

Enhanced resolution of FY4 remote sensing visible spectrum images utilizing super-resolution and transfer learning techniques

Bo Zhang, Muyuan Ma, Mingqing Wang, Danfeng Hong, Le Yu, Jie Wang, Peng Gong and Xiaomeng Huang*

Abstract—Remote sensing images acquired by the FY4 satellite are crucial for regional cloud monitoring and meteorological services. Inspired by the success of deep learning networks in image super-resolution, we applied image super-resolution to FY4 visible spectrum (VIS) images. However, training a robust network directly for FY4 VIS image super-resolution remains challenging due to the limited provision of high resolution FY4 sample data. Here, we propose a super-resolution and transfer learning model, FY4-SR-Net. It is composed of pre-training and fine-tuning models. The pre-training model was developed using a deep residual network and a large number of FY4 A 4km and 1km resolution VIS images as the training data. The knowledge derived from 4 km to 1 km resolution images was incorporated into FY4 B 1 km to 0.25 km resolution VIS images. The FY4-SR-Net is fine-tuned by incorporating limited 1km and 0.25km resolution panchromatic (PAN) images, and then producing 1km super-resolution VIS images of the FY4 satellite. Using the one-day FY4 test dataset for qualitative and quantitative evaluations, the FY4-SR-Net outperformed the classic bicubic interpolation approach with a 16.12% reduction in root mean square error (RMSE) and a 2.97% rise in peak signal-to-noise ratio (PSNR) averages. The structural similarity (SSIM) value average increased by 0.0026. This work provides a new precedent for improving the spatial resolution of FY4 series meteorological satellites, which has important scientific significance and application properties.

Index Terms—Remote sensing images, super-resolution, transfer learning, pre-training, fine-tuning.

I. INTRODUCTION

As part of a new generation of China's geostationary orbit satellites, the FY4 meteorological satellite helps with weather forecast analysis, short-term climate prediction, environmental management, resource development, disaster prevention and mitigation, and scientific research. This satellite

is an essential component of China's comprehensive meteorological observation satellite network [1]. The spatial resolution of a remote sensing image is an important quality measure and a leading indicator of a country's aeronautical capabilities. Remote sensing images with metric and submetric spatial resolution is already widely used, but their temporal resolution is still rather poor. Some remote sensing images have a low spatial resolution but high temporal resolution, such as the FY4 satellite VIS images with a temporal resolution of 5 minutes and a spatial resolution between 0.5 and 1 kilometer. Such a mismatch between spatial and temporal resolution scarcely satisfies the demand for FY4 images with increasing production applications; thus, the advancement of meteorological remote sensing technology is greatly impeded [2].

Significant progress has been made in the area of remote sensing image process with various advanced deep networks [3-6]. Deep learning algorithms offer a cost-effective and efficient solution to this mismatching problem [7]. Super-resolution (SR) is a technique for enhancing image spatial resolution by reconstructing high-resolution (HR) images from single or multiple low-resolution (LR) images. It is classified into two categories: classical interpolation methods [8-10] and deep-learning-based methods [11-13], with the latter subdivided further into single-frame SR methods [14, 15] and multi-frame SR methods [16, 17]. MFSR constructs HR images by acquiring multiple LR images of the identical scene with the same or distinct sensors [18]. Merino and Núñez [19] presented a technique called Super-Resolution Variable-Pixel Linear Reconstruction (SRVPLR) for reconstructing a high-resolution (HR) image from many low-resolution (LR) images recorded over an extended period of time. This technique was adapted from drizzle [20], which was meant to operate with dithered, under sampled astronomical images. Before image reconstruction, Shen *et al.* [21] developed an SR approach for

This work is supported by the National Key Research and Development Program of China (2020YFA0607900, 2020YFA0608003), the National Natural Science Foundation of China (42125503, 42075137), and in part by the Key Research and Development Program of Shaanxi Province, China (2020SF-434). (Corresponding author: Xiaomeng Huang.)

B. Zhang, M. Ma, M. Wang, L. Yu and X. Huang are with the Ministry of Education Key Laboratory for Earth System Modeling, and Department of Earth System Science, Tsinghua University, Beijing 10084, China (e-mail: zhangbo2020@tsinghua.edu.cn; mamuyuan@tju.edu.cn; wangmq18@tsinghua.edu.cn; leyu@tsinghua.edu.cn; hxm@tsinghua.edu.cn).

D. Hong is with the key Laboratory of Computational Optical Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: hongdf@aircas.ac.cn).

J. Wang is Peng Cheng Laboratory, Shenzhen, Guangdong 518000, China (e-mail: wangj10@pcl.ac.cn).

P. Gong is with the department of Geography and Department of Earth Sciences, University of Hong Kong, Hong Kong, China (e-mail: penggong@hku.hk).

MODIS remote sensing images where image registration was conducted in the range and spatial domains. Using panchromatic Landsat7 images captured on several days, Li *et al.* [22] proposed a MAP-based SR approach with a general hidden Markov tree model and tested it. Fan *et al.* [23] developed a POCS-based SR algorithm with a slant knife-edge technique. The approach was evaluated using Airborne Digital Sensor 40 (ADS40) three-line images and overlapping Gaofen-2 images. Because of the simplicity and excellent performance acquired through intense supervised training, the single-frame SR has become the workhorse.

Initial studies on single-frame SR for remote sensing images used simple structures like 3D Full Convolutional Neural Network (3D-FCNN) and Remote Sensing Image Convolutional Neural Network (RSCNN) with direct superposition of convolutional layers [24-26] as their basis. An unsupervised depth generating network was developed by Haut *et al.* [27] in order to overcome the lack of benchmark training data for super-resolution remote sensing images. With the advent of super-resolution networks in the realm of computer vision, modular designs, such as the Multi-Perception Attention Network (MPSR) [28] and Dense-Sampling Super-Resolution Network (DSSR) [29] were implemented for single-frame SR. For better segmentation accuracy after super-resolution, Lei *et al.* [30] proposed the S2Net network, which can simultaneously accomplish SR and image segmentation for remote sensing images. Coupling remote sensing images and the generated high-resolution images for a discriminative network, Lei *et al.* [31] developed a coupled discriminative generative adversarial network (CDGAN), which was able to perform well on an ultra-resolution task in the low-frequency region.

