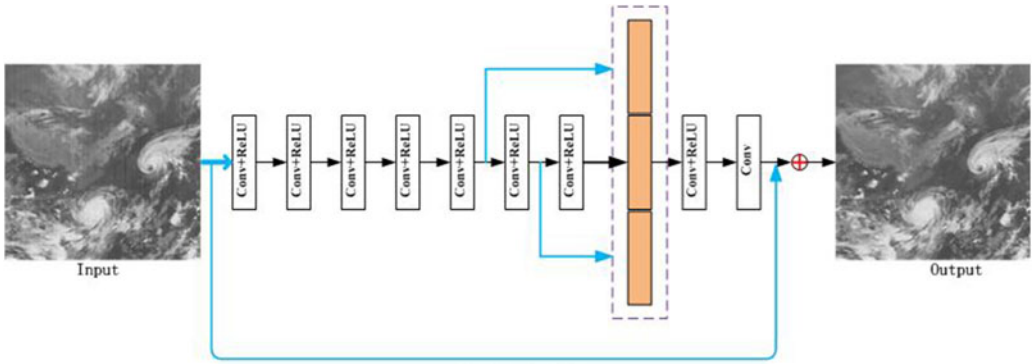
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

Removing Stripe Noise From Infrared Cloud Images via Deep Convolutional Networks

Volume 10, Number 4, August 2018

Pengfei Xiao
Yecai Guo
Peixian Zhuang



Graphic Abstract

DOI: 10.1109/JPHOT.2018.2854303

Removing Stripe Noise From Infrared Cloud Images via Deep Convolutional Networks

Pengfei Xiao, Yecai Guo, and Peixian Zhuang

Nanjing University of Information Science and Technology, Nanjing 210044, China

DOI:10.1109/JPHOT.2018.2854303

Manuscript received February 7, 2018; revised May 27, 2018; accepted July 4, 2018. Date of publication July 25, 2018; date of current version July 30, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61673222 and 61701245. Corresponding author: Yecai Guo (e-mail: guo-yecai@163.com).

Abstract: We propose a new deep network architecture for removing a stripe noise from a single meteorological satellite infrared cloud image. In the proposed framework, a residual learning is utilized to directly reduce the mapping range from input to output, which speeds up the training process as well as boosts the destriping performance. Inspired by the wide inference networks, we use wider CNNs with more convolutions in the first part of the proposed network, which is helpful for learning the similar pixel-distribution features from noisy images. To further improve the performance, we propose a local-global combination structure model, which combines the representations of different layers for recovering the rich details of infrared cloud images. Moreover, we extend our method to remove rain streaks from single images, which provides a new idea for rain-removal task. In addition, we provide a new meteorological satellite infrared cloud image dataset for training and validating the proposed network. Final extensive experiments demonstrate that the proposed method can achieve both comparable restoration quality and computational efficiency with several state-of-the-art approaches.

Index Terms: Stripe noise removal, infrared cloud image, deep convolutional neural network, local-global combination, residual learning.

1. Introduction

As an important branch of remote sensing technology, meteorological satellite cloud images have been widely used to predict the weather changes [1]. Meteorological satellite cloud images mainly consist of infrared cloud images and visible cloud images. Compared with visible cloud images, infrared cloud images can reflect the thermal radiation characteristic of the target and can be obtained at night. Therefore, infrared cloud images are widely used. However, due to the inconsistent responses among different detectors [2], infrared cloud images are easily contaminated with stripe noise, which makes the structure of cloud hard to recognize and affects the subsequent higher level data applications, such as cloud classification [3]. In order to improve the visual quality of infrared cloud images for weather forecasting, it is crucial to remove stripe noise and to recover original structure of cloud.

From the viewpoint of periodicity, stripe noise can be classified into two categories, periodic stripe noise and aperiodic stripe noise. Compared with periodic stripe noise, aperiodic stripe noise is more common and more difficult to deal with, especially when the original images are rich in

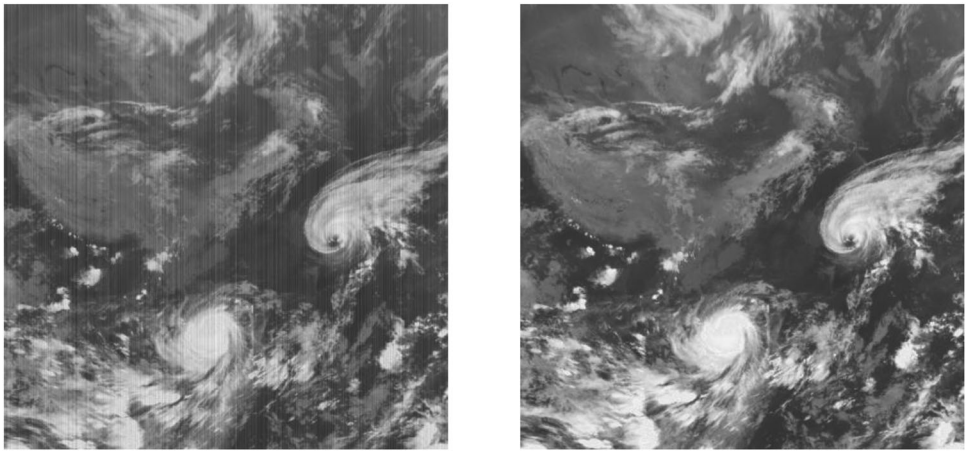


Fig. 1. An example of a noisy test image and our result.

details such as meteorological satellite cloud images. An end-to-end deep network architecture for removing stripe noise from infrared cloud images is developed. Fig. 1 shows an example of a noisy test image and our result. Until now, a certain number of methods have been proposed to remove stripe noise from remote sensing images. From the viewpoint of underlying similarity, these methods can be classified into several groups. We briefly review these approaches and then discuss the contributions of our proposed framework.

1.1 Related Work

The first family of destriping approaches is filter-based methods, including Fourier domain filter [4]–[6], wavelet analysis [7], [8], and the combined wavelet-Fourier filter [9], [10]. These methods are based on the assumption that stripe noise are periodic, and then they analyze the property of stripe noise in a transformed domain. Since the periodic feature of stripe noise can be easily identified in a power spectrum analysis, the filter-based methods hold the advantage in processing periodic stripe noise over aperiodic stripe noise. However, original image structures can also be filtered if they possess the same frequency as stripe noise, further resulting in excessive blurring or ring artifacts. [4], [5], [11]. To conquer this drawback, more accurate approaches of stripe information truncation [12], [13] have been found.

