Infrared Aerothermal Nonuniform Correction via Deep Multiscale Residual Network

Yi Chang, *Student Member, IEEE*, Luxin Yan, *Member, IEEE*, Li Liu, Houzhang Fang, and Sheng Zhong

Abstract—In the infrared focal plane arrays imaging systems, the temperature-dependent nonuniformity effects severely degrade the image quality. In this letter, we propose a very deep convolutional neural network for unified infrared aerothermal nonuniform correction. Our network is built with the multiscale and residual training. The multiscale subnetworks utilize the multiscale property in the images, and the long–short-term residual learning contributes to the information propagation. Compared with the previous methods, the proposed method is more robust to various nonuniform artifacts and more efficient at processing time. Experimental results validate the superiority of our method for infrared nonuniform correction.

Index Terms—Convolutional neural network (CNN), infrared image, nonuniform correction.

I. Introduction

FOR the infrared imaging systems equipped on the high-speed aircraft, due to the temperatures fluctuations, the resulting images mainly contain two kinds of fixed pattern noise (FPN): a smooth nonuniform bias field that looks like bright and large spot and line pattern nonuniform stripe noise, as shown in Fig. 1(a). This aerothermal nonuniform effect severely influences the image quality for subsequent application. Therefore, it is necessary for us to remove these artifacts before the succeeding image interpretation processes are performed. In this letter, we mathematically formulate the degradation as follows:

$$Y = X + B + E + N \qquad (1)$$

where $Y \in \mathbb{R}^{R \times C}$ is the observed image, R and C stand for the number of the rows and columns, respectively, X is the clear image [Fig. 1(e)], B is the bias field [Fig. 1(c)], E is the stripe noise [Fig. 1(b)] (B and E are the nonuniform FPNs), and N is the random noise [Fig. 1(d)]. The goal of our work is to obtain the clear image X from the degraded image Y.
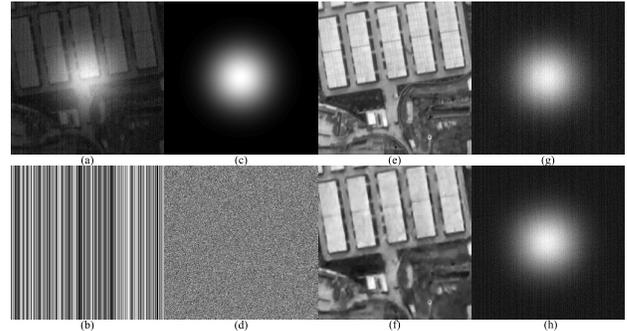
Fig. 1. Illustration of the main degradation factors in the IFPA detectors. (a) Degradation image Y. (b) Stripe noise E. (c) Bias field B. (d) Random noise N. (e) Ground truth image X. (f) Estimated image by the proposed method. (g) Whole error ($B + E + N$). (h) Estimated error by the proposed method.

The bias field is highly related to optical window temperature. The bias caused by thermal radiation looks like a bright and smooth spot, which is very similar to 2-D Gaussian distribution. For the bias field removal, we mainly introduce the optimization-based restoration methods [1]–[5]. Based on the correction model [6], Cao and Tisse [2] locally fitted the derivatives of correction model to the gradient components with a subsequent bilateral filter for refinement. Zheng and Gee [1] first proposed an image decomposition-based dual L_1 sparsity representation constraints for both the image and bias field component, respectively. Most of the optimization-based methods followed this framework. Later, Liu and Zhang [5] introduced the sparser L_p gradient regularization for the images with better performance.