In regard to implementing deep learning-based super resolution networks, the quality and quantity of available high- and low-resolution samples, as well as training images, are all important considerations. FY4 remote sensing VIS images' inherent low-resolution qualities make obtaining sufficient high-resolution training data problematic. To address issue, we proposed a novel methodology with two steps. The first is to design a deep residual model for pre-training chunked 4-fold super-resolution feature knowledge. The second measure is to fine-tune the system with higher-resolution remote sensing image pairs. Based on this approach, the resolution of FY4 visible spectrum (VIS) can be enhanced. Several image super-resolution experiments use image data instead of raw grayscale values as model inputs to build training networks, resulting in a loss of pixel information and a reduction in model generalizability [32]. In addition, to our knowledge, there have been few super-resolution studies of the FY4 global-scale meteorological satellite.

The FY4-SR-Net model, which we created in accordance with the pre-training and fine-tuning of image transfer learning theory, effectively addresses these issues. By using FY4 4 km-1 km monitoring VIS image data (gray values) as pre-training data, we can avoid the scarcity of high-resolution remote sensing training datasets. Training with higher-resolution remote sensing image pairs is then used to fine-tune the system. Our proposed method for improving resolution is more accurate and works better than classical interpolation, both in terms of quality and clarity. A higher resolution FY4 satellite VIS image can visualize the finer structure and type of cloud masses.

The rest of this paper is organized as follows: first, some details about the data and our network structure (Data and Methods). Then, we demonstrate the effectiveness of the proposed method using experimental results (Results). Finally, we conclude this paper with some comments on future work (Conclusions).

II. DATA AND METHODS

A. Data Collection and Preprocessing

1) *FY4 A*: FY4 A belongs to the second generation of geostationary meteorological satellites in China. It is made to meet the country's wide range of environmental and space science needs [33], such as those in the oceans, agriculture, forestry, and hydropower. The FY4 A satellite was formally launched on November 11, 2016, and is equipped with an Advanced Geostationary Radiation Imager (AGRI), a Geostationary Interferometric Infrared Sounder (GIIRS), a Lightning Mapping Imager (LMI), and a Space Environment Package (SEP) instrument. A scanning imaging radiometer (SIR) is responsible for collecting cloud data. Fig. 1 shows the FY4-A satellite band and resolution information. With its fourteen channels, the satellite can detect aerosols and snow, as well as the different stages of clouds and high and mid-level water vapor. When compared to the single visible channel of FY2, FY4 A for the first time produces color satellite cloud maps and generates regional observation photos in one minute.

Users can access FY4 A data from the official website of the National Meteorological Satellite Center [34]. The FY4 star's main payload, the Advanced Geostationary Radiation Imager (AGRI), has a complex double-scanning mirror mechanism that enables precise and flexible two-dimensional pointing, allowing for quick scanning of areas at a minute rate. It includes a) Full disk data with 4km resolution; b) Full disk data with 1km resolution; c) China zone data with 4km resolution; d) China zone data with 1km resolution. Full disk identified as DISK, China zone marker as REGC [35]. China zone REGC with a 5-minute temporal resolution was chosen as the training data for this work.







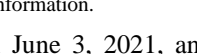
		Spectral band		Spatial and temporal resolution		
Channel		Wavelength(μm)		Visible band	Near Infrared	
1	Visible band	0.45-0.49	Visible band 	Visible band 0.5-1km	Near Infrared 1-4km	
2		0.55-0.75				
3	Near Infrared	0.75-0.90	Near Infrared 	DISK observation time		
4		1.36-1.39				
5		1.58-1.64				
6		2.1-2.25				
7	Infrared	3.5-4.0 (high)	Infrared 	REGC observation time		
8		3.5-4.0 (low)				
9	5.8-6.7	4 Channels	5 mins		5 mins	5 mins
10	6.9-7.3		REGC observation time			
11	Infrared	8.0-9.0	Infrared 	REGC observation time		
12		10.3-11.3				
13		11.5-12.5				
14		13.2-13.8			1h	1h
			14 Channels			

Fig. 1. FY4 A satellite band and resolution information.

2) *FY4 B*: FY4 B was launched on June 3, 2021, and is primarily used for operational meteorological satellite monitoring. The average life expectancy of a B star is seven years longer than that of A star. Its successful launch is crucial

for ensuring the upgradation of China's geostationary meteorological satellites and their continued reliability and stability. The addition of the Geostationary Orbit High-speed Imager (GHI) to FY4 B provides the capability for high-speed, high-resolution regional imaging. The Geo Interferometric Infrared Sounder (GIIRS) increases the spatial resolution in the visible spectrum to 1 km.

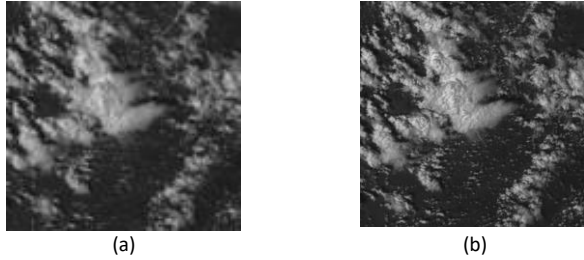


Fig. 2. FY4 B (a)1 km and (b) 0.25 km data

FY4 B has the capability of rapid imaging, i.e., 1-minute interval with a spatial resolution up to 250 m in a region. It includes: : a) China zone data with 2 km resolution; b) China zone data with 500 m resolution; c) China zone data with 250 m resolution. The 2 km resolution data were not utilized in this study. Datasets with a resolution of one kilometer were created by downsampling data with a resolution of 500 meters using an interpolation technique typical for image super-resolution. Fig. 2 depicts a tiny portion of the FY4 B imaging data. With the inclusion of multichannel AGRI the FY4 B spatial resolution in the 2.1 μm and 3.5 μm bands has been enhanced to 2 km. The FY4 B and A twin-star networks provide China and other countries along the "Belt and Road" with weather monitoring and forecasting, emergency disaster prevention and mitigation services.

3) *Data Preprocessing*: The FY4 data were projected and converted into a unified coordinate system in order to eliminate position bias. To minimize sensor errors, the relationship between digital quantization values and radiation brightness values was established after radiometric calibration [36]. The blue band (from 0.45 μm to 0.49 μm) of FY4 helps to obtain clear cloud boundary information when drawing cloud cover maps [37]. So we used the NOMChannel01 0.47 μm VIS channel of FY4 A data, collected from Sept. 13 to 20, 2021 for model pre-training. Deep learning has been shown to be effective in improving the resolution of low-resolution multi-spectral images with high spatial resolution PAN images [38]. Transfer learning and model testing used the NOMChannel01 PAN channel of FY4 B data from Nov.25 to 30, 2021. Gray values of the FY4 data were normalized from 0 to 1 before being entered into the model in order to remove the influence of local perception characteristics. High-resolution (HR) FY4 1000 m images are split into 128x128 data points while low-resolution (LR) FY4 4000 m images are split into 32x32 data points during the preprocessing phase. The HR FY4 250 m images of the B star and the LR FY4 1000 m images performed the same processing to generate the fine-tuning dataset simultaneously. In addition to fragmentation, FY4 data with missing values were removed from the dataset. The final dataset consists of 316,199 groups, on which a random 80/20

training/testing split was performed. Python GDAL was used for preprocessing.

