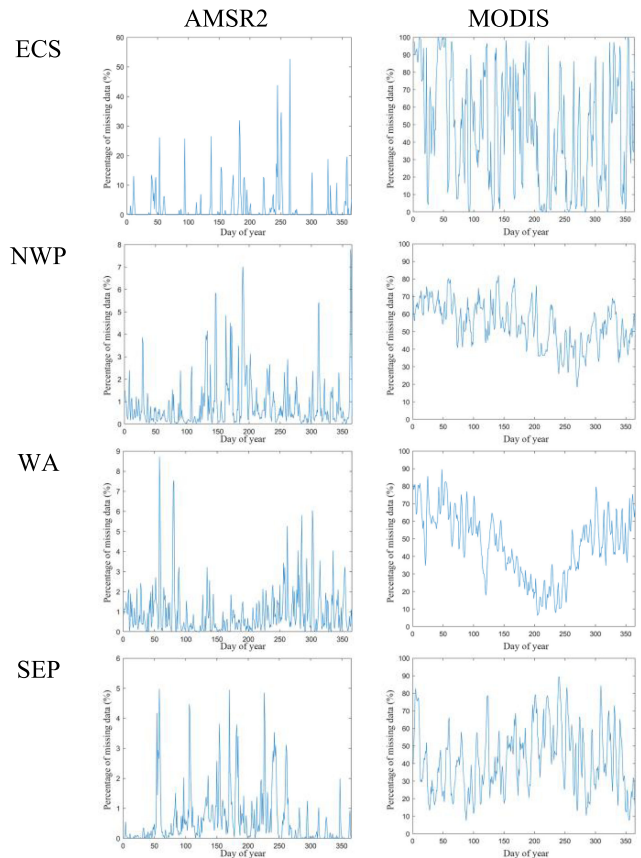


Fig. 2. Distributions of percentages of missing data in the four selected study areas.

percentages of missing data of AMSR2 and MODIS datasets are approximately 2.26% and 46.90%; in NWP, these two values are 0.89% and 54.19%; in WA, these two values are 0.94% and 46.74%; in SEP, these two values are 0.61% and 41.77%. Because of the influence of microwave radiation from the land, in the marginal sea area, such as ECS, more missing pixels can be found in the AMSR2 SST data, which may affect the performance of the SR model.

III. ODR NETWORK

A. Fundamental Theory of CNNs

The CNN, one of the most popular networks in deep learning, has been widely employed in many remote sensing research fields, such as image classification [32], [33] and image fusion [34]–[38]. Three important architecture ideas including the local receptive field, weight sharing and subsampling, enable CNNs to achieve shift, scale and distortion invariant properties [34]. However, these three ideas are not all helpful in SR. The subsampling that is often fulfilled by using a pooling function is harmful to SR [39], so the subsampling is not discussed in this study. It is time-consuming, even impractical, to calculate a neuron value within the objective layer based on all neurons from its previous layer, so in CNNs, the neuron value is computed only using its neighboring neuron values from its previous layer. Therefore, the local receptive field is introduced and defined as the region in the

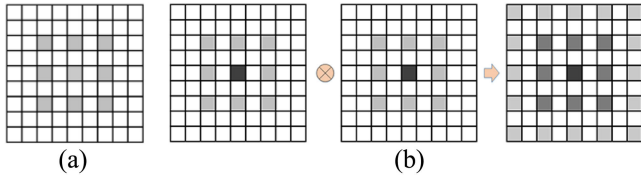


Fig. 3. Two examples of a dilated filter. (a) Single dilated convolution with a dilation factor of 2. (b) Two dilated convolutions with the same dilation factor of 2.

input space that a particular feature is affected by. In addition, weight sharing means the weights of a convolution kernel are unchanged for feature map generation at a certain layer, which can seriously reduce the number of learnable parameters. After each convolution operation, a nonlinear activation function is generally employed. Mathematically, a single convolution operation can be expressed as follows:

$$f^p = a \left(\sum_s w^{(p,q)} * m^{(q)} + b^{(p)} \right) \quad (1)$$

where f^p and m^q are the p th generated feature map of a convolution layer and the q th input feature map, $w^{(p,q)}$ represents the weights in the convolution kernel, s is the size of local receptive field, $b^{(p)}$ represents the bias, the symbol $*$ means the convolution operator, and a denotes the nonlinear activation function. In this study, the rectified linear unit (ReLU) layer was used as the nonlinear activation function as follows:

$$a(h) = \max(0, h) \quad (2)$$

where \max stands for the maximum value.

B. Receptive Field Calculation

It has been proven that a large receptive field can be helpful for SR performance [23], [28]. Generally, three ways are commonly used to enlarge the receptive field: 1) enlarge the convolution kernel; 2) deepen the network; 3) employ dilated convolution. The first two ways can definitely add more learnable parameters, while the third one can enhance the receptive field and maintain the number of learnable parameters. Hence, the dilated convolution was adopted in this study. A dilated filter with dilation factor d can be deemed as a sparse filter of size $(2d+1) \times (2d+1)$. However, a single dilated filter cannot fully use all pixels covered by the receptive field. For example, as shown in Fig. 3(a). Theoretically, the local receptive field is 5×5 by using the dilated filter with a dilation factor of 2, but the number of valid pixels for feature computation is only 9. Moreover, if all layers have the same dilation factor, as shown in Fig. 3(b), the final receptive field is still not complete and resembles a checkerboard, which can also lose several instances of information. Wang *et al.* [40] proposed a hybrid dilated convolution method to deal with this “gridding” problem, which uses a group of dilated convolution layers with different dilation factors. Hence, in practice, a series of dilated convolutions with different dilation factors are commonly used in combination.