The stripe noise is mainly caused by the nonuniform response of adjacent detectors. The stripe noise has significantly directional characteristic due to its imaging mechanism. For the FPN stripe removal, there exist various kinds of methods: reference-based approaches [7], filtering-based methods [8], [9], the scene-based optimization methods [10]–[12], and the learning-based method [13], [14]. Ratliff *et al.* [7] proposed an algebraic-based algorithm which assumed that each infrared focal plane arrays (IFPAs) detector output obeys an approximate linear irradiance voltage model. Cao *et al.* [9] took advantage of the directional characteristic of the FPN and proposed a 1-D guided filter for stripe noise removal in infrared images. The optimization-based sparsity methods have been popular in recent years. Vera *et al.* [10] proposed an isotropic total variation approach making use of an alternating minimization strategy for FPN removal. Furthermore, the L_0 gradient-based iterative adaptive nonuniform
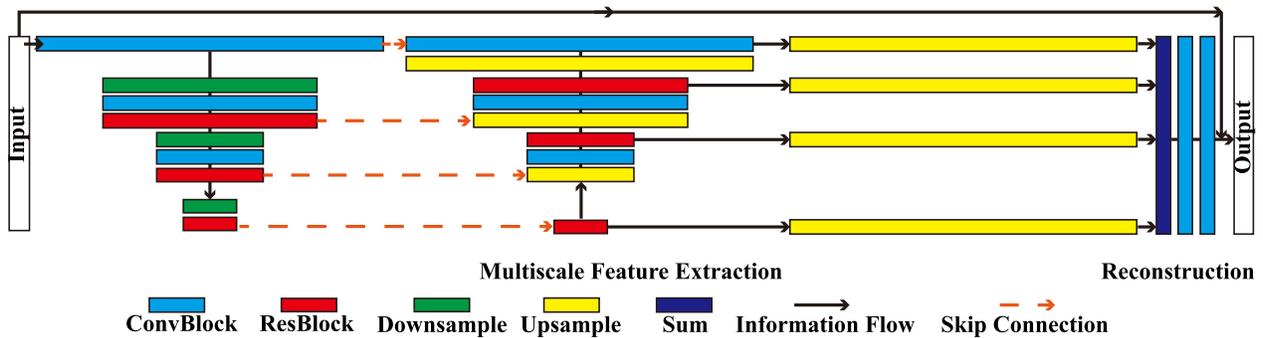
Fig. 2. Overview of the proposed DMRN for nonuniform correction in the IFPA. Our method contains two main parts: multiscale feature extraction and image reconstruction. In the first stage, we gradually downsample and upsample the features with skip connections. Then, we fuse the multiscale features for better representation of the IFPA images. In the second stage, we reconstruct the image from the output of the first stage. The residual learning is utilized both in short and long distance for better information propagation. The two stages are trained in an end-to-end manner.

correction method has been proposed [11], which assumed the gradient of the image is much sparser. Recently, Kuang *et al.* [13] presented a three-layer deep convolutional neural network (CNN) for single infrared image stripe noise removal.

Although numerous FPN correction methods have been proposed in the past decades, all of them are designed for one specific task only, such as the nonuniform bias field [1]–[5] or nonuniform stripe noise [7], [9]–[14]. Moreover, most of the previous works utilized the hand-crafted features, such as the gradient or the dictionary coefficients. These hand-crafted priors only explore the locally shallow feature information of the images while ignoring the global high-level feature of the images. To overcome these limitations, we propose a deep multiscale residual network (DMRN) for infrared nonuniform correction. We introduce a very deep CNN model for extracting the high-level contextual information of the images, in which the discriminative features are benefit to distinguish each component in (1) from each other. Our model does not rely on any predefined statistical assumption of the FPN or the image, which makes it very robust for arbitrary degradations. In addition, we utilize the multiscale information in the images via the CNN for better representation of both the large-scale edge and fine texture in the images. Overall, the contributions of this letter are as follows.

1) This letter proposes a very deep CNN model for unified infrared nonuniform bias field and stripe correction. Compared with the previous methods, the deep features are more representative for both the IFPA image and artifacts.

2) We introduce the long–short-term residual learning strategy, which significantly reduces the difficulty of training. In addition, the multiscale network could utilize the multiscale property in the images; meanwhile, it can obviously reduce the computational load and memory.