B. Methods

1) *The structure of the pre-training network*: Due to a lack of high-resolution monitoring data, direct training of a robust network for FY4-SR remains a difficulty. A pre-training FY4 model is built using a deep residual model and a large number of FY4 4 km resolution and 1km resolution images as input train data, taking inspiration from transfer learning. As shown in Fig. 3, the FY4 model knowledge gained from 4 km resolution and 1 km resolution data is incorporated into 1km resolution and 0.25 km resolution data, and the network is fine-tuned by taking restricted 1km resolution and 0.25 km resolution data to build an FY4-SR-Net that meets the FY4 satellite's 1km super-resolution. The final input is the low resolution FY4 1 km FY_L and the high resolution FY4 250 m FY_H is obtained.

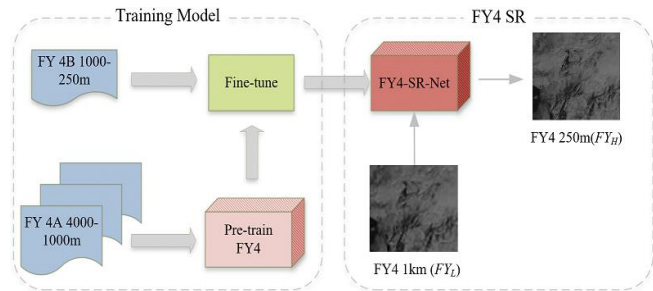


Fig. 3. The working process of FY4-SR-Net

As illustrated in Figure 4, the pre-trained FY4 network consists of low-level feature extraction, element-wise summation, and upsampling layers. A collection of features are extracted by the first convolutional layer of the pre-trained FY4 network. The model is trained with a residual network, where the input in each residual block is supposed to be x , the expected output is $f(x)$, and the residual map $g(x) = f(x) - x$. The element-wise summation layer then accumulates the original inputs x and $g(x)$ to produce the desired output (x). The residual network structure can provide more precision by implementing deeper network layers [39].

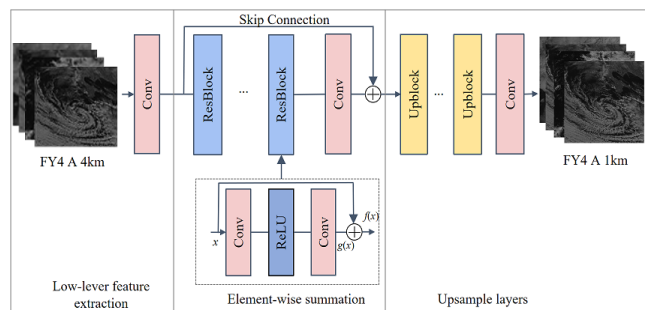


Fig. 4. The pre-trained FY4 network

The sub-pixel convolution method proposed by [40] is used for upsampling, as depicted in Fig. 5. First, the features of r^2 channels are obtained by convolution, and then the high-resolution image is obtained by cycle screening.

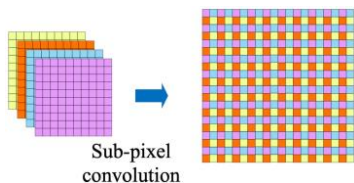


Fig. 5. The sub-pixel convolution structure

In comparison to the typical deconvolution layer, this sub-pixel convolution can decompose low-resolution data into high-resolution space without additional computation. Since ReLU has good nonlinear fitting performance [41], it is used as the model's activation function.

2) *Fine-tuning and loss function*: Transfer learning is the process of adapting the knowledge or patterns acquired in one area or task to a different but similar domain or work. As shown in Fig.6, the knowledge gained by solving the super-resolution of FY 4A 4 km resolution and 1km resolution data is used to construct a higher resolution FY4-SR-Net model. Some layers of the pre-training network were frozen, and then the finite high-resolution FY4 B 1 km to 0.25 km resolution data pairs were used to fine-tune the network.

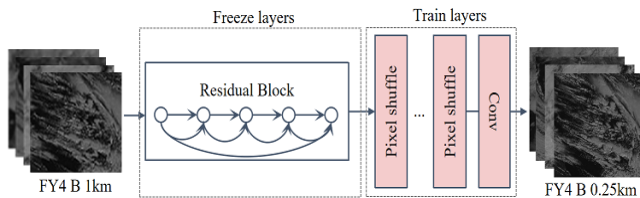


Fig. 6. Fine-tuning the network with finite high-resolution FY4 B dataset

In the field of single-image SR reconstruction, the pixel loss function and the perceptual loss function are the most commonly employed loss functions. This study uses the former function to minimize the pixel error between the output and input target for training the model. The FY4-SR-Net optimizes the network with the MSE (mean square error). In addition, we analyzed additional metrics, e.g., MAE (mean absolute error), SSIM and PSNR. The MSE, which is insensitive to outliers, safeguards the model's stability. This facilitates the reconstruction of FY4 data with a high resolution.

3) *Evaluating metrics*: Quantitative evaluation and visual effects were used to assess the quality of super resolution image data. The RMSE is a widely used numerical accuracy metric for evaluating data accuracy. It does not indicate the number of individual point errors but describes the overall dispersion of the FY4 data. where H_i is the value of the original image, h_i is the value of the super-resolution image, and n is the number of sampling points.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (H_i - h_i)^2} \quad (1)$$

The PSNR measures the pixel difference between reconstructed high-resolution images and actual images, and is the most popular metric for assessing image quality. MSE is calculated before PSNR, and it is frequently used to construct loss functions. Given two $m \times n$ monochromatic images I and K , one of which is the noise approximation of the other, MSE and PSNR are calculated as:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (2)$$

$$PSNR = 10 \cdot \log_{10} \left(\frac{\Delta S^2}{MSE} \right) \quad (3)$$

where ΔS represents the maximum color value of the image points. The ΔS value equals 255 if each sampling point is represented by 8 bits. However, the grayscale values of the FY4 data were frequently far above 255. The original ΔS formula is modified to reflect the difference between the maximum and minimum gray values in the FY4 data. The higher the PSNR value is, the better the image quality.