Here, the size of the local receptive field and the requirement of a dilated convolution combination are discussed. First, the

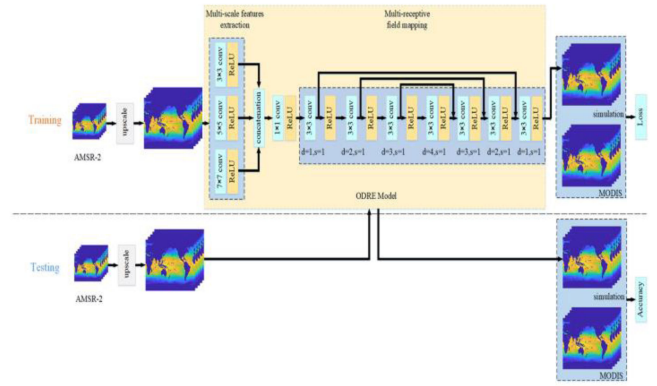


Fig. 4. Proposed ODRE network. d and s represent the dilation factor and stride, respectively.

total receptive field of a dilated convolution combination can be calculated as follows:

$$T_{-s_n} = (l_n - 1) \times d_n + 1 + 2 \times \sum_{i=1}^{n-1} fl \left(\frac{(l_i - 1) \times d_i + 1}{2} \right) \quad (3)$$

where T_{-s_n} is the size of the total receptive field in the n th convolution layer, l_n is the size of convolution kernel in the n th layer, d_n is the corresponding dilation factor, and the fl function returns the value of a number rounded downward to the nearest integer. Generally, the kernel size is often set to an odd number, i.e., $l_n = 3, 5, 7, \dots$, which means $(l_i - 1) \times d_i$ is an even number, so the sum value can be rewritten as $\sum_{i=1}^{n-1} \frac{(l_i - 1) \times d_i}{2}$. Hence, the receptive field can be calculated as

$$\begin{aligned} T_{-s_n} &= 1 + (l_n - 1) \times d_n + \sum_{i=1}^{n-1} (l_i - 1) \times d_i \\ &= 1 + \sum_{i=1}^n (l_i - 1) \times d_i. \end{aligned} \quad (4)$$

A large receptive field size cannot guarantee that more information is included (see Fig. 3). To avoid the “gridding” problem and to effectively use all pixels in the receptive field, Wang *et al.* [40] defined the maximum dilation factor as

$$O_i = \max(O_{i+1} - 2d_i, -O_{i+1} + 2d_i, d_i) \quad O_n = d_n. \quad (5)$$

When $O_2 \leq l_2$, the final receptive field size of a series of dilated convolutions can cover a complete square region. Additionally, the dilation factor within a group should not have a common factor relationship, such as 3, 6, and 9; otherwise, the “gridding” problem will still exist. For a more detailed description about the maximum dilation factor definition, see [40].

C. Architecture of ODRE Network

1) *Multiscale Feature Extraction.*: The first part of the ODRE network is the multiscale feature extraction (see Fig. 4). The distributions of SST in low- and high-resolution maps are similar, so contextual information in different scales may supply the SR procedure with many useful spatial features. Therefore,

three convolution operations of 3×3 , 5×5 , and 7×7 kernel sizes were first conducted simultaneously on the input data, and each convolution operation can generate a single-scale 64-channel feature map. A nonlinear activation ReLU was followed after each convolution. Then, the three feature maps were concatenated to produce a multiscale feature map and to reduce the computational load, a 1×1 convolution layer and an ReLU layer were used to decrease the number of channels to 64. The final multiscale 64-channel feature map containing the contexture information with different scales can be used for posterior processing.

2) *Multireceptive Field Mapping*: The second part of ODRE is a series of dilated convolutions with skip connections. The dilation factors in the dilated convolution group increase from 1 to 4 and symmetrically decrease from 4 to 1. The kernel size of all layers is 3×3 with a stride of 1 pixel. According to (4), the receptive field of each layer is 3, 7, 13, 21, 27, 31, and 33. We found that “the larger the receptive field is, the better the SR performance” is not always valid in SST SR. The influence of the number of dilated convolution layers for SR performance is discussed in Section V.

Skip connections have been successfully used in semantic segmentation [41], SR [23], and missing data removal [29]. With the deepening of the network, a number of image details may be lost, causing blurring of the outputs of the SR model. Hence, putting more details obtained from the previous layers into the posterior layers by using skip connections can enhance the recovery accuracy. On the other hand, the vanishing gradient problem is harmful for SR. It has been proven that skip connections can effectively decrease the influence of gradient vanishing problem. Additionally, details from the previous convolution layers can be better preserved by using the symmetrical structure. Li *et al.* [31] proposed a symmetrical concatenation network for thin cloud removal and the symmetrical structure, which resembles the process of information complementation, was used to better preserve details in cloud-free regions. Therefore, in this study, two corresponding dilated convolution layers were linked by using the skip connections (see Fig. 4).

3) *Network Training*: Because of the similarities between microwave-based and infrared-based SST, instead of directly using the MODIS SST as labels to calculate the loss function, the residual images between upscaled AMSR2 and MODIS SST were employed as labels to train the ODRE network. In the test part, by adding the residual images to the corresponding upscaled AMSR2 SST maps, the simulated high-resolution SST can be generated. The average mean squared error (MAE) was used as the loss function

$$\text{Loss} = \frac{1}{N} \sum_{j=1}^N \|g(x_j) - (y_j - x_j)\|^2 \quad (6)$$

where x and y present the upscaled AMSR2 and MODIS SST, respectively, g indicates the SR model, N is the number of patches in a mini-batch for stochastic gradient descent, and in this study, N is set to 64. The mini-batch gradient descent algorithm with back propagation was used to optimize the loss function. Except for the ODRE model, the other four SR models were also trained

based on the same datasets in this study to validate SR methods for SST reconstruction and to make a comparison with the proposed ODRE model. The codes of SRCNN, VDSR, and DRRN models can be downloaded from GitHub³ and the code of the FDSR model was achieved according to the published manuscript. The learning rate and the momentum were fixed to $1e-6$ and 0.9 , respectively. The models were trained for up to 40 epochs and all the convolutional filters were initialized by using “MSRA” [39]. The proposed network and the other four networks employing the Caffe framework were trained using an NVIDIA GeForce GTX 1050Ti GPU.