Scene-based techniques, the second category of methods, utilize selected stable scenes to compute the relationship between different detectors before correcting the other images. For example, based on the assumption that the response of the sensor is linear, Fischer *et al.* [14] computed the correction coefficients for hyperspectral SWIR data by using selected stable scene images; with the help of ice-sheet images, Bindschadler and Choi [15] estimated the detector errors in Hyperion data. However, the limitation of scene-based methods is the artifacts produced by the fixed coefficients, since the valid coefficients may vary widely with different scene contents [14].

The interpolation-based methods constitute another category of destriping methods. These techniques first detect the locations of stripes, and then the pixels affected by stripes are replaced with reasonable values calculated by designed interpolation function [17]. The main characteristic of these methods is simple and robust when the type of stripes is single [17]. However, one limitation of these methods is their poor destriping capacity when the width of the stripe is relatively large, since interpolation is impossible in this case [18]. Moreover, it may be difficult to detect the stripes when the striped images are complex [18], [19].

Another category of destriping methods, the statistical-based methods, utilize the statistical properties of digital numbers of the fine sensors. These methods first examine distributions of each detector and then rectify them to the reference one [20], [21]. These techniques, such as equal-

ization methods [22], histogram modification [23], [24] and moment matching [19], are typical and widely applied in the field of stripe removal. Despite the efficiency of these methods, the efficacy of them is poor when the stripes are nonlinear. To conquer this drawback, a piece-wise approach was proposed to remove the nonlinear and irregular stripes in MODIS data [25]. Nevertheless, these methods are just applicable for specific stripes.

Recently, optimization-based models, which view destriping issue as an ill-posed problem, have become a mainstream category of destriping techniques [26]–[32]. Combining prior information with regularization terms, the latent clean image can be estimated by minimizing the energy function. Assuming that normal pixels obey Huber-Markov distribution, Shen and Zhang [26] first tried to solve the destriping problem under a maximum a posteriori framework. In [27], a more complex model was built by considering the direction of stripes. Subsequently, based on sparse representation and unidirectional total variation, a joint model was proposed in [28] to remove random stripe noise. Utilizing the latent sparsity of stripes in data, stripe noises were separated from true image in [31]. Furthermore, in order to remove highly dense stripes, a variational model based on the statistical properties of stripes was proposed in [32]. Liu *et al.* [33] proposed a universal destriping framework. The clean image can be estimated by 2-D variational optimization, which is guided by the statistical feature-based information estimated by 1-D variation method. Compared with other methods, optimization-based methods produce relatively balanced results both quantitatively and qualitatively. However, when images contain complex structures, the destriped images tend to be over-smoothed, even though the stripes are almost removed. Besides, the computation time is still a problem.

Until now, Kuang *et al.* [16] was the first to handle the infrared image stripe removal task via a deep learning method (named SNRCNN). SNRCNN is a three-layer convolution neural network, which directly learns a mapping from noisy image to clean image in an end-to-end manner. It is hard for SNRCNN to remove stripes thoroughly and recover rich details due to the limited capacity of the three-layer network. Therefore, remaining stripes and over-smoothing are common in the results of SNRCNN.

1.2 Our Contributions

Motivated by the success of deep convolutional neural network (CNN) on several low level vision tasks, such as image denoising [34], deconvolution [35], super-resolution [36], [37], [45] and in-painting [38], a novel network is designed for stripe noise removal from single infrared cloud images based on deep CNN, and it produces excellent performance. To the best of our knowledge, this is the first method based on deep learning methods for addressing this problem. We name the proposed model Infrared Cloud image Stripe Removal Network (ICSRN). Our main contributions are as follows:

1) ICSRN adopts the residual learning strategy to predict the residual (difference between clean and striped image), since it can significantly reduce the mapping range, which is helpful for fast convergence and better performance. Motivated by the discovery in [44], we increase kernel size of the first several layers to learn similar pixel distribution from striped images, which is greatly helpful for predicting the residual. Moreover, we propose a local-global combination structure model to combine representations of different layers for recovering the rich details of infrared cloud images.
2) We create and use a new dataset of 200 striped/clean image pairs for training and validating the network. Although without re-training, the model trained on our infrared cloud image dataset shows excellent performance on common infrared images, since infrared cloud images contain more complex features than common infrared images, and these complex features extracted by CNN are general for dealing with common infrared images.
3) ICSRN is applied to remove vertical rain streaks from single images and outperforms deep detail network [51], which provides a new idea for rain-removal task that real-world rain streaks can be decomposed into horizontal and vertical components. Meanwhile, ICSRN can be utilized to deal with each component separately.

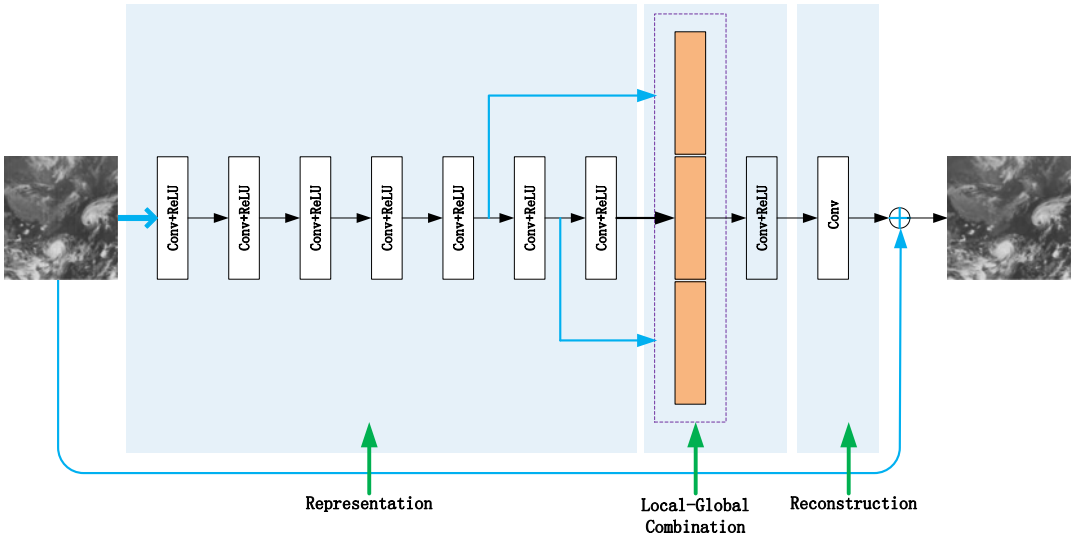


Fig. 2. The architecture of the proposed ICSRN network.