3) The proposed method outperforms the state-of-the-art (SOTA) methods by a large margin in terms of the speed, qualitative, and quantitative assessments. Moreover, our method is very robust to the random noise in IFPA images.

II. DEEP MULTISCALE RESIDUAL NETWORK

Here, we will give our unified DMRN for FPN removal. Although there are several deep learning-based methods for infrared stripe noise removal, the depth of them is too shallow, such as three layers in [13] and ten layers in [14]. On the contrary, we take the multiscale information of the image into consideration and enlarge the receptive field via the multiscale feature extraction module.

A. Multiscale Feature Extraction

It is well-known that the images contain different scales of information, such as the large-scale edges and the fine textures. On the one hand, the fine textures can be well represented by the shallow features with a relative small receptive field. The large-scale edges can be globally captured by the deep features with a relative larger receptive field. This motivates us to build a very deep network to globally capture both the local fine texture and global large edges. On the other hand, we argue that the multiscale structural information, which is benefit for image representation, need to be explicitly modeled by different scales of feature maps. This motivates us to construct a multiscale pyramid network to explicitly depict different scale structures.

The deep pyramid strategy has been extensively used in image segmentation [15], image restoration [16], due to its powerful representation ability. In this letter, we introduce a DMRN for infrared nonuniform FPN removal, as shown in Fig. 2. We first downsample the image gradually via the convolutional layer with the stride 2. In this letter, we extract four scale features from the original size 256×256 to 32×32. This part can be regarded as an encoder to extract the compact and multiscale representations. Then, we gradually upsample the image via the deconvolution layer with the scale 2. This part can be regarded as a decoder to reconstruct the signal. The skip connection is introduced, where all the features with the same size are concatenated together, to improve the information flow among different layers. We also introduce the residual blocks [17], which create short paths among neighborhood features, to alleviate gradient vanishing and train a very deep model.

To leverage the multiscale information to guide the nonuniform correction, a connectivity fusion layer is further
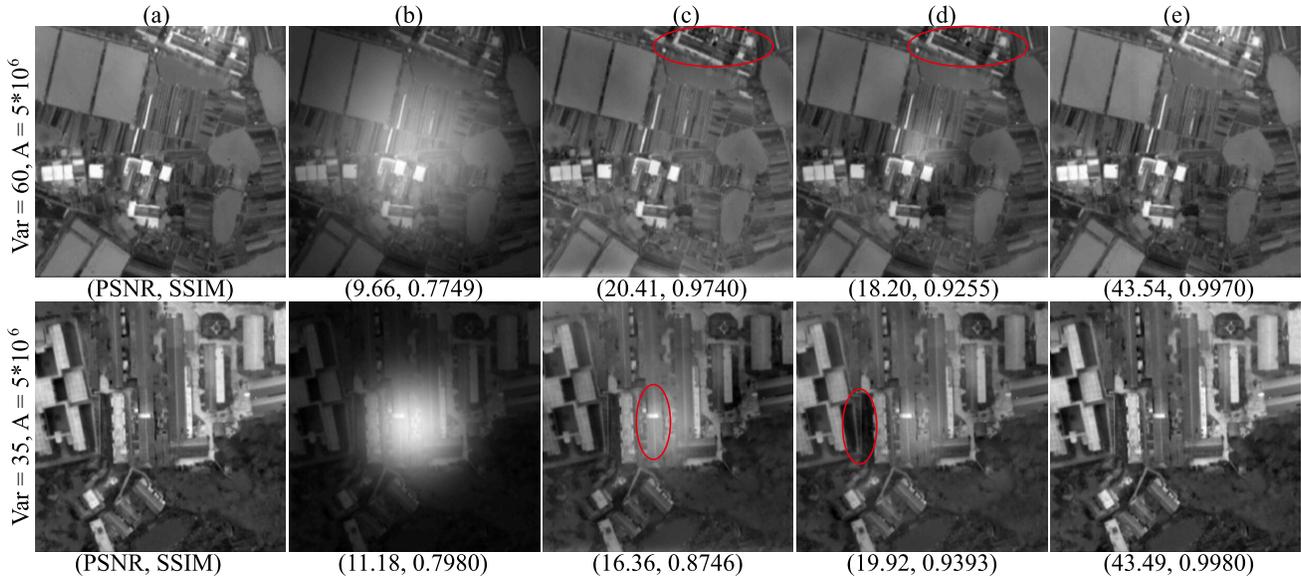
Fig. 3.   Simulated nonuniform bias field removal results under different degradation levels. The first and second rows show different bias fields. (Left to Right columns) (a) Original IR image. (b) Biased field noise image. Correction results by (c) GCF, (d) TVD, and (e) DMRN.