D. Quantitative Evaluation

To validate the SR models for SST reconstruction and to make a comparison between the ODRE model and the other four remarkable SR models, three indices including MAE, root-mean-squared error (RMSE), and signal-to-noise ratio (SNR), were used. The actual MODIS SST values and the corresponding SST values obtained from SR methods at the existing points were evaluated by using these three statistical parameters. Similar to the previous studies [43], [44], the SNR used in this study is defined as the ratio of the standard deviation of the simulated values and the standard deviation of the difference between actual values and simulated values for the existing data points. For clarification, the smaller MAE and RMSE and the higher SNR values indicate better SR performance. In addition, the peak SNR (PSNR) and the structural similarity index (SSIM) that have been commonly used in SR evaluation were also used to analyze the performance of ODRE. The MAE, RMSE, SNR, PSNR, and SSIM can be calculated as follows:

$$\begin{aligned} \text{MAE} &= \frac{\sum_{\{ep\}} |k(ep) - r(ep)|}{\{ep\}} \\ \text{RMSE} &= \sqrt{\frac{\sum_{\{ep\}} [k(ep) - r(ep)]^2}{\{ep\}}} \\ \text{SNR} &= \frac{\text{std}[k(ep)]}{\text{std}[|k(ep) - r(ep)|]} \\ \text{PSNR} &= 10 \times \log_{10} \frac{\max(\max[k(ep)], \max[r(ep)])^2 \times \{ep\}}{\sum [k(ep) - r(ep)]^2} \\ \text{SSIM} &= \frac{(2 \times \overline{k(ep)} \times \overline{r(ep)} + C_1) \times (2 \times \text{cov}(k(ep), r(ep)) + C_2)}{(\overline{k(ep)}^2 + \overline{r(ep)}^2 + C_1) \times (\text{std}(k(ep)), \text{std}(r(ep)) + C_2)} \end{aligned} \quad (7)$$

where k and r present the simulated SST obtained from SR models and the actual MODIS SST, ep stands for the existing pixels, std , cov , and max indicate the standard deviation value, the covariance value and the maximum value, $\overline{k(ep)}$ is the mean value of the simulated existing SST pixels, C_1 , C_2 are two small constants to avoid outliers, and they are both equal to 0.001 in this study, and $\{ \}$ is the number of elements in a set.

³[Online]. Available: <https://github.com/>

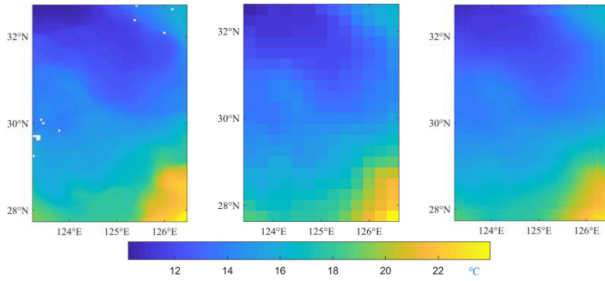


Fig. 5. (a) Illustrations of MODIS SST, (b) original AMSR2 SST, and (c) output from the ODRE model in ECS on April 2, 2019.

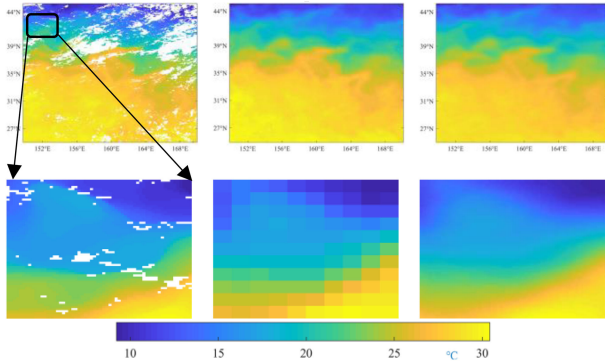


Fig. 6. (a) Illustrations of MODIS SST, (b) original AMSR2 SST, and (c) output from the ODRE model in NWP on September 27, 2019. A subarea marked in black in MODIS SST is used to present the details.

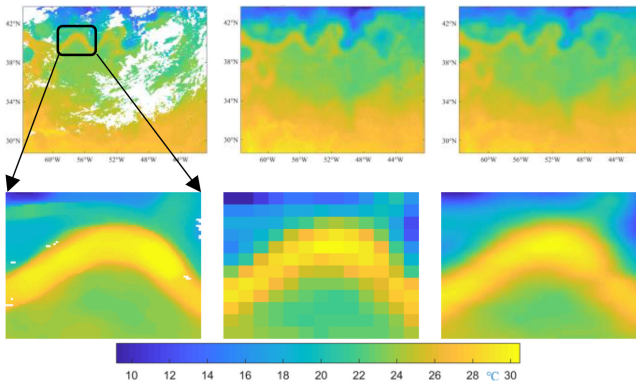


Fig. 7. (a) Illustrations of MODIS SST, (b) original AMSR2 SST, (c) and the output from the ODRE model in WA on July 6, 2019. A subarea marked in black in MODIS SST is used to present the details.

IV. RESULTS

A. Validation of ODRE

First, the validity of the ODRE model on SST reconstruction is discussed. Figs. 5–8 show the illustrations of the original MODIS SST, AMSR2 SST and the corresponding SR SST obtained in ECS, NWP, WA, and SEP. Three subareas marked in the original MODIS SST were also selected to present the details. As a whole, the SST distributions of MODIS and AMSR2 data are similar, but due to the lower resolution, the details cannot be perfectly reflected in the AMSR2 SST, especially in the selected subareas. In addition, the AMSR2 SST data in these four study areas are relatively complete, while the MODIS SST data suffer

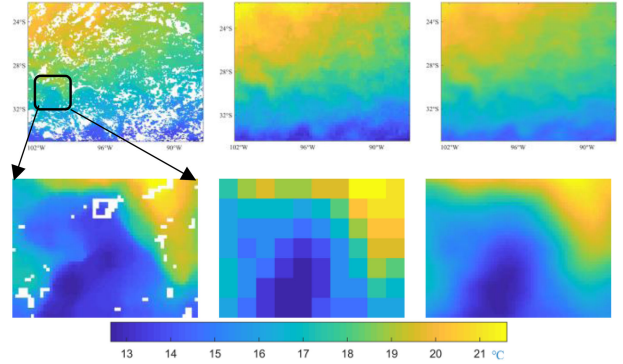


Fig. 8. (a) Illustrations of MODIS SST, (b) original AMSR2 SST, and (c) output from the ODRE model in SEP on September 6, 2019. A subarea marked in black in MODIS SST is used to present the details.