2. Proposed Destriping Model

In this section, we present the proposed destriping CNN model (ICSRN) and extend it for handling rain-removal task. Generally, network design and model learning from training data are two necessary steps of a deep CNN model. By simultaneously taking the properties of stripe noise and infrared cloud images into consideration, we design a deep CNN network for stripe removal of infrared cloud images. For model learning, we adopt the residual learning strategy for fast convergence and improved destriping performance. Finally, we extend ICSRN for rain-removal task.

2.1 Network Architecture

The input of ICSRN is an observed striped infrared cloud image and the stripes are considered to be additive noise [19], [27], [28], the degradation model can be described as

$$\mathbf{Y} = \mathbf{X} + \mathbf{N} \tag{1}$$

where $\mathbf{Y}$ is the observed striped infrared cloud image and $\mathbf{X}$ is the latent clean infrared cloud image. The term $\mathbf{N}$ represents the stripe component and minor random noise. Since ICSRN adopts residual learning strategy, the output of the network is the estimated $\mathrm{R}(\mathbf{Y}) \approx -\mathbf{N}$, and then we have $\mathbf{X} \approx \mathbf{Y} + \mathrm{R}(\mathbf{Y})$. We adopt mean square error (MSE) as the cost function, and the optimization objective can be represented as

$$\mathrm{L}(\theta) = \frac{1}{2n} \sum_{i=1}^{n} \left\| \mathrm{R}(\mathbf{Y}_i; \theta) - (\mathbf{X}_i - \mathbf{Y}_i) \right\|_F^2 \tag{2}$$

where $\mathbf{Y}_i$ and $\mathbf{X}_i$ are the noisy-clean sub-image pair in a patch, θ are the trainable parameters, "F" means Frobenius norm. The proposed destriping CNN model (ICSRN) is illustrated in Fig. 2 and the detailed configurations is presented in Table 1. ICSRN consists of three main parts: representation, local-global combination and reconstruction. In the following, we explain the architecture of ICSRN.

2.1.1 Representation: Seven convolutional layers followed by rectified linear units (ReLU) are utilized by this part. Batch normalization is not used in ICSRN as Lim B *et al.* [42] presented in their super-resolution work, since batch normalization limits range flexibility by normalizing the features. Detailed configurations are presented in Table 1. The network structure of this part can

TABLE 1
Detailed Configurations of ICSRN

Three Main Parts	Layer	Filter Size	Filter Number	Stride	Pad
Representation	conv1	7×7	64	1	3
	conv2	$64\times7\times7$	64	1	3
	conv3	$64\times7\times7$	64	1	3
	conv4	$64\times7\times7$	64	1	3
	conv5	$64\times3\times3$	64	1	1
	conv6	$64\times3\times3$	64	1	1
	conv7	$64\times3\times3$	64	1	1
Local-Global Combination	concat	conv5+conv6+conv7			
	conv8	$64\times3\times3$	64	1	1
Reconstruction	conv9	$64\times3\times3$	1	1	1

be expressed as:

$$f_i = \sigma(\mathbf{W}_i * f_{i-1} + \mathbf{b}_i) \tag{3}$$

where i = 1, ...,n1, n1 represents the total number of layers in this part, f_0 is the input, '$*$' indicates the convolution operation, **W** contains weights and **b** biases, σ is a rectified linear units (ReLU) for non-linearity.

For the first four layers, 64 filters of size c $\times$ 7 $\times$ 7 are used to generate 64 feature maps and rectified linear units (ReLU) are utilized for non-linearity. Since the infrared cloud images in this paper have only one channel, so c is set to 1 for the first layer and 64 for the rest. Architecture design of first four layers is inspired by the contributions of Wide Inference Network [44]. Instead of depth, "width" is the key in low-level vision tasks like image denoising. Inference mapping primarily relies on the priors behind the noise property and wider CNNs can learn the similar pixel-distribution features from noisy images. Many types of noise follow a certain distribution. Taking additive white Gaussian noise for example, as long as noise level and size of original images are the same, different noisy images tend to follow similar pixel distributions. In infrared cloud image destriping task, we discover the similar phenomenon. As shown in Fig. 3, two noisy infrared cloud images have similar pixel distributions despite the great different features in original images. Therefore, Wide Inference Network can be utilized for infrared cloud image stripe noise removal. In order to achieve balance between performance and computation complexity, we compare average PSNR on the same validation dataset during training with different filter size (F), different number of layers (L) and different number of filters (K), respectively in this part. When training with different filter size, we ensure that the receptive field is the same. All comparing experiments are added aperiodic stripe noise with the noise level $\alpha = 35$ (standard deviation of noise), and the data sets we provide are introduced in the Section 3.1.

1) *Filter Size (F):* In our network, $F = 7 \times 7$ is used which is the same as Wide Inference Network. As we can see in Fig. 4(a), it can remarkably improve performance compared with smaller F.

2) *Layer Number (L):* As shown in Fig. 4(b), the network with $L = 4$ achieves remarkable performance gains than the network with $L = 3$ and $L = 5$. $L = 4$ is most suitable for high performance in the destriping task.

3) *Filter Number (K):* From Fig. 4(c), Network with $K = 128$ achieves better performance than the network with $K = 64$. However, the training time and computation complexity grow vastly compared to the network with $K = 64$. Moreover, these four layers are just one part of our network, a large number of parameters of these layers will affect the design of the next part. Therefore, $K = 64$ is found to be the optimal value.

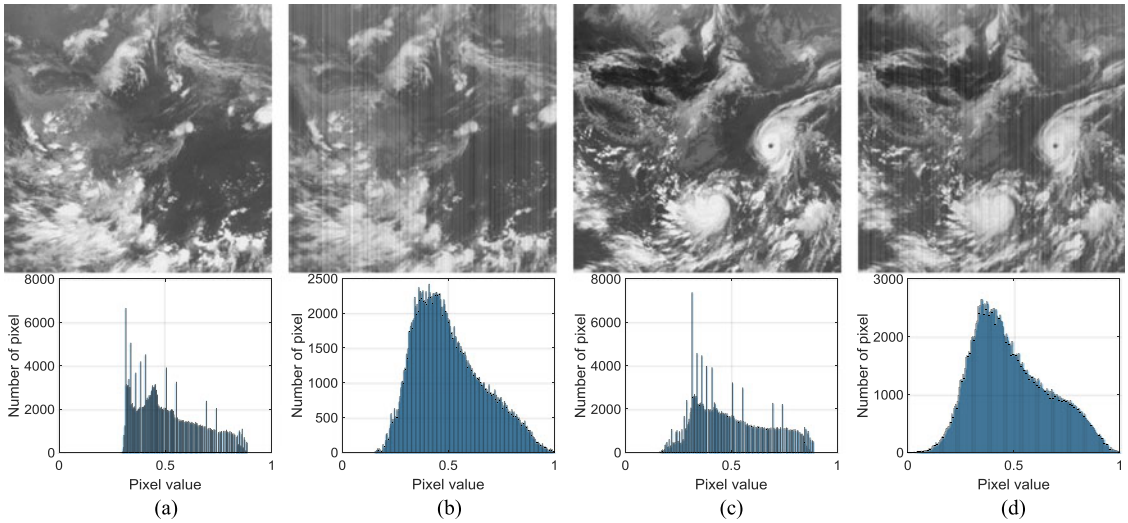


Fig. 3. Similar histogram distributions of two different images added stripe noise with the same noise level. (a) original image-1 (b) noisy-1 (c) original image-2 (d) noisy-2.