The SSIM (Eq. 4) is a measurement of the similarity of two images and has a wide range of applications in both image deblurring and image super-resolution.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y+c_1)(\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \quad (4)$$

where x and y are real images and super-resolution images, respectively, μ_x and μ_y represent the standard deviations of x and y respectively. σ_x and σ_y are the standard deviations and σ_{xy} is the covariance. SSIM is a number ranging from 0 to 1. The larger the value, the smaller the difference between the output image and the distortion-free image, and hence the better the image quality. When the two images are identical, SSIM equals 1.

III. RESULTS

The network training and experiments were implemented with PyTorch [42]. The model consists of 89 layers of convolutional networks, with the convolutional kernel size in each layer set to 3 and the padding set to 1. The input data were passed through the convolutional layers, and the output is resized with the zero-padding method to achieve a constant output matrix size for the subsequent layer. We utilized the Adam optimizer with a learning rate of 0.0001 to train the model with the large dataset. The exponential decay formula is employed to adjust the learning rate. The size of the training batch is set at 128. The models under went 200 epochs of training, with each epoch consisting of 2400 training steps. In order to prevent overfitting, the early stopping approach is used to terminate training when the model's performance begins to drop in the validation set (patience is 7). For model training and evaluation, a workstation with four 2080ti 11G GPUs was employed.

A. Pre-training and fine-tuning performance

For our proposed FY4-SR-Net model, the pre-training network was trained with 1318 remote sensing image samples from Sept.13 to 20, 2021 of FY4 A. In other words, pre-training to acquire chunked 4-fold super-resolution feature information. FY4 B data from Nov. 25 to 29, 2021 were utilized to fine-tune the pre-trained network. The remaining FY4 B monitoring data with an interval of 2 hours on Nov. 30, 2021, were selected for model testing, as shown in Fig. 7. FY4 B is a geostationary satellite that orbits the equator, and UTC 12:00-20:00 corresponds to the the nighttime of Eastern Hemisphere. It is therefore impossible for active remote sensing satellites to gather reflectance spectra, rendering the image maps nearly dark.

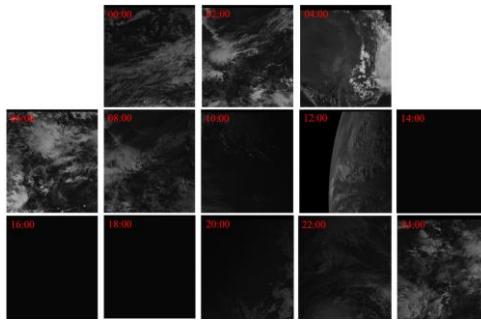


Fig.7. Monitoring data with an interval of 2 hours on Nov. 30, 2021.

TABLE I

THE RMSE INTERNAL TEST RESULTS WITH TRADITIONAL BICUBIC METHOD ON NOV. 30, 2021

Times	R ₁ (Bicubic)	R ₂ (FY4 A Pre-training)	R ₃ (FY4 B Pre-training)	R ₄ FY4-SR-Net	ΔR_2 =(R ₁ -R ₂)/R ₁ %	ΔR_3 =(R ₁ -R ₃)/R ₁ %	ΔR_4 =(R ₁ -R ₄)/R ₁ %
00:00	24.25	20.50	20.50	20.07	15.46	15.46	17.24
02:00	28.46	23.22	23.55	22.80	18.41	17.25	19.89
04:00	37.47	30.58	30.37	29.50	18.39	18.95	21.27
06:00	21.12	17.75	17.56	17.28	15.96	16.86	18.18
08:00	19.63	16.69	16.72	16.27	14.98	14.82	17.12
10:00	10.03	8.96	9.21	8.93	10.67	8.18	10.97
*12:00	1.43	1.70	1.34	1.32	-18.88	6.29	7.69
*14:00	1.41	1.69	1.28	1.29	-19.86	9.22	8.51
*16:00	1.34	1.60	1.10	1.04	-19.40	17.91	22.39
*18:00	1.28	1.57	1.07	1.06	-22.66	16.41	17.19
*20:00	1.26	1.56	1.06	1.05	-23.81	15.87	16.67
22:00	12.95	11.04	11.08	10.75	14.75	14.44	16.99
24:00	14.22	12.07	12.27	12.03	15.12	13.71	15.40
all Avg					1.47	14.26	16.12

* marked :the near-dark period.

TABLE II

THE PSNR INTERNAL TEST RESULTS WITH TRADITIONAL BICUBIC METHOD ON NOV. 30, 2021

Times	P ₁ (Bicubic)	P ₂ (FY4 A Pre-training)	P ₃ (FY4 B Pre-training)	P ₄ FY4-SR-Net	ΔP_2 =(P ₂ -P ₁)/P ₁ %	ΔP_3 =(P ₃ -P ₁)/P ₁ %	ΔP_4 =(P ₄ -P ₁)/P ₁ %
00:00	44.55	46.00	46.00	46.19	3.25	3.25	3.68
02:00	43.15	44.92	44.80	45.08	4.10	3.82	4.47
04:00	40.77	42.53	42.59	42.84	4.32	4.46	5.08
06:00	45.74	47.26	47.35	47.48	3.32	3.52	3.80
08:00	46.38	47.79	47.77	48.01	3.04	3.00	3.51
10:00	52.21	53.19	52.95	53.22	1.88	1.42	1.93
*12:00	69.12	67.59	69.78	69.83	-2.21	0.95	1.03
*14:00	69.21	67.68	70.07	70.01	-2.21	1.24	1.16
*16:00	69.67	68.13	71.39	71.90	-2.21	2.47	3.20
*18:00	70.06	68.29	71.64	71.66	-2.53	2.26	2.28
*20:00	70.17	68.36	71.71	71.76	-2.58	2.19	2.27
22:00	49.99	51.38	51.34	51.61	2.78	2.70	3.24
24:00	49.18	50.60	50.46	50.63	2.89	2.60	2.95
all Avg					1.06	2.61	2.97

* marked :the near-dark period.