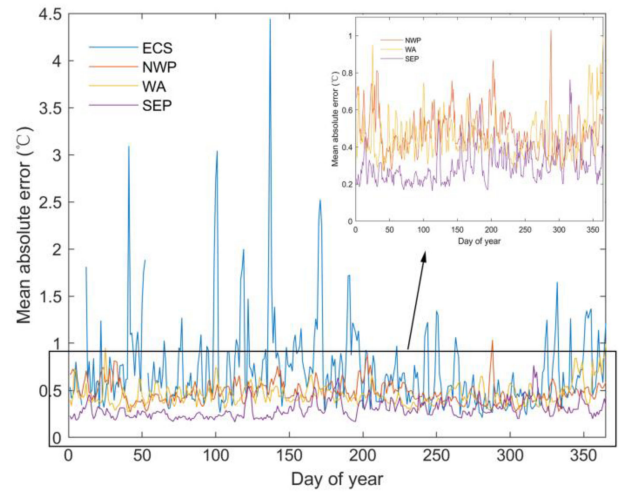


Fig. 9. Distributions of MAEs in the four selected study areas in 2019.

from the serious missing pixels that are randomly distributed in these four study areas.

The outputs from the ODRE model can obtain smooth SST structures that are similar to the actual MODIS SST, and some spatial details can also be found in the reconstructed results. All of these visual results can qualitatively prove that the ODRE model can be effectively used in SST SR. In addition, compared to the output in ECS, the spatial details are better reconstructed in NWP, WA, and SEP, which means the reconstruction accuracy in ECS is relatively low. The reason is likely that the zero padding was used for each layer to keep the same size of the feature maps in each convolution layer, which may cause some errors in boundary regions. The ECS is the smallest region in the four selected study areas, so the border effects may have a more obvious influence on the SR performance.

B. Distributions of MAE and PSNR

The distributions of MAEs and PSNRs obtained from the ODRE model for the four selected study areas were analyzed. As shown in Fig. 9, the MAEs in ECS are generally larger than those in the other three study areas, which corresponds to the conclusion in Section V-A. The large MAEs in ECS may be attributed to the greater amount of missing pixels in the marginal

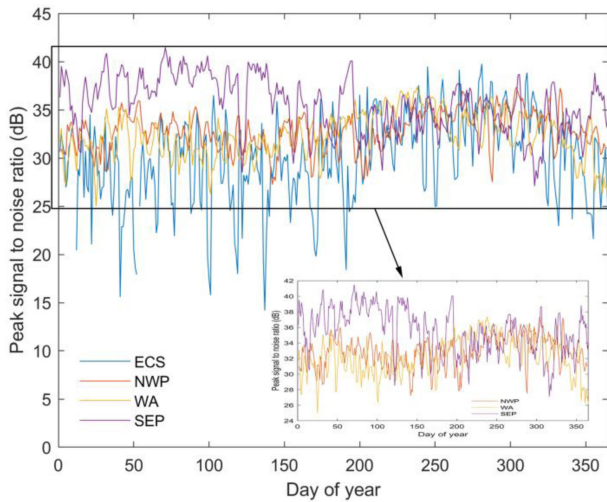


Fig. 10. Distributions of PSNRs in the four selected study areas in 2019.

sea area and the border effects. On the other hand, except for ECS, the MAEs in the other three areas are much smaller. Relatively, the MAEs in SEP are the smallest, which means the ODRE model can achieve better performance in areas covered by stable SST. Furthermore, the distributions of MAEs in NWP and WA are similar. Therefore, in addition to the influence of the missing data, the SST variety is another factor that affects the SR performance. Generally, the more obvious the SSR variety is, the larger the MAEs that can be obtained.

Similar to the conclusions drawn from the MAE distributions, the PSNRs in ECS are the lowest in the four selected study areas. However, in summer, compared with other seasons, the PSNRs are relatively large in ECS. In winter, the PSNRs in SEP have the largest values, while in summer and autumn, especially from July to early November, the distributions of PSNRs in NWP, WA, and SEP are comparable. According to the PSNR formula [see (7)], the maximum values of reconstructed SST and MODIS SST are essential for the final PSNR value. In summer and autumn, the maximum SST in NWP, WA, and SEP are similar, while in winter, the SST in SEP are more stable and higher than that in NWP and WA. In addition, similar to the MAE calculation, the average difference between the simulated SST and MODIS SST at the existing pixels (denominator in the PSNR formula) is relatively smaller in SEP. Therefore, the PSNRs in SEP are relatively larger than those in NWP and WA in winter, while with the enhancement of SST, the PSNRs become similar in these three regions.