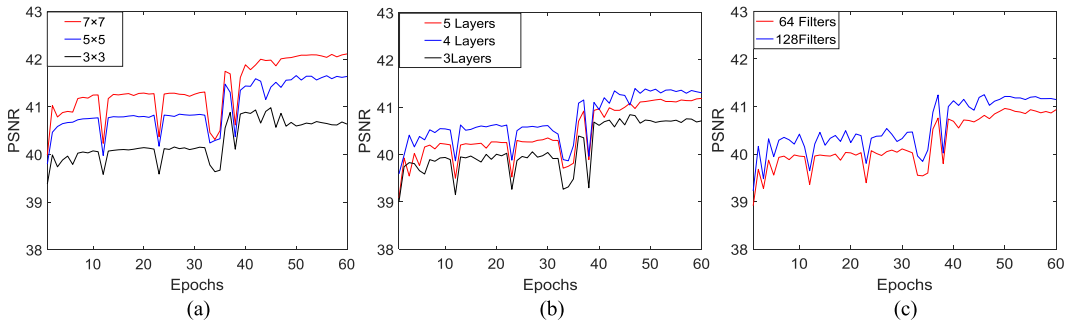


Fig. 4. Comparison of average PSNR on validation dataset training for CNNs with different structure components. (a) filter size; (b) number of layers; (c) number of filters.

For the next three layers, 64 filters of size $64 \times 3 \times 3$ are used to generate 64 feature maps and rectified linear units (ReLU) are utilized for non-linearity. Different from common natural images, infrared cloud images are rich in details due to the complex structure of cloud which determines that our task is different from common denoising tasks. Stacking layers with filters of the same size $64 \times 7 \times 7$ will lead to filters that become rapidly global, which makes local details hard to recover. Inspired by super-resolution networks such as VDSR [45] in recovering details, we use filters of size $64 \times 3 \times 3$ in these three layers to represent more relatively local features, which helps ICSRN better infer high-frequency components. The representations of these three layers are utilized by the next part for local-global combination.

2.1.2 Local-Global Combination: This part first concatenates convolutional results of several layers, then utilizes a convolutional layer to merge these combined representations for final reconstruction. The concatenated result is expressed as:

$$f_c = [f_i, f_j, f_k, \ldots] \tag{4}$$

where f_i, f_j, f_k are representations of different layers. Then the overall structure can be expressed as:

$$f_{\lg c} = \sigma(\mathbf{W}_{\lg c} * f_c + \mathbf{b}_{\lg c}) \tag{5}$$

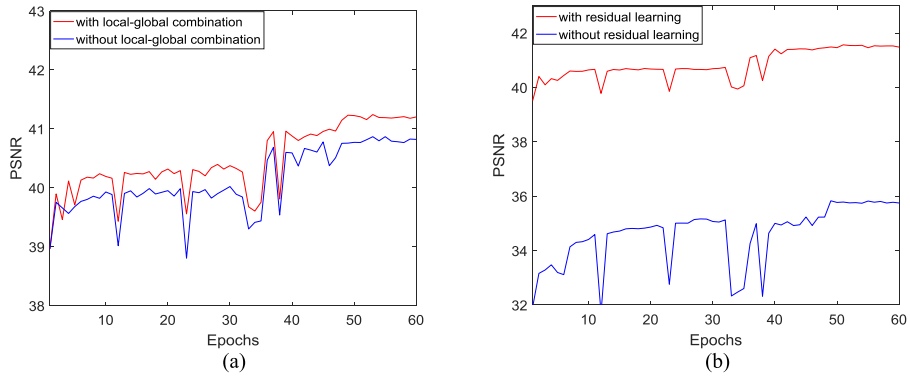


Fig. 5. Comparison of average PSNR on validation dataset training for CNNs with (a) the network without local-global combination (b) the network without residual learning.

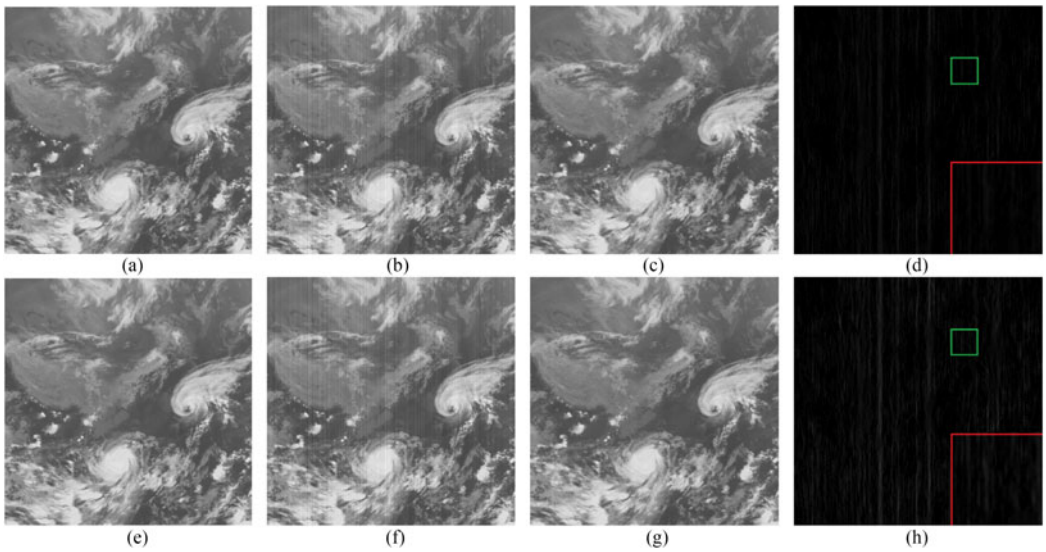


Fig. 6. Example of destriping results. (a), (e): original image $\mathbf{x}$. (b), (f): noisy image. (c): destriping result by the proposed network $\mathbf{x_I}$ (g) destriping result by the network without local-global combination $\mathbf{x_J}$. (d): $|\mathbf{x_I} - \mathbf{x}|$. (h): $|\mathbf{x_J} - \mathbf{x}|$.