TABLE III

THE SSIM INTERNAL TEST RESULTS WITH TRADITIONAL BICUBIC METHOD ON NOV. 30, 2021

Times	S ₁ (Bicubic)	S ₂ (FY4 A Pre-training)	S ₃ (FY4 B Pre-training)	S ₄ FY4-SR-Net	ΔS_2 =S ₂ -S ₁	ΔS_3 =S ₃ -S ₁	ΔS_4 =S ₄ -S ₁
00:00	0.9800	0.9856	0.9856	0.9862	0.0056	0.0056	0.0062
02:00	0.9805	0.9866	0.9863	0.9872	0.0061	0.0058	0.0067
04:00	0.9790	0.9856	0.9856	0.9865	0.0066	0.0066	0.0075
06:00	0.9887	0.9922	0.9921	0.9924	0.0035	0.0034	0.0037
08:00	0.9862	0.9900	0.9899	0.9904	0.0038	0.0037	0.0042
10:00	0.9953	0.9963	0.9961	0.9964	0.0011	0.0011	0.0011
*12:00	0.9998	0.9998	0.9998	0.9999	0	0	0.0001
*14:00	0.9998	0.9998	0.9998	0.9999	0	0	0.0001
*16:00	0.9998	0.9998	0.9998	0.9999	0	0	0.0001
*18:00	0.9998	0.9998	0.9999	0.9999	0	0.0001	0.0001
*20:00	0.9998	0.9998	0.9999	0.9999	0	0.0001	0.0001
22:00	0.9926	0.9946	0.9945	0.9949	0.0023	0.0023	0.0023
24:00	0.9932	0.9951	0.9951	0.9952	0.0020	0.0020	0.0020
all Avg					0.0023	0.0023	0.0026

* marked :the near-dark period.

In addition, we trained two deep residual network based pre-training networks with FY4 A and FY4 B data respectively, referred to as FY4 A pre-training and FY4 B pre-training. After training these networks under identical settings, the testing dataset D_{test} was used to compare the performance of FY4 SR with a fourfold upscaling factor. We compared the performance of these training strategies with the bicubic interpolation method as a baseline [43]. Note that interpolation was performed on low-resolution images of each sample to create high-resolution images. We quantitatively evaluate the three training strategy methods by evaluating metrics, including the RMSE, PSNR and SSIM.

As illustrated in Tables I-III, FY4-SR-Net outperformed bicubic and the other two training strategies for most of the test samples in terms of the RMSE, PSNR and SSIM. The FY4-SR-Net RMSE was lowered by an average of 16.12%, while the FY4 A pre-training and FY4 B pre-training were reduced by average of 1.47% and 14.26%, respectively. The PSNR averages increased by 2.97%, and the FY4 A pre-training and FY4 B pre-training increased by an average of 1.06% and 2.61%, respectively. The SSIM value average increased by 0.0026, and the FY4 A pre-training and FY4 B pre-training average increased by the same average of 0.0023.

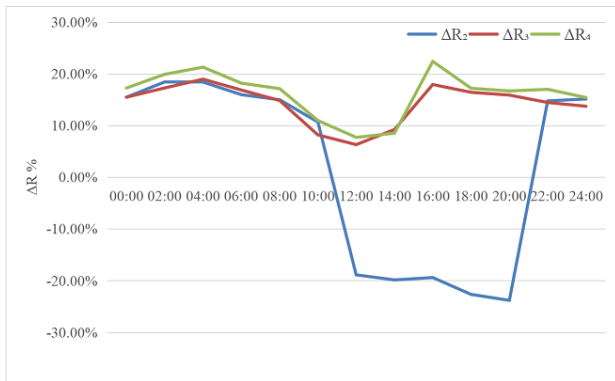


Fig. 8. Difference comparison results of RMSE with traditional bicubic method on Nov. 30, 2021

Fig. 8 shows the percentage of the relative difference line chart of the RMSE between the SR results and bicubic method for one day of testing. The level of promotion varies with time, as test data change substantially between data acquisition moments, * marked in the Tables I-III. For remote sensing images acquired from 12:00 to 20:00 UTC during the near-dark period, the FY4 A pre-training model has difficulty enhancing the resolution well, even worse than bicubic. In contrast, the FY4-SR-Net model with fine-tuning achieved better super-resolution results than the FY4 B pre-training strategy.

The results of comparing the PSNR with the classic bicubic approach on Nov. 30, 2021 are depicted in Fig. 9. The PSNR performance of the three training strategies is comparable to the RMSE. The large performance discrepancy between FY4-SR-Net and FY4 B pre-training implies that transfer learning can significantly enhance the super-resolution impact.

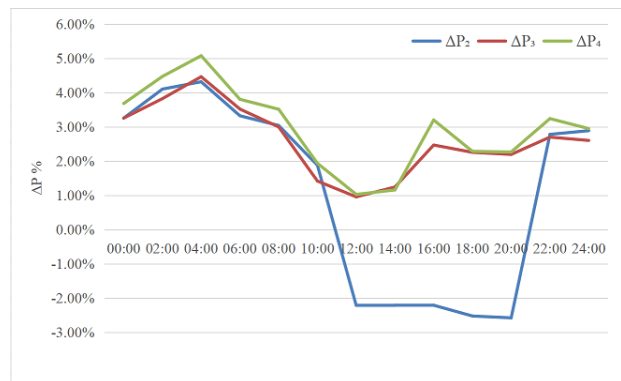


Fig. 9. Difference comparison results of PSNR with traditional bicubic method on Nov. 30, 2021

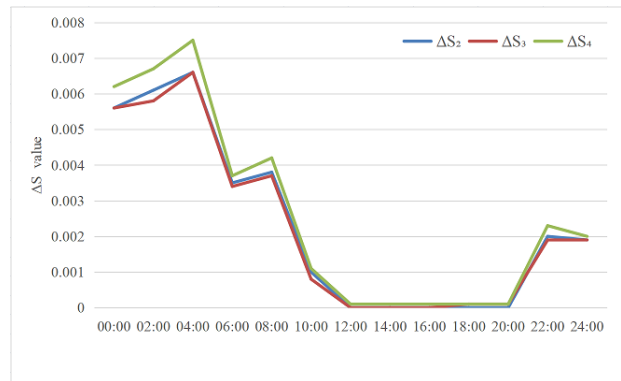


Fig. 10. Different value comparison results of SSIM with traditional bicubic method on Nov. 30, 2021

The SSIM metric might be used to quantify degradation in image quality due to processing, such as data interpolation or data compression. As illustrated in Fig. 10, the FY4-SR-Net model exhibits the greatest super-resolution effect, followed by FY4 B pre-training and FY4 A pre-training. These results demonstrate that transfer learning is effective, particularly when the number of high resolution training samples is limited.

B. Quality of SR result for FY4

We compare the qualitative results of our method to those of the original low-resolution images and bicubic interpolation. In general, bicubic may increase spatial resolution effectively, but the resulting image has a polished texture and limited recovery of cloud information. Figure 11 depicts the SR outcomes for FY4 B 1 km resolution PAN images at a scale of 4. Several cloud-based details have been properly reconstructed in comparison to the initial FY4 B low-resolution data. As demonstrated in Figure 12, the benefit of our proposed approach is also apparent in shaping the cloud textures of FY4 A 1 km resolution VIS images. Visually and quantitatively, the FY4-SR-Net model surpasses the conventional bicubic technique. The FY4-SR-Net method provides improved performance in detail restoration. With this technology, high-resolution FY4 satellite VIS images can be utilized to determine the various types of clouds and their structures.