In addition, we can see that in ECS, there are top three peak MAEs and bottom three peak PSNRs on February 10, April 11, and May 16, 2019. We took April 11, 2019 as an example to explain these peaks. Because the SR models are trained based on multilevel structure information, the difference in structure between input scene and objective scene can significantly affect the SR performance. Instead of downscaling the high-resolution image to simulate low-resolution image, which is commonly applied in natural image SR, the high- and low-resolution SSTs are obtained from different sensors in this study, so the

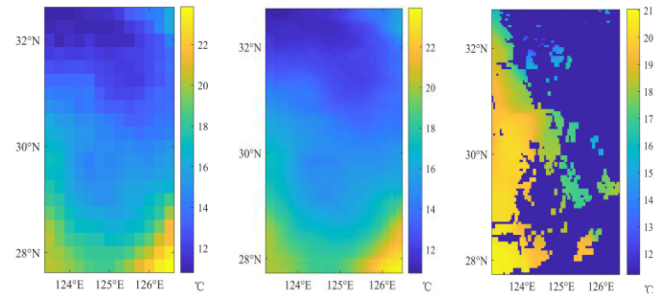


Fig. 11. Illustrations of original AMSR2 SST, (b) output from the ODRE model, and (c) MODIS SST in ECS on April 11, 2019.

difference in spatial structure of high- and low-resolution SSTs may introduce errors. As shown in Fig. 11, we can see that due to large difference in spatial structure between the original AMSR2 SST and the corresponding MODIS SST, the MAE of the reconstructive SST becomes relatively large.

C. Comparisons With Other Super-Resolution Methods

In this section, the performance of the proposed ODRE model, bicubic interpolation and the other four remarkable SR models in SST reconstruction is discussed. As shown in Table I, compared to the statistical results obtained using bicubic interpolation, most SR models can obtain lower MAE and RMSE values and higher SNR, PSNR, and SSIM values in the four selected study areas. Because the SR models used in this study all need a bicubic interpolation preprocess before training and the trained models are implemented based on the bicubic interpolated AMSR2 SST data, we can deduce that the SR methods can enhance the SST reconstruction performance.

The proposed ODRE model can achieve the lowest MAEs in ECS, NWP, and WA, the second-lowest MAE in SEP; the lowest RMSEs in ECS and WA, the second-lowest RMSEs in NWP and SEP. Additionally, ODRE model can acquire the highest or the second-highest PSNRs and SSIMs in ECS, NWP, and SEP. As a whole, the ODRE, FDSR, and DRRN models can achieve relatively better statistical values than the other two SR models, while the SRCNN model acquires the worst results, likely because of its relatively shallow network. Hence, “the deeper, the better” is also correct in SST reconstruction.

Furthermore, the most suitable SR model for various regions is different. For example, in ECS, the ODRE model is the best choice for reconstruction, while in SEP, even though the difference between the statistical results obtained from the ODRE model and the FDSR model is not obvious, the best option is the FDSR model. Because the models are trained globally in this study, one single model may not be the best one for all sea areas. In addition, similar to previous research [45], we can deduce that models trained specifically for the objective region can achieve a higher accuracy than models trained globally.

Finally, even though the proposed ODRE model cannot achieve all of the best statistical values in the four selected study areas, it is still a suitable option for SST reconstruction.

As shown in Fig. 12, the density scatterplots of MODIS and simulated SST values in ECS, NWP, WA, and SEP obtained

TABLE I

COMPARISONS BETWEEN ODRE, BICUBIC INTERPOLATION AND THE OTHER FOUR REMARKABLE SR MODELS FOR RECONSTRUCTING SST IN THE FOUR SELECTED STUDY AREAS. RED AND BLUE INDICATE THE BEST AND THE SECOND-BEST PERFORMANCE, RESPECTIVELY

		ODRE	Bicubic	FDSR	DRRN	VDSR	SRCNN
ECS	MAE	0.6262	0.6919	0.6382	0.6395	0.6487	0.6644
	RMSE	0.9286	1.0104	0.9474	0.9718	0.9631	0.9820
	SNR	7.2306	6.9745	7.1366	6.9217	7.0203	6.9248
	PSNR	31.9961	31.3620	31.8630	34.1671	31.6546	31.3862
	SSIM	0.9837	0.9828	0.9832	0.9827	0.9833	0.9827
NWP	MAE	0.4745	0.5134	0.4775	0.4768	0.4747	0.5335
	RMSE	0.6694	0.7089	0.6692	0.6722	0.6724	0.8245
	SNR	12.0481	12.0489	12.1805	12.2870	12.0608	8.9181
	PSNR	34.0100	33.5611	33.9532	34.8473	33.9143	32.0945
	SSIM	0.9932	0.9930	0.9932	0.9933	0.9933	0.9898
WA	MAE	0.4436	0.5173	0.4437	0.4670	0.4680	0.4922
	RMSE	0.6590	0.7208	0.6601	0.6775	0.6786	0.7597
	SNR	8.3592	8.4256	8.3904	8.4860	8.3671	6.9374
	PSNR	34.5563	33.8616	34.5231	35.8273	34.2612	33.2363
	SSIM	0.9872	0.9869	0.9871	0.9872	0.9873	0.9836
SEP	MAE	0.2886	0.3143	0.2885	0.2949	0.3240	0.3087
	RMSE	0.4104	0.4399	0.4091	0.4192	0.4504	0.4379
	SNR	7.4608	7.3157	7.5231	7.5539	7.0374	7.0890
	PSNR	36.3977	35.9616	36.4437	36.5633	35.6781	35.9118
	SSIM	0.9845	0.9854	0.9842	0.9854	0.9846	0.9839

TABLE II

COMPARISONS BETWEEN THE ODRE MODEL WITH AND WITHOUT MULTISCALE FEATURE EXTRACTION IN THE FOUR SELECTED STUDY AREAS. RED INDICATES THE BEST PERFORMANCE

	ECS		NWP	
	With	Without	With	Without
MAE	0.6262	0.6683	0.4745	0.4858
RMSE	0.9286	0.9803	0.6694	0.6865
SNR	7.2306	6.9453	12.0481	11.7813
PSNR	31.9961	31.5632	34.0100	33.7851
SSIM	0.9837	0.9832	0.9932	0.9932
	WA		SEP	
	With	Without	With	Without
MAE	0.4436	0.4739	0.2886	0.3197
RMSE	0.6590	0.6833	0.4104	0.4459
SNR	8.3592	8.3218	7.4608	7.0537
PSNR	34.5563	34.2471	36.3977	35.7782
SSIM	0.9872	0.9873	0.9845	0.9844