In the first part, different features are represented by utilizing filters of different size and this part combines different level representations, including relatively local details and global environmental prior for better destriping performance. To verify the effectiveness of this part, we design a set of experiments to compare the performance of our ICSRN and the network without local-global combination both qualitatively and quantitatively. We compare the average PSNR on the same validation dataset during training. All comparing experiments are added stripe noise with the noise level $\alpha = 35$, and the data sets we provide are introduced in the Section 3.1.

1) *Quantitative Comparison:* As we can see in Fig. 5(a), ICSRN achieves remarkable performance gains than the network without local-global combination. Local-global combination is demonstrated to be suitable for our network.

2) *Qualitative Comparison:* Fig. 6 illustrates the visual results of our proposed model and the network without local-global combination. From Fig. 6(d) and (h), we can see that ICSRN preserves more original details by local-global combination, since more white pixels remain in Fig. 6(h) than in Fig. 6(d), which denotes the difference between estimated result and original image.

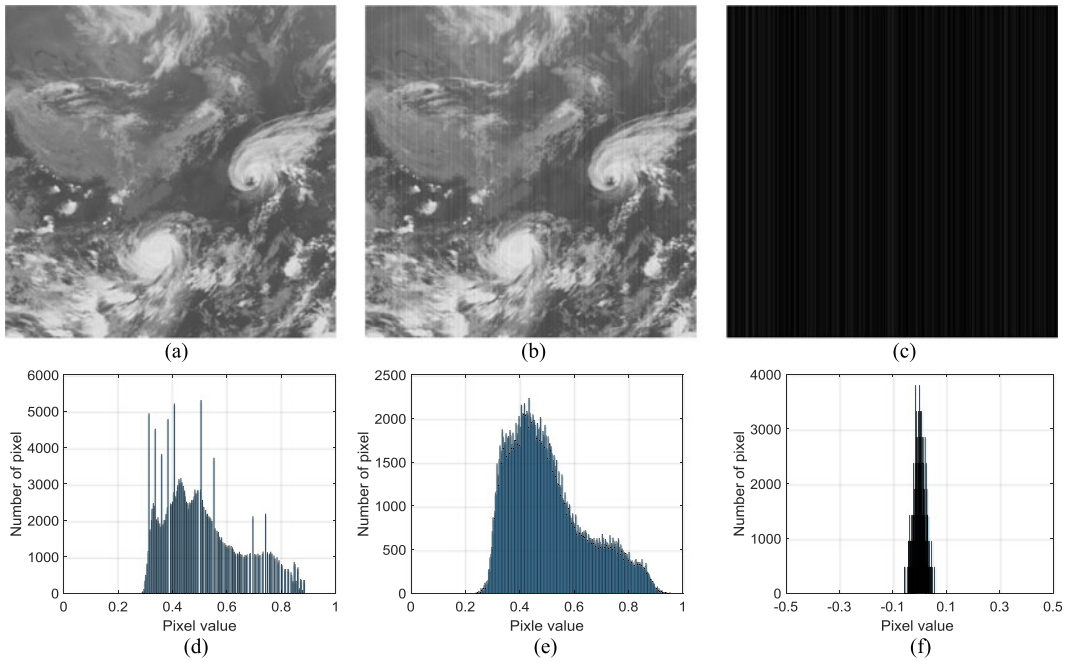


Fig. 7. Range reduction of the residual layer. (a): original image **x** (b): noisy image **y** (c): **x**–**y** (d):histogram of **x** (e): histogram of **y** (f): histogram of **x**–**y**.

2.1.3 Reconstruction: In this part, only one convolutional layer is utilized to output the residual

$$\mathbf{R} = \mathbf{W} * f_{\lg c} + \mathbf{b}_f \tag{6}$$

and the final destriped image $\hat{\mathbf{X}}$ can be obtained by adding the noisy image

$$\hat{\mathbf{X}} = \mathbf{Y} + \mathbf{R} \tag{7}$$

2.2 Learning Strategy: Residual Learning

Learning a residual representation is able to reduce the solution space by compressing the mapping range, since the noisy image is much more like the latent clean image than the residual image, which is helpful for optimizing the residual mapping [40].We can see from Fig. 7(d) and (f), the original image **X** has a much wider range in pixel values than residual of noisy image $\mathbf{X} - \mathbf{Y}$. Skip connection from input to end, like the one employed in VDSR [46], is able to propagate the lossless information through the entire network, which is helpful for the final destriped image. As we can see in Fig. 5(b), residual learning can result in more stable convergence and can exhibit improved destriping performance than learning original mapping.

2.3 Extension to Rain-Removal Task

Since our model shows excellent performance in destriping task, we try to apply ICSRN to deal with the noise which is similar with stripe noise. As we can see from Fig. 8(b) and (f), the vertical rain streaks and stripes are similar in shape. Therefore, we utilize ICSRN to remove vertical rain streaks from single images and achieve excellent performance.

We re-train a model with the synthetic dataset provided [51] and compare the results with method [51], which is the state-of-the-art de-raining method. As we can see in Fig. 8, deep detail network [51] remains some rain streaks and results in blurring while our method can remove rain streaks completely and preserve structure. We use PSNR and SSIM [47] for a quantitative evaluation. From

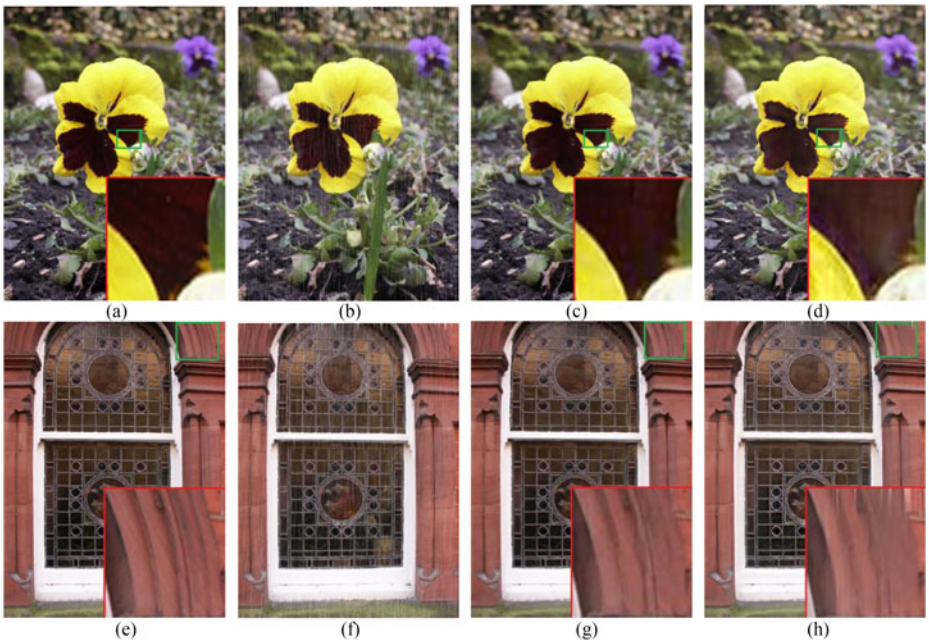


Fig. 8. Visual results of two images [51] along with PSNR/SSIM. (a) original image (b) rainy image (c) ICSRN/**35.08**/**0.95** (d) method [51]/27.49/0.89 (e) original image (f) rainy image (g) ICSRN/ **33.84**/**0.91** (h) method [51]/27.57/0.85.