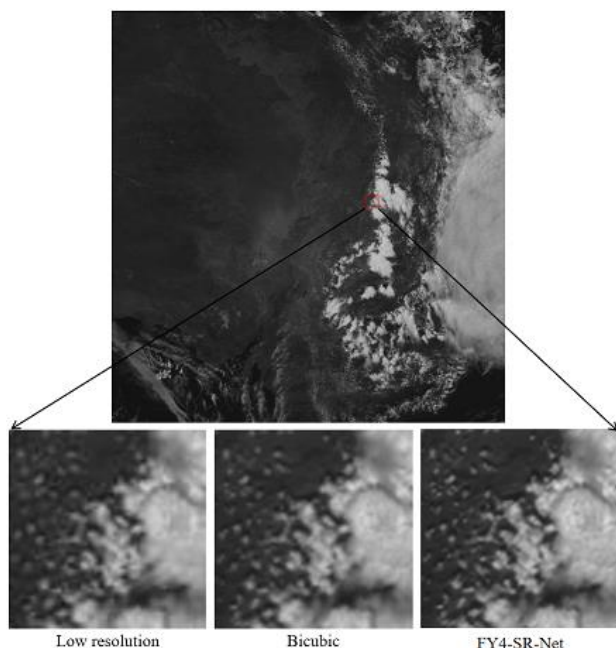


Fig. 11. SR result of FY4 B 1 km data with scale of 4, at UTC 04:00 (12:00 a.m. local time) on Nov.30, 2021

IV. CONCLUSIONS

To meet the super-resolution requirements of the FY4 satellites, a transfer learning method based on pre-training and fine-tuning is presented in this paper. A deep residual network is pre-trained with a large number of FY4 4km resolution and 1 km resolution VIS images and then fine-tuned with the limited 1 km resolution and 0.25 km resolution data. Experiments show that our proposed FY4-SR-Net is better than the current baseline bicubic methods, with an improvement of 16.12% in RMSE, an increase of 2.97% in average PSNR, and an increase of 0.0026 in average SSIM values for monitoring data with a 2-hour interval on November 30, 2021.

With its advantages of high super-resolution accuracy and low cost, FY4-SR-Net has the potential to be widely used for FY4 series satellite VIS image SR in various locations, particularly in mainland China. As a result of this super-resolution technique, the FY4 series of satellites have improved the spatial and temporal resolution of images. This enhances their ability to monitor regional weather events and provide meteorological services.

In conclusion, the development of FY4-SR-Net based on transfer learning can be more precise and effective than interpolation in terms of both qualitative and quantitative resolution enhancement. This is the first time, to our knowledge, that a pre-training and fine-tuning structured SR method has been proposed for the FY4 remote sensing VIS image data. We anticipate that our research will establish a precedent for super-

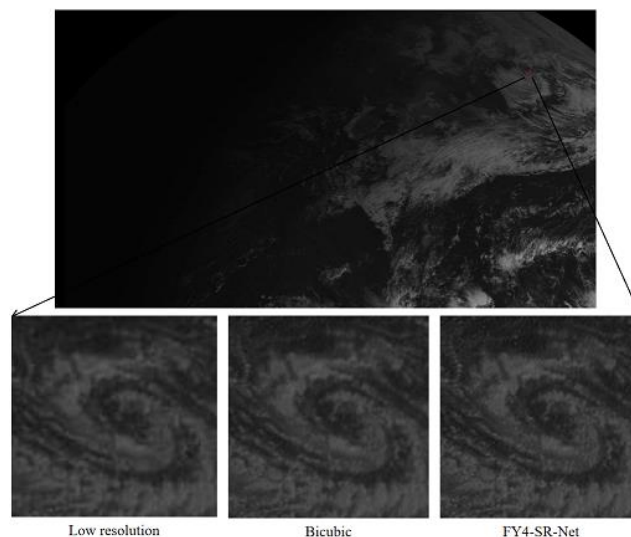


Fig. 12. SR result of FY4 A 1 km data with scale of 4, at UTC 01:23 (09:23 a.m. local time) on Jan. 18, 2022

In short, all three training methods, with the exception of the dark time period, produce superior super-resolution outcomes than bicubic interpolation. The addition of fine-tuning to our model can effectively compensate for and enhance the pre-training outcomes. The pre-training and fine-tuning based FY4-SR-Net model achieved greater super-resolution outcomes than the direct training method.

resolution image processing based on transfer learning. As a result, numerous meteorological satellites will be able to deliver more precise and clearer high-resolution images. In the future, we will perform comparisons with other single-frame SR methods. Our research will focus on enhancing satellite image SR performance at certain time intervals and examining data input and architecture design to further reduce the model prediction errors.

ACKNOWLEDGMENT

The authors would like to thank the National Satellite Meteorological Centre (NSMC) for providing FY4 data sets for the super-resolution task.

REFERENCES

- [1] X. Cai *et al.*, "Temperature and humidity profile retrieval from FY4-GIIRS hyperspectral data using artificial neural networks," *Remote Sens.*, vol. 12, no. 11, pp. 1872, May. 2020, DOI: 10.3390/rs12111872.
- [2] H. Xie *et al.*, "A bibliometric analysis on land degradation: Current status, development, and future directions," *Land.*, vol. 9, no.1, pp. 28, January. 2020, DOI: 10.3390/land9010028.
- [3] Y. Han, C. Deng, B. Zhao and B. Zhao, "Spatial-Temporal Context-Aware Tracking," *IEEE Signal Process. Lett.*, vol. 26, no. 3, pp. 500-504, March. 2019, DOI: 10.1109/LSP.2019.2895962.
- [4] D. Hong *et al.*, "More Diverse Means Better: Multimodal Deep Learning Meets Remote-Sensing Imagery Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340-4354, May. 2021, DOI: 10.1109/TGRS.2020.3016820.
- [5] D. Hong *et al.*, "SpectralFormer: Rethinking Hyperspectral Image Classification With Transformers," *IEEE Trans. Geosci. Remote Sens.*, vol.