TABLE III

COMPARISONS BETWEEN THE ODRE MODEL WITH DIFFERENT NUMBERS OF DILATED CONVOLUTIONS IN THE FOUR SELECTED STUDY AREAS. RED INDICATES THE BEST PERFORMANCE

	dilation factor	MAE	RMSE	SNR	PSNR	SSIM
ECS	3	0.6251	0.9132	7.4459	31.8880	0.9844
	4	0.6262	0.9286	7.2306	31.9961	0.9837
	5	0.6517	0.9629	7.0265	31.7005	0.9832
NWP	3	0.4763	0.6715	12.0292	33.7812	0.9932
	4	0.4745	0.6694	12.0481	34.0100	0.9932
	5	0.4808	0.6797	11.8927	33.8610	0.9932
WA	3	0.4522	0.6734	8.1691	34.1518	0.9868
	4	0.4436	0.6590	8.3592	34.5563	0.9872
	5	0.4657	0.6782	8.3196	34.3160	0.9872
SWP	3	0.2899	0.4123	7.4316	36.2433	0.9849
	4	0.2886	0.4104	7.4608	36.3977	0.9845
	5	0.3143	0.4400	7.1074	35.8352	0.9844

from (a) ODRE, (b) bicubic (c) DRRN (d) FDSR (e) SRCNN, and (f) VDSR methods are similar in general, which means the SR model can effectively fulfill the SST reconstruction. In detail, the simulated SST may be slightly higher than the MODIS SST by using the DRRN model in the regions covered by high SST in ECS, NWP, and WA, while the simulated SST is lower by using the SRCNN model in the low SST regions in NWP and WA.

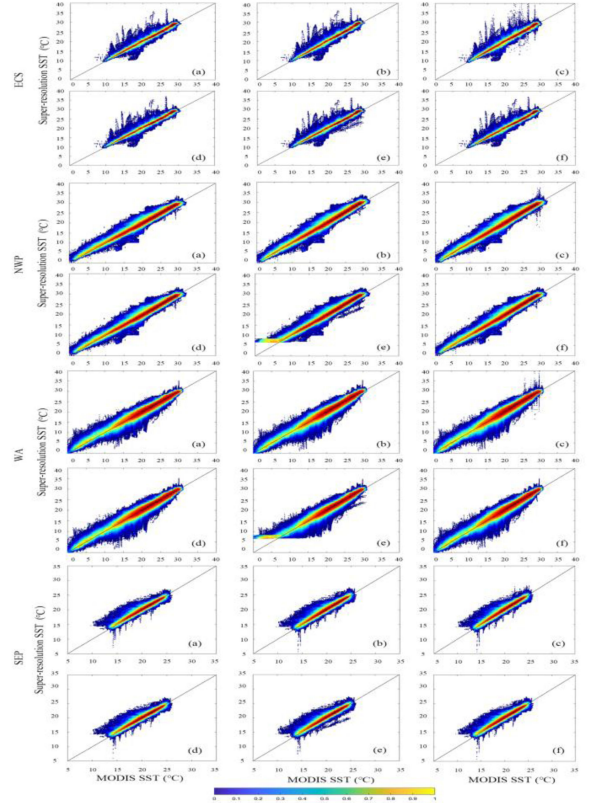


Fig. 12. Density scatterplots of MODIS and simulated SST values in ECS, NWP, WA, and SEP obtained by using (a) ODRE, (b) bicubic, (c) DRRN, (d) FDSR, (e) SRCNN, and (f) VDSR, where the line is the 1:1 line.

Generally, the simulated SST has relatively large errors in ECS, while small errors appear in SEP. Hence, the missing pixels in the original AMSR2 SST data and the variety of SSTs are two key factors in SR SST reconstruction.

V. DISCUSSION

A. Validation of Feature Extraction

The ODRE network has two constituents the so-called multiscale feature extraction and multireceptive field mapping. Here, the necessity of feature extraction was discussed. As shown in Table II, in the four selected study areas, the ODRE network with the multiscale feature extraction part can obtain lower MAEs and RMSEs while obtaining higher SNRs, PSNRs, and SSIMs, which means the multiscale feature extraction part is beneficial to the SST reconstruction. In the four areas, the decreases of MAEs and RMSEs are 0.0421 °C, 0.0113 °C, 0.0303 °C, 0.0311 °C and 0.0517 °C, 0.0171 °C, 0.0243 °C, 0.0355 °C, respectively, and the enhancements of SNRs and PSNRs are 0.2853, 0.2668, 0.0374, 0.4071 and 0.4329, 0.3092, 0.6195. The differences of SSIMs are not obvious in the four study areas. Hence, the incorporation of multiscale feature extraction can be more useful for ECS and SEP regions, likely because more extracted features can add more information from the small region (ECS) and the stable region (SEP) into the SR process.

B. Number of Dilated Convolutions

The number of dilated convolutions in the multireceptive field mapping part can also have an influence on the SR performance. Here, we constructed three networks with different numbers of dilated convolutions, and the dilation factor was used to indicate the number of dilated convolutions. For example, if the dilation factor is 3, the network will have 5 dilation convolution layers with dilation factors of 1, 2, 3, 2, and 1, and the skip connections are also added between two corresponding dilated convolution layers.

As shown in Table III, first, compared to the network with a dilation factor of 3, the network with a dilation factor of 4 can achieve better statistical results, except for ECS, which means a larger receptive field can enhance the SR performance, likely because more spatial information can be used in the SR process. Moreover, the network with a dilation factor of 5 acquires the worst results. A larger receptive field means more zero-paddings are implemented at the boundary regions, which may introduce some errors. Similarly, in ECS, when the dilation factor is set to 3, the SR reconstruction can achieve the best statistical values. On the other hand, more dilation convolutions can carry more information in the SR model, but this approach may also introduce some unnecessary information that is irrelevant to the local spatial structure and may cause an overfitting problem. Hence, simply adding more dilation convolutions may not enhance the SR performance.