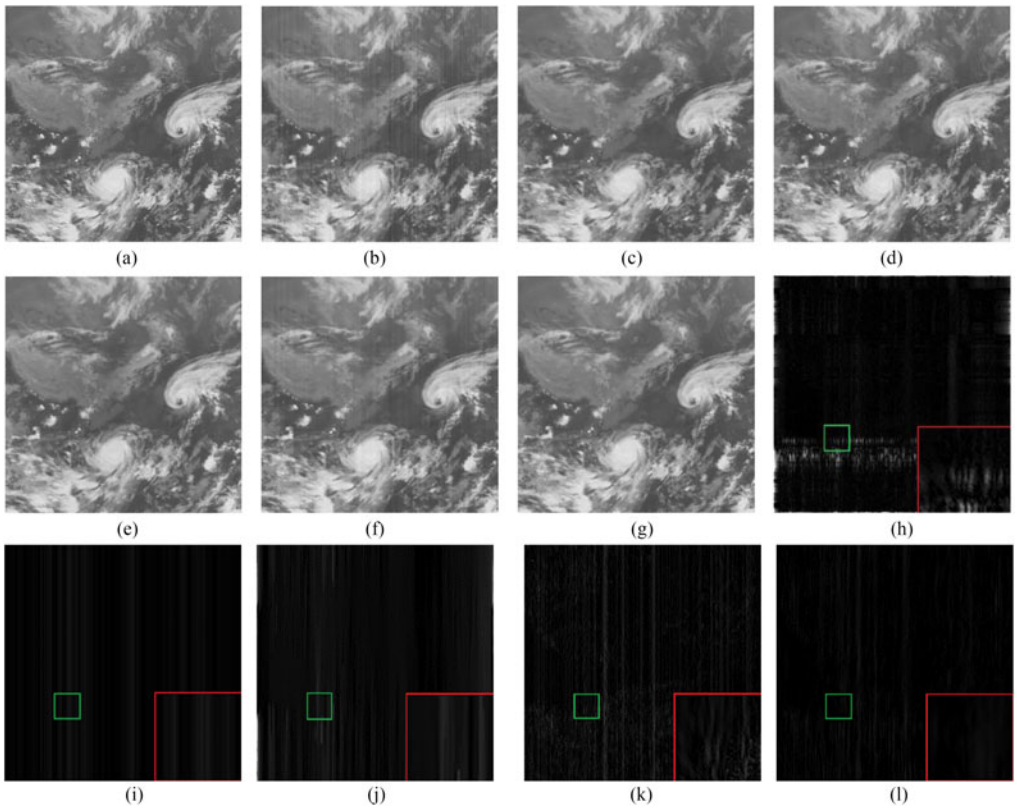


Fig. 9. Destriping results of one example from the proposed test dataset with noise level $\alpha = 35$ and (h)–(l) are the respective absolute values of the difference between the estimated images and original images. (a) original image (b) noisy (c) LRSID (d) WAFT (e) UTV (f) SNRCNN (g) ICSRN (h) LRSID (i) WAFT (j) UTV (k) SNRCNN (l) ICSRN.

Fig. 8, we can see that our method is able to obtain remarkable PSNR and SSIM gains compared to the state-of-the-art method [51].

The excellent performance of ICSRN demonstrate the strong capacity of ICSRN in dealing with noise that is similar with stripe noise. Meanwhile, ICSRN provides a new idea for rain-removal task since real-world rain streaks can be decomposed into horizontal and vertical components. Meanwhile, ICSRN can be utilized to deal with each component separately.

3. Experiments

3.1 Dataset and Similarity Metrics

We provide a new meteorological satellite infrared cloud image dataset for training and testing our network. The source data are downloaded from National Satellite Meteorological Center(NSMC).[1] Infrared band 1 of FY-2G meteorological satellite data is chosen as the source data, and the resolution of each pixel is 5 km. Our base training set consists of 200 infrared cloud images of size 575×575 and 50 infrared cloud images of size 575×575 are used for validation. We use 14 standard test images of size 475×475 for evaluation. We train ICSRN for destriping with three noise levels, i.e., $\alpha = 25, 30, 35$. We set the patch size equal to the receptive field as 35×35. Data augmentation is used and 203496 patches are cropped with stride 35 to train the model. We add aperiodic stripe noise by the method in [16], since aperiodic stripe noise is much more difficult to deal with than periodic stripe noise, which can be almost completely removed from remote sensing images by the method [33]. We use peak signal-to-noise ratio(PSNR) and structure similarity index measure(SSIM) [48] to measure the performance of several different destriping methods.

3.2 Implementation Details

We initialize the weights by the method in [48] which is a theoretically sound procedure for network utilizing rectified linear units (ReLU). We use stochastic gradient descent(SGD) with weight decay of 0.0001, a momentum of 0.9 and a mini-batch size of 64. We train all experiments over 60 epochs. Learning rate is initially set to 0.1 to obtain a fast convergence and then decreases by a factor of 10 every 30 epochs. Meanwhile, gradient clipping [49] is used to prevent gradient explosion.

We use the Caffe package [50] to implement our ICSRN.[2] All these experiments are carried out on an Intel i7 CPU 4.0GHz with 16GB RAM and Nvidia 1080 GPU. It takes about 8 hours to train our ICSRN.

3.3 Comparison With Several State-of-the-Art Methods

We compare the proposed ICSRN with several state-of-the-art destriping methods, including combined wavelet-Fourier filtering (WAFT) [8], the unidirectional TV model [27], low-rank-based single-image decomposition model (LRSID) [43] and Stripe Noise Removal Convolutional Neural Network (SNRCNN) [16]. We re-train SNRCNN on our infrared cloud image dataset. The implementation codes are downloaded from authors' websites or implemented by our own with comparable performance and the default parameter settings are used in our experiments.