- 60, no. 5518615, pp. 1-15, 25 Nov. 2022, DOI: 10.1109/TGRS.2021.3130716.
- [6] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza and J. Chanussot, "Graph Convolutional Networks for Hyperspectral Image Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966-5978, July. 2021, DOI: 10.1109/TGRS.2020.3015157.
- [7] Z. Wang, J. Chen and S. C. H. Hoi, "Deep Learning for Image Super-Resolution: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3365-3387, 1 Oct. 2021, DOI: 10.1109/TPAMI.2020.2982166.
- [8] Z. Lu, C. Wu, X. Yu and C. Hong, "Single image super resolution based on multi-scale structural self similarity and neighborhood regression," *Proc. SPIE 10806, Tenth International Conference on Digital Image Processing (ICDIP 2018)*. 2018. International Society for Optics and Photonics, DOI:10.1117/12.2502973
- [9] Y. Zhang, Q. Fan, F. Bao, Y. Liu and C. Zhang, "Single-Image Super-Resolution Based on Rational Fractal Interpolation," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 3782-3797, Aug. 2018, DOI: 10.1109/TIP.2018.2826139.
- [10] D. Hong *et al.*, "Interpretable hyperspectral artificial intelligence: When nonconvex modeling meets hyperspectral," *Remote Sens.*, vol. 9, no. 2, pp. 52 - 87, June. 2021, DOI: 10.1109/MGRS.2021.3064051.
- [11] L. Wang *et al.*, "Parallax Attention for Unsupervised Stereo Correspondence Learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 4, pp. 2108-2125, 1 April 2022, DOI: 10.1109/TPAMI.2020.3026899.
- [12] L. Wang *et al.*, "Exploring sparsity in image super-resolution for efficient inference," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp.4917-4926, 2021, DOI: 10.48550/arXiv.2006.09603.
- [13] M. Liang *et al.*, "Video Super-Resolution Reconstruction Based on Deep Learning and Spatio-Temporal Feature Self-similarity," *IEEE Trans. Knowl. Data Eng.*, 2022, pp. 1, DOI: 10.1109/TKDE.2020.3034261.
- [14] J. Xin, Q. Xiong, C. Xiong, Z. Li and Z. Gao "Single image super-resolution with multi-level feature fusion recursive network," *eurocomputing*, vol. 370, no. 22, pp. 166-173, April 2019, DOI: 0.1016/j.neucom.2019.06.102.
- [15] W. Wang, Y. Hu, Y. Luo and T. Zhang, "Brief Survey of Single Image Super-Resolution Reconstruction Based on Deep Learning Approaches," *Sens Imaging*, vol. 21, no. 21, April 2020, DOI: 10.1007/s11220-020-00285-4.
- [16] R. Hu and X.Cui, "Application of single frame image super-resolution algorithm based on generative adversarial network in tennis motion image resolution," *J Ambient Intell Human Comput*, vol. 21, pp. 1-11, April 2021, DOI: 10.1007/s12652-021-03100-4.
- [17] X. Du, X. Qu, Y. He and D. Guo, "Single Image Super-Resolution Based on Multi-Scale Competitive Convolutional Neural Network," *Sensors*, vol. 370, no. 3, pp. 789, March 2018, DOI: 10.3390/s18030789.
- [18] J. Xu, Y. Liang, J. Liu, and Z. Huang, "Multi-Frame Super-Resolution of Gaofen-4 Remote Sensing Images," *Sensors*, vol. 17, no. 9, pp. 2142, Sep. 2017, DOI: 10.3390/s17092142.
- [19] M. Merino, J. Nunez. "Super-resolution of remotely sensed images with variable-pixel linear reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1446-1457, May 2007, DOI: 10.1109/TGRS.2007.893271.
- [20] A. Fruchter, R. Hook. "Drizzle: A method for the linear reconstruction of under sampled images," *Publ. Astron. Soc. Pac.*, vol. 114, pp. 144, May. 2002, DOI: 10.1086/338393
- [21] H. Shen, M. K. Ng, P. Li and L. Zhang, "Super-Resolution Reconstruction Algorithm To MODIS Remote Sensing Images," *Comput. J.*, vol. 52, no. 1, pp. 90-100, Jan. 2009, DOI: 10.1093/comjnl/bxm028.
- [22] F. Li, X. Jia, D. Fraser and A. Lambert, "Super Resolution for Remote Sensing Images Based on a Universal Hidden Markov Tree Model," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1270-1278, March. 2010, DOI: 10.1109/TGRS.2009.2031636.
- [23] C. Fan, C.Wu, G. Li, J. Ma, "Projections onto Convex Sets Super-Resolution Reconstruction Based on Point Spread Function Estimation of Low-Resolution Remote Sensing Images," *Sensors (Basel)*, vol. 17, no. 2, pp. 362, Feb. 2017, DOI: 10.3390/s17020362.
- [24] M. Shao, X. Yuan, J. Ji, Y. Zhang, S. Wan and Q. Du, "Hyperspectral Image Spatial Super-Resolution via 3D Full Convolutional Neural Network," *Remote Sens.*, vol. 9, no. 11, pp. 1139, Nov. 2017, DOI: 10.3390/rs9111139.
- [25] J. Fu, Y. Liu and F.Li, "Single frame super resolution with convolutional neural network for remote sensing imagery," *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, pp. 8014-8017, 2018, DOI: 10.1109/IGARSS.2018.8518584.
- [26] D. Hong *et al.*, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923-1938, April 2019, DOI: 10.1109/TIP.2018.2878958.
- [27] J. M. Haut *et al.*, "A new deep generative network for unsupervised remote sensing single-image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6792-6810, Nov. 2018, DOI: 10.1109/TGRS.2018.2843525.
- [28] X. Dong, Z. Xi, X. Sun and L. Gao, "Transferred Multi-Perception Attention Networks for Remote Sensing Image Super-Resolution," *Remote Sens.*, vol. 11, no. 23, pp. 2875, Dec. 2019, DOI: 10.3390/rs11232857.
- [29] X. Dong. *et al.*, "Remote sensing image super-resolution using novel dense-sampling networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1618-1633, Feb. 2021, DOI: 10.1109/TGRS.2020.2994253.
- [30] S. Lei, Z. Shi, X. Wu, B. Pan, X. Xu and H. Hao, "Simultaneous Super-Resolution and Segmentation for Remote Sensing Images," *Proc. IEEE Int. Geosci. Remote Sens. Symp. IGARSS*, pp. 3121-3124, 2019, DOI: 10.1109/IGARSS.2019.8900402.
- [31] S. Lei, Z. Shi and Z. Zou, "Coupled Adversarial Training for Remote Sensing Image Super-Resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3633-3643, May 2020, DOI: 10.1109/TGRS.2019.2959020.
- [32] C. Shorten, T.M. Khoshgftaar. "A survey on Image Data Augmentation for Deep Learning," *J Big Data*, vol. 6, pp. 60, April. 2019, DOI: 10.1186/s40537-019-0197-0.
- [33] J. Yang, Z. Zhang, C. Wei, F. Lu and Q. Guo, "Introducing the new generation of Chinese geostationary weather satellites, Fengyun-4," *Bull. Amer. Meteorol. Soc.*, vol. 98, no. 8, pp. 1637-1658, Aug. 2017, DOI: 10.1175/BAMS-D-16-0065.1.
- [34] H. Fu, Y. Shen, J. Liu, G. He, J. Chen, P. Liu, J. Qian and J. Li, "Cloud Detection for FY Meteorology Satellite Based on Ensemble Thresholds and Random Forests Approach," *Remote Sens.* vol. 11, no. 1, pp. 44, Jan. 2019, DOI: 10.3390/rs11010044.
- [35] X. Wang, M. Min, F. Wang, J. Guo, B. Li and S. Tang, "Intercomparisons of Cloud Mask Products Among Fengyun-4A, Himawari-8, and MODIS," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8827-8839, Nov. 2019, DOI: 10.1109/TGRS.2019.2923247.
- [36] X. Wu, D. Hong and J. Chanussot, "Convolutional Neural Networks for Multimodal Remote Sensing Data Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1-10, 2022, Art no. 5517010, DOI: 10.1109/TGRS.2021.3124913.
- [37] Z.Zhu and C. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83-94, March 2012, DOI: 10.1016/j.rse.2011.10.028.
- [38] F. Ye, Y. Guo, and P. Zhuang, "Pan-sharpening via a gradient-based deep network prior," *Signal Process. Image Commun.*, vol. 74, pp. 322-331, May 2019, DOI:10.1016/j.image.2019.03.004
- [39] J. Kim, J. K. Lee and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks", *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 1646-1654, Jun. 2016, DOI: 10.1109/CVPR.2016.182.
- [40] W. Shi *et al.*, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," *Proc. IEEE Conf. CVPR*, 2016, pp. 1874-1883.
- [41] S. Dubey and S. Chakraborty, "Average biased ReLU based CNN descriptor for improved face retrieval," *Multimed Tools Appl.*, vol. 80, no. 15, pp. 23181-23206, Dec. 2021, DOI: 10.1007/s11042-020-10269-x.
- [42] N. Ketkar and J. Moolayil, "Introduction to PyTorch. In: Deep Learning with Python," *Apress, Berkeley, CA*, pp. 27-91, April 2021, DOI: 10.1007/978-1-4842-5364-9_2.
- [43] Z. Huang and L. Cao, "Bicubic interpolation and extrapolation iteration method for high resolution digital holographic reconstruction." *Opt. Lasers Eng.* vol. 130, pp. 106090, July 2020, DOI: 10.1016/j.optlaseng.2020.106090.