VI. CONCLUSION

Deep learning-based SR models have been proposed in recent years and widely used in natural images and remote sensing images. However, their applications in satellite-derived SST are not fully discussed. Additionally, the low- and high-resolution data used in SST SR are obtained from microwave-based and infrared-based SST datasets instead of downscaling the high-resolution images to acquire the corresponding low-resolution images. Hence, the correlation between low- and high-resolution data is relatively low. In this study, an SR model called ODRE that includes multiscale feature extraction and multireceptive field mapping was proposed. The validity of ODRE and the other four remarkable deep learning SR models on SST reconstruction was analyzed. Finally, some details of ODRE were discussed. Overall, the major conclusions can be summarized as follows.

- 1) Compared with bicubic interpolation, the results from SR models have lower MAE and RMSE values and higher SNR, PSNR, and SSIM values in the four selected study areas.
- 2) Compared with four other SR models, the proposed ODRE model can achieve the lowest MAEs in ECS, NWP, and WA, the second-lowest MAE in SEP; the lowest RMSEs in ECS and WA, the second-lowest RMSEs in NWP and SEP. Additionally, ODRE model can acquire the highest or the second-highest PSNRs and SSIMs in ECS, NWP, and SEP.
- 3) The SR performance in the marginal sea areas is relatively poor because of the missing pixels, and in the regions covered by stable SST, the SR can acquire better reconstruction results. In addition, models trained specifically

for the objective region can achieve a higher accuracy than models trained globally.

- 4) A multiscale feature extraction process can enhance the SR performance, especially for the small region (ECS) and the stable SST region (SEP).
- 5) Finally, while a deeper network can be helpful for the SR performance, simply adding more dilation convolutions may not enhance the SR performance.

ACKNOWLEDGMENT

The authors would like to thank the Institute of Geographic Sciences and Natural Resources Research, University of the Chinese Academy of Sciences, the National Marine Data and Information Service, and Tianjin University for their support and also like to thank the anonymous reviewers who provided valuable comments, which greatly enhanced the quality of the manuscript.

REFERENCES

- [1] F. J. Wentz, C. Gentemann, D. Smith, and D. Chelton, "Satellite measurements of sea surface temperature through clouds," *Science*, vol. 288, pp. 847–850, 2000.
- [2] C. J. Xiao, N. C. Chen, C. L. Hu, K. Wang, J. Y. Gong, and Z. Q. Chen, "Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach," *Remote Sens. Environ.*, vol. 233, 2019, Art. no. 111358.
- [3] H. F. Shen *et al.*, "Missing information reconstruction of remote sensing data: A technical review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 61–85, Sep. 2015.
- [4] A. Alvera-Azcarate, A. Barth, J. M. Beckers, and R. H. Weisberg, "Multivariate reconstruction of missing data in sea surface temperature, chlorophyll, and wind satellite fields," *J. Geophysical Res.-Oceans*, vol. 112, 2007, Art. no. C03008.
- [5] X. H. Li, L. Y. Wang, Q. Cheng, P. H. Wu, W. X. Gan, and L. N. Fang, "Cloud removal in remote sensing images using nonnegative matrix factorization and error correction," *ISPRS J. Photogrammetry Remote Sens.*, vol. 148, pp. 103–113, 2019.
- [6] L. Guan and H. Kawamura, "Merging satellite infrared and microwave SSTs: Methodology and evaluation of the new SST," *J. Oceanogr.*, vol. 60, no. 5, pp. 905–912, 2004.
- [7] Y. Z. Wang, L. Guan, and L. Q. Qu, "Merging sea surface temperature observed by satellite infrared and microwave radiometers using Kalman filter," *Periodical Ocean Univ. China*, vol. 40, no. 12, pp. 126–130, 2010.
- [8] A. H. Li *et al.*, "Blending multi-resolution satellite sea surface temperature (SST) products using Bayesian maximum entropy method," *Remote Sens. Environ.*, vol. 135, pp. 52–63, 2013.
- [9] R. W. Reynolds and T. M. Smith, "Improved global sea surface temperature analyses using optimum interpolation," *J. Climate*, vol. 7, no. 6, pp. 929–948, 1994.
- [10] R. W. Reynolds, N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Q. Wang, "An improved in situ and satellite SST analysis for climate," *J. Climate*, vol. 15, no. 13, pp. 1609–1625, 2002.
- [11] R. W. Reynolds, T. M. Smith, C. Y. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, "Daily high-resolution-blended analyses for sea surface temperature," *J. Climate*, vol. 20, no. 22, pp. 5473–5496, 2007.
- [12] W. F. Sun *et al.*, "A new global gridded sea surface temperature product constructed from infrared and microwave radiometer data using the optimum interpolation method," *Acta Oceanologica Sinica*, vol. 37, no. 9, pp. 41–49, 2018.
- [13] J. Thiebaut, E. Rogers, W. Q. Wang, and B. Katz, "A new high-resolution blended real-time global sea surface temperature analysis," *Bull. Amer. Meteorological Soc.*, vol. 84, no. 5, pp. 645–656, 2003.
- [14] Y. Chao, Z. J. Li, J. D. Farrara, and P. Hung, "Blending sea surface temperatures from multiple satellites and in situ observations for coastal oceans," *J. Atmospheric Ocean. Technol.*, vol. 26, no. 7, pp. 1415–1426, 2009.