Fig. 9 illustrates the visual results of different methods, and Table 2 lists the PSNR/SSIM results of different methods on the proposed test dataset. As one can see, our proposed ICSRN can achieve the best PSNR results than the competing methods and can yield the highest SSIM except UTV. Since the structure of cloud is so complex that it is hard for us to evaluate the visual results directly. Therefore, we show the absolute value of difference between the estimated images and original images within a certain range for better visualization. LRSID loses significant details which can be obviously observed in Fig. 9(h). As shown in Fig. 9(i), the result of WAFT contains obvious unprocessed or over-processed stripes. Even though UTV can thoroughly remove stripe noise, the

[1] http://satellite.nsmc.org.cn/portalsite/default.aspx
[2] The implementation and dataset are available at https://github.com/NUIST-xiaopengfei/ICSRN

TABLE 2
Average PSNR/SSIM of Different Methods on the Proposed Dataset

	PSNR/SSIM				
σ	LRSID	WAFT	UTV	SNRCNN	ICSRN
25	36.82/0.9749	45.01/0.9910	34.13/0.9913	34.13/0.9913	**45.03/0.9925**
30	36.64/0.9741	43.05/0.9741	34.03/**0.9913**	39.29/0.9706	**43.19**/0.9902
35	36.78/0.9751	41.59/0.9751	34.01/**0.9913**	37.93/0.9606	**41.63**/0.9876

TABLE 3
Average PSNR/SSIM of Different Methods on the Set of 18 Infrared Images [16]

	PSNR/SSIM	
σ	SNRCNN	ICSRN
25	44.40/0.9876	45.82/0.9971
30	42.64/0.9858	43.94/0.9942
35	40.65/0.9791	42.02/0.9869

(a) (b) (c) (d)

Fig. 10. Destriping results of one image [16] with noise level $\alpha = 35$ along with PSNR/SSIM. (a) original image (b) noisy (c) SNRCNN/44.42/0.9770 (d) ICSRN/**45.80**/**0.9867**.

lowest PSNR and the white area in Fig. 9(j) reveals the drawback of brightness distortion, which is mainly due to over-smoothing. As we can see from Fig. 9(k), SNRCNN remains significant stripe noise and loses many details. Since the limited capability of the three-layer convolution network is hard to produce pleasing results. Fig. 9(g) and (l) shows the destriping results of the proposed method which are visually pleasing, with few residual effects and local distortion.

We also compare the performance of our proposed method with SNRCNN on the set of 18 infrared images [16]. The model is trained on infrared cloud image dataset without re-training. From Fig. 10(c), the result of SNRCNN remains more residual noise than the proposed method which is consistent with the low PSNR and SSIM shown in Table 3. However, without re-training, our ICSRN produces more pleasing performance than SNRCNN, which demonstrates that our model is general.

3.4 Run time

Testing speed is another essential aspect for an image destriping method. Except for SNRCNN and ICSRN, other methods are not suited for parallel computation on GPU. In order to make a fair comparison between different methods, all methods are implemented on CPU. Table 4 shows the

TABLE 4
Average Run Time of Different Methods on Test Dataset With Different Noise Level

	Run Time(s)				
σ	LRSID	WAFT	UTV	SNRCNN	ICSRN
25	11.21	0.48	0.81	0.02	0.14
30	11.24	0.47	0.81	0.02	0.13
35	11.2	0.46	0.82	0.02	0.13

average run time of different methods on test dataset with different noise level. It can be seen that our ICSRN have a sub-second test speed, and it is faster than LRSID, WAFT, UTV. Although it is slower than SNRCNN, by considering image quality improvement, our ICSRN is more effective in image quality improvement. Therefore, ICSRN is suitable for real-time applications, among which the most important is weather forecasting.

4. Conclusion

In this paper, we proposed a deep convolutional neural network for meteorological satellite infrared cloud image destriping, where residual learning is adopted for speeding up the training process as well as boosting the destriping performance. Since the cloud are rich in details, and it is hard to completely recover details by simply using high-level features, the local-global combination structure model is designed for recovering more details by combining the representations of different layers. Moreover, we showed the feasibility to extend our ICSRN to remove vertical rain streaks from individual image, which provides a new idea for rain-removal task by decomposing real-world rain streaks to horizontal and vertical components. The results of extensive experiments demonstrated that the proposed method can produce favorable destriping performance quantitatively and qualitatively when dealing with meteorological satellite infrared cloud images and common infrared images. The proposed method still works when the resolution of infrared cloud images changes. In addition, the proposed method has promising run time, making it well-suited for real-time applications, such as Precipitation Nowcasting. In the near future, we will investigate our model for denoising of meteorological satellite infrared cloud images with real complex noise and extend our model for handling other image restoration tasks.

References

[1] T. Fiolleau and R. Roca, “An algorithm for the detection and tracking of tropical mesoscale convective systems using infrared images from geostationary satellite,” *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 4302–4315, Jul. 2013.
[2] H. Carfantan and J. Idier, “Statistical linear destriping of satellite-based pushbroom-type images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1860–1871, Apr. 2010.
[3] F. Aires, F. Marquisseau, C. Prigent, and G. Sèze, “A land and ocean microwave cloud classification algorithm derived from AMSU-A and -B, trained using MSG-SEVIRI infrared and visible observations,” *Mon Weather Rev.*, vol. 139, no. 139, pp. 2347–2366, 2015.
[4] J. J. Simpson, J. I. Gobat, and R. Frouin, “Improved destriping of GOES images using finite impulse response filters,” *Remote Sens. Environ.*, vol. 52, no. 1, pp. 15–35, Apr. 1995.
[5] J. Chen, Y. Shao, H. Guo, W. Wang, and B. Zhu, “Destriping CMODIS data by power filtering,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 2119–2124, Sep. 2003.
[6] J. Pan and C. Chang, “Destriping of Landsat MSS images by filtering techniques,” *Photogramm. Eng. Remote Sens.*, vol. 58, no. 10, pp. 1417–1423, Oct. 1992.
[7] J. Torres and S. Infante, “Wavelet analysis for elimination of striping noise in satellite images,” *Opt. Eng.*, vol. 40, no. 7, pp. 1309–1314, Jul. 2001.
[8] J. Chen, H. Xie, Y. Shao, and L. Yang, “Oblique striping removal in remote sensing imagery based on wavelet transform,” *Int. J. Remote Sens.*, vol. 27, no. 8, pp. 1717–1723, Apr. 2006.
[9] B. Munch, P. Trtik, F. Marone, and M. Stampanoni, “Stripe and ring artifact removal with combined wavelet-Fourier filtering,” *Opt. Exp.*, vol. 17, no. 10, pp. 8567–8591, May 2009.