Bo Zhang received the B.S. degree from Anhui Normal University, Anhui, China, in 2014. He received the Ph.D. degree from Xi'an Jiaotong University, Xi'an, China, in 2020. He is currently a post-doctor at Tsinghua University, Beijing, China. His research interests include GIS, remote sensing, matching learning, remote sensing image processing, and pattern recognition.



Muyuan Ma received his B.E. Degree from the School of Marine Science and Technology, Tianjin University, Tianjin, China. He is going to pursue the Ph.D degree with the Department of Earth System Science, Tsinghua University, Beijing, China. His research interests include machine learning, remote sensing, and deep learning in remote sensing applications.



Mingqing Wang received the B.S degree in Information and Computing Sciences and the M.S. degree in Applied Mathematics from Taiyuan University of Technology, Taiyuan, China in 2011 and 2014, respectively. He is currently a Ph.D candidate with the Department of Earth System Science, Tsinghua University, Beijing, China. His research interests include scientific computation, machine learning, and AI in geoscience.



Danfeng Hong (S'16--M'19--SM'21) received the M.Sc. degree (summa cum laude) in computer vision from the College of Information Engineering, Qingdao University, Qingdao, China, in 2015, the Dr. -Ing degree (summa cum laude) from the Signal Processing in Earth Observation (SiPEO), Technical University of Munich (TUM), Munich, Germany, in 2019. He is currently a Professor with the Key Laboratory of Computational Optical Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences (CAS). Before joining CAS, he has been a Research Scientist and led a Spectral Vision Working Group at the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany. He was also an Adjunct Scientist at GIPSA-lab, Grenoble INP, CNRS, Univ. Grenoble Alpes, Grenoble, France. His research interests include signal / image processing, hyperspectral remote sensing, machine / deep learning, artificial intelligence, and their applications in Earth Vision.



Le Yu (Senior Member, IEEE) received the B.S. degree in geographical information system and the Ph.D. degree in remote sensing from Zhejiang University, Hangzhou, China, in 2005 and 2010, respectively. He is currently an Associate Professor with the Department of Earth System Science, Tsinghua University, Beijing, China. His research interests include global land cover/use mapping, simulation, and applications. Dr. Yu is an Associate Editor of International Journal of Remote Sensing and the Section Editor-in-Chief of Land. He has served as a member of Editorial Board and a referee for many international journals.



Jie Wang received his Ph.D. degree from the School of Mathematical Sciences, Beijing Normal University, Beijing China in 2009. He received a B.A degree in 2001 and a M.S. degree in 2004 from Beijing Normal University, Beijing China. He is currently an associate professor in the Peng Cheng Laboratory. His major research interest is machine learning for remote sensing data processing and land cover change monitoring.



Peng Gong is Chair Professor of Global Sustainability at the University of Hong Kong. He built the Department of Earth System Science and served as Dean of School of Sciences at Tsinghua University. He also served as the founding director of Tsinghua Urban Institute. He had previously taught at the University of Calgary and the University of California, Berkeley. His major research interests include mapping, monitoring and modeling of global environmental change, and modelling of environmentally related infectious diseases such as schistosomiasis, avian influenza, dengue and COVID-19, and healthy and sustainable cities. He is the author/co-author of over 600 articles and 8 books. He chaired/co-chaired 9 Lancet Commission reports on climate change and health, and healthy cities in China.



Xiaomeng Huang received his Ph.D degree from the Department of Computer Science and Technology, Tsinghua University, Beijing, China in 2007. He received a B.A degree in 2000 from Wuhan University, Wuhan, China and a M.S. degree in 2003 from Huazhong University of Science and Technology, Wuhan, China. He is currently an associate professor at Tsinghua University. He is especially interested in big Earth system models, big data, and AI in geoscience.