- [15] M. Martin *et al.*, "Group for high resolution sea surface temperature (GHRSSST) analysis fields inter-comparisons. Part I: A GHRSSST multi-product ensemble (GMPE)," *Deep Sea Res. II, Topical Stud. Oceanogr.*, vol. 77–80, pp. 21–30, 2012.
- [16] H. Chang, D. Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Washington DC, USA, 2004, pp. 275–282.
- [17] M. Bevilacqua, A. Roumy, C. Guillemot, and M. L. A. Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in *Proc. Brit. Mach. Vis. Conf.*, University of Surrey, Guildford, U.K., 2012, pp. 1–12.
- [18] J. C. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
- [19] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Proc. Int. Conf. Curves Surfaces*, Springer, Avignon, France, 2010, pp. 711–730.
- [20] S. Schuler, C. Leistner, and H. Bischof, "Fast and accurate image up-scaling with super-resolution forests," in *Proc IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, 2015, pp. 3791–3799.
- [21] C. Dong, C. C. Loy, K. M. He, and X. O. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.
- [22] C. Dong, C. C. Loy, and X. O. Tang, "Accelerating the super-resolution convolutional neural network," in *Proc. Eur. Conf. Comput. Vis.*, Springer, Amsterdam, The Netherlands, 2016, pp. 391–407.
- [23] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc IEEE Conf. Comput. Vis. Pattern Recognit.*, Seattle, WA, USA, 2016, pp. 1646–1654.
- [24] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, 2016, pp. 770–778.
- [25] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proc IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, 2016, pp. 1637–1645.
- [26] Y. Tai, J. Yang, and X. M. Liu, "Image super-resolution via deep recursive residual network," in *Proc IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, HI, USA, 2017, pp. 2790–2798.
- [27] Y. P. Chang and B. Luo, "Bidirectional convolutional LSTM neural network for remote sensing image super-resolution," *Remote Sens.*, vol. 11, no. 20, 2019, Art. no. 2333.
- [28] L. Zhang *et al.*, "Fast single image super-resolution via dilated residual networks," *IEEE Access*, vol. 7, pp. 109729–109738, 2019.
- [29] Q. Zhang, Q. Q. Yuan, C. Zeng, X. H. Li, and Y. C. Wei, "Missing data reconstruction in remote sensing image with a unified spatial-temporal-spectral deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, Aug. 2018.
- [30] X. J. Mao, C. H. Shen, and Y. B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Proc. Adv. Neural Inf. Process. Syst.*, Barcelona, Spain, vol. 29, 2016.
- [31] W. B. Li, Y. Li, D. Chen, and J. C. W. Chan, "Thin cloud removal with residual symmetrical concatenation network," *ISPRS J. Photogrammetry Remote Sens.*, vol. 153, pp. 137–150, 2019.
- [32] R. Gaetano, D. Ienco, K. Ose, and R. Cresson, "A two-branch CNN architecture for land cover classification of PAN and MS imagery," *Remote Sens.*, vol. 10, no. 11, 2018, Art. no. 1746.
- [33] L. Zhu, Y. S. Chen, P. Ghamisi, and J. A. Benediktsson, "Generative adversarial networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5046–5063, Sep. 2018.
- [34] Z. F. Shao and J. J. Cai, "Remote sensing image fusion with deep convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1656–1669, May 2018.
- [35] Z. F. Shao, J. J. Cai, P. Fu, L. Q. Hu, and T. Liu, "Deep learning-based fusion of Landsat-8 and Sentinel-2 images for a harmonized surface reflectance product," *Remote Sens. Environ.*, vol. 235, 2019, Art. no. 111425.
- [36] Z. Y. Tan, L. P. Di, M. D. Zhang, L. Y. Guo, and M. L. Gao, "An enhanced deep convolutional model for spatiotemporal image fusion," *Remote Sens.*, vol. 11, no. 24, 2019, Art. no. 2898.
- [37] M. H. Jiang, H. F. Shen, J. Li, Q. Q. Yuan, and L. P. Zhang, "A differential information residual convolution neural network for pansharpening," *ISPRS J. Photogrammetry Remote Sens.*, vol. 163, pp. 257–271, 2020.
- [38] X. Y. Liu, Q. J. Liu, and Y. H. Wang, "Remote sensing image fusion based on two-stream fusion network," *Inf. Fusion*, vol. 55, pp. 1–15, 2020.
- [39] A. P. Yang, B. W. Yang, Z. Ji, Y. W. Pang, and L. Shao, "Lightweight group convolutional network for single image super-resolution," *Inf. Sci.*, vol. 516, pp. 220–233, 2020.
- [40] P. Q. Wang *et al.*, "Understanding convolution for semantic segmentation," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, New York, NY, USA, 2018, pp. 1451–1460.
- [41] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, 2015, pp. 3431–3440.
- [42] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Santiago, Chile, 2015, pp. 1026–1034.
- [43] B. Ping, F. Z. Su, and Y. S. Meng, "Reconstruction of satellite-derived sea surface temperature data based on an improved DINEOF algorithm," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 4181–4188, Aug. 2015.
- [44] B. Ping, F. Z. Su, and Y. S. Meng, "An improved DINEOF algorithm for filling missing values in spatio-temporal sea surface temperature data," *PLOS One*, vol. 11, no. 5, 2016, Art. no. e0155928.
- [45] A. Ducournau and R. Fablet, "Deep learning for ocean remote sensing: An application of convolutional neural networks for super-resolution on satellite-derived SST data," in *Proc. 9th IAPR Workshop Pattern Recognit. Remote Sens.*, Cancun, Mexico, 2016.



Bo Ping received the B.E. and Ph.D. degrees in remote sensing science and technology from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2009 and 2015, respectively.

In 2013, he was with the University of Rhode Island, USA, as a Visiting Student for one year.

His current research interests include information extraction based on satellite data and satellite-derived data recovery.



Fenzhen Su received the B.E. degree in remote sensing science and technology from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 1994, and the Ph.D. degree in remote sensing science and technology from the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, in 2001.

His current research interests include spatio-temporal ocean information detection, oil spilling detection, and Ocean GIS model research.



Xingxing Han received the Ph.D. degree in remote sensing science and technology from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China, in 2017.

In 2015, she was with the University of South Florida, USA as a Visiting Student for two years. Her current research interests include wetland mapping and water quality monitoring based on satellite data.



Yunshan Meng received the B.E. degree in geomatics engineering from Changchun Institute of Technology, Changchun, China, in 2009, the M.Sc. degree in remote sensing science and technology from Wuhan University, Wuhan, China, in 2011, and the Ph.D. degree in mapping and geographic information system from the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, in 2016.

Her current research interests include high resolution remote sensing image segmentation and geological disaster identification and monitoring using remote sensing technology.