[10] R. Pande-Chhetri and A. Abd-Elrahman, "De-striping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 5, pp. 620–636, Sep. 2011.
[11] J. Torres and S. Infante, "Wavelet analysis for elimination of striping noise in satellite images," *Opt. Eng.*, vol. 40, no. 7, pp. 1309–1314, Jul. 2001.
[12] B. Munch, P. Trtik, F. Marone, and M. Stampanoni, "Stripe and ring artifact removal with combined wavelet-Fourier filtering," *Opt. Exp.*, vol. 17, no. 10, pp. 8567–8591, May 2009.
[13] R. Pande-Chhetri and A. Abd-Elrahman, "De-striping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 5, pp. 620–636, Sep. 2011.
[14] A. D. Fischer, T. J. Thomas, R. A. Leathers, and T. V. Downes, "Stable scene-based non-uniformity correction coefficients for hyperspectral SWIR sensors," in *Proc. Int. Conf. IEEE Aerospace,* Mar. 2007, pp. 1–14.
[15] R. Bindschadler and H. Choi, "Characterizing and correcting hyperion detectors using ice-sheet images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1189–1193, Jun. 2003.
[16] X. Kuang, X. Sui, Q. Chen, and G. Gu, "Single infrared image stripe noise removal using deep convolutional networks," *IEEE Photon. J.*, vol. 9, no. 4, Aug. 2017 Art. no. 3900913.
[17] H. S. Jung, J. S. Won, M. H. Kang, and Y. W. Lee, "Detection and restoration of defective lines in the SPOT 4 SWIR band," *IEEE Trans. Image Process.*, vol. 19, no. 8, pp. 2143–2156, Aug. 2010.
[18] F. Tsai and W. W. Chen, "Striping noise detection and correction of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 4122–4131, Dec. 2008.
[19] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1852–1866, Jun. 2015.
[20] P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Stripe noise reduction in MODIS data by combining histogram matching with facet filter," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1844–1856, Jun. 2007.
[21] F. L. Gadallah, F. Csillag, and E. J. M. Smith, "Destriping multisensor imagery with moment matching," *Int. J. Remote Sens.*, vol. 21, no. 12, pp. 2505–2511, 2000.
[22] V. R. Algazi and G. E. Ford, "Radiometric equalization of nonperiodic striping in satellite data," *Comput. Graph. Image Process.*, vol. 16, no. 3, pp. 287–295, 1981.
[23] B. K. P. Horn and R. J. W. Oodham, "Destriping LANDSAT MSS images by histogram modification," *Comput. Graph. Image Process.*, vol. 10, no. 1, pp. 69–83, 1979.
[24] M. Wegener, "Destriping multiple sensor imagery by improved histogram matching," *Int. J. Remote Sens.*, vol. 11, no. 5, pp. 859–875, 1990.
[25] H. Shen, W. Jiang, H. Zhang, and L. Zhang, "A piece-wise approach to removing the nonlinear and irregular stripes in MODIS data," *Int. J. Remote Sens.*, vol. 35, no. 1, pp. 44–53, 2014.
[26] H. Shen and L. Zhang, "A MAP-based algorithm for destriping and inpainting of remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, May 2009.
[27] M. Bouali and S. Ladjal, "Toward optimal destriping of MODIS data using a unidirectional variational model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2924–2935, 2011.
[28] Y. Chang, L. Yan, H. Fang, and H. Liu, "Simultaneous destriping and denoising for remote sensing images with unidirectional total variation and sparse representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 6, pp. 1051–1055, Jun. 2014.
[29] G. Zhou, H. Fang, L. Yan, T. Zhang, and J. Hu, "Removal of stripe noise with spatially adaptive unidirectional total variation," *Optik*, vol. 125, no. 12, pp. 2756–2762, 2014.
[30] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.
[31] X. Liu, X. Lu, H. Shen, Q. Yuan, Y. Jiao, and L. Zhang, "Stripe noise separation and removal in remote sensing images by consideration of the global sparsity and local variational properties," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 3049–3060, May 2016.
[32] X. Liu, H. Shen, Q. Yuan, L. Zhang, and Q. Cheng, "A novel removal method for dense stripes in remote sensing images," in *Proc. ISPRS Annals Photogrammetry, Remote Sens. Spatial Inf. Sci.*, Prague, Czech Republic, 2016, pp. 57–61, vol. 3.
[33] X. Liu, H. Shen, Q. Yuan, X. Lu, and C. Zhou, "A universal destriping framework combining 1-D and 2-D variational optimization methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 808–822, Feb. 2018.
[34] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
[35] L. Xu, J. S. J. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2014, pp. 1790–1798, vol. 2.
[36] C. Dong, C. L. Chen, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.
[37] Y. Shi, K. Wang, C. Cheng, L. Xu, and L. Lin, "Structure-preserving image super-resolution via contextualized multi-task learning," *IEEE Trans. Mulrimedia.*, vol. 19, no. 12, pp. 2804–2815, Dec. 2017.
[38] Y. Liu, J. Pan, and Z. Su, "Deep blind image inpainting," arxiv: 1712.09078.
[39] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–409, Jun. 2013.
[40] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. Int. Conf. IEEE Comput. Vis. Pattern Recongnit.*, Boston, MA, USA, 2015, pp. 770–778.
[41] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, Lille, France, 2015, pp. 448–456.
[42] B. Lim, S. Son, H. Kim, H. Kim, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. Int. Conf. Comput. Vis. Pattern Recognit.*, Honolulu, Hawaii, USA, 2017, pp. 1131–1140.
[43] Y. Chang, L. Yan, T. Wu, and S. Zhong, "Remote sensing image stripe noise removal: From image decomposition perspective," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7018–7031, Dec. 2016.

[44] P. Liu and R. Fang, "Wide inference network for image denoising," arxiv:1707.05414.
[45] J. Kim, J. K. Lee, and K. M. Lee," Accurate image super-resolution using very deep convolutional networks," in *Proc. Int. Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, 2016, pp. 1646–1654.
[46] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Clearing the Skies: A deep network architecture for single-image rain removal," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2944–2956, Jun. 2017.
[47] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[48] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," arxiv:1502.01852.
[49] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training Recurrent Neural Networks," arxiv:1211.5063.
[50] Q. Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. Int. Conf. ACM Multimedia*, 2014, pp. 675–678.
[51] X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding, and J. Paisley, "Removing rain from single images via a deep detail network," in *Proc. Int. Conf. IEEE Comput. Vis. Pattern Recognit.*, Honolulu, HI, USA, 2017, pp. 1715–1723.