introduced, where the multiscale features from decoders are fused together. The intuitions behind this are twofolds. On the one hand, fusing the low-resolution features with the high-resolution features actually creates an information flow for better information propagation between low and high levels. On the other hand, this could better compensate low-level fine details to high-level large-scale features with joint and powerful representation for the images.

B. Reconstruction

In the second stage, the extracted multiscale features are the input of the reconstruction module which is used to obtain the clear image. Here, we use two convolutional blocks to achieve this goal. Note that we do not learn the mapping of the clear image directly. We introduce the residual learning strategy by adding a skip connection between the input and output. Thus, the network actually learns the whole error ($B + E + N$), it guarantee that sparser gradient of the residual errors is easier to propagate. To reconstruct the image, we introduce the L_2-based loss function

$$J = \frac{1}{2}||\mathcal{F}(Y;W) - (B + E + N)||^2 \qquad (2)$$

where W is the mapping parameters to be learned. We introduce both the resblock (short-term residual) and residual learning (long-term residual) to increase the representation ability of the network. The residual learning greatly improves the depth of the network and avoids gradient vanishing issue.

C. Training Details

The MatConvnet toolbox [18] is employed to train the model. The training code and IR data sets of our DMRN have been released at the homepage of the author.[1] We initialize the convolutional filters with the Xavier method [17]. The learning rate starts from 0.0005 and is divided by 2 after each 20 epoch.

[1]http://www.escience.cn/people/changyi/index.html

TABLE I
AVERAGE QUANTITATIVE RESULTS OF DIFFERENT METHODS UNDER SEVERAL BIAS FIELD LEVELS ON 20 TEST IR IMAGES

	Index	Bias Field	GCF	TVD	DMRN	DMRN+
Var = 60	PSNR	9.84	18.25	19.27	35.97	37.47
$A = 5*10^5$	SSIM	0.761	0.9201	0.9301	0.9924	0.9945
Var = 35	PSNR	11.16	17.36	19.27	35.24	36.78
$A = 5*10^6$	SSIM	0.8029	0.8954	0.9301	0.9944	0.9957

The momentum and decay are fixed as 0.9 and 0, respectively. Adam solver [19] is used as optimization algorithm with a minibatch size of 48. We train the model with 100 epoches. We randomly choose 10 000 samples from the Place 2 data set with size 256×256 for pretraining. Then, we fine-tune the pretrained models on the collected 1500 midwave IR images.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Experimental Setting

For the bias field removal, we compare with the gradient components-based filtering (GCF) method [2] and TV-based decomposition (TVD) methods [3]. For the stripe noise removal, we compare with the low-rank single image decomposition (LRSID) [12] and deep-learning-based stripe nonuniformity correction [14]. We use the codes provided by the authors and fine-tune the hyperparameters by default to achieve the best performance. The PSNR and SSIM are employed for the quantitative index. The visual correction and cross profile are used as the qualitative assessment. Due to the page limitation, more results and analysis are placed in the supplementary.

B. Comparison With STOA

1) Bias Field Correction: We compare DMRN with SOTA for nonuniform bias field correction in the IR images. In Fig. 3, the GCF and TVD always introduce unexpected artifacts (marked by the red ellipse). The results of the proposed method are almost the same as the original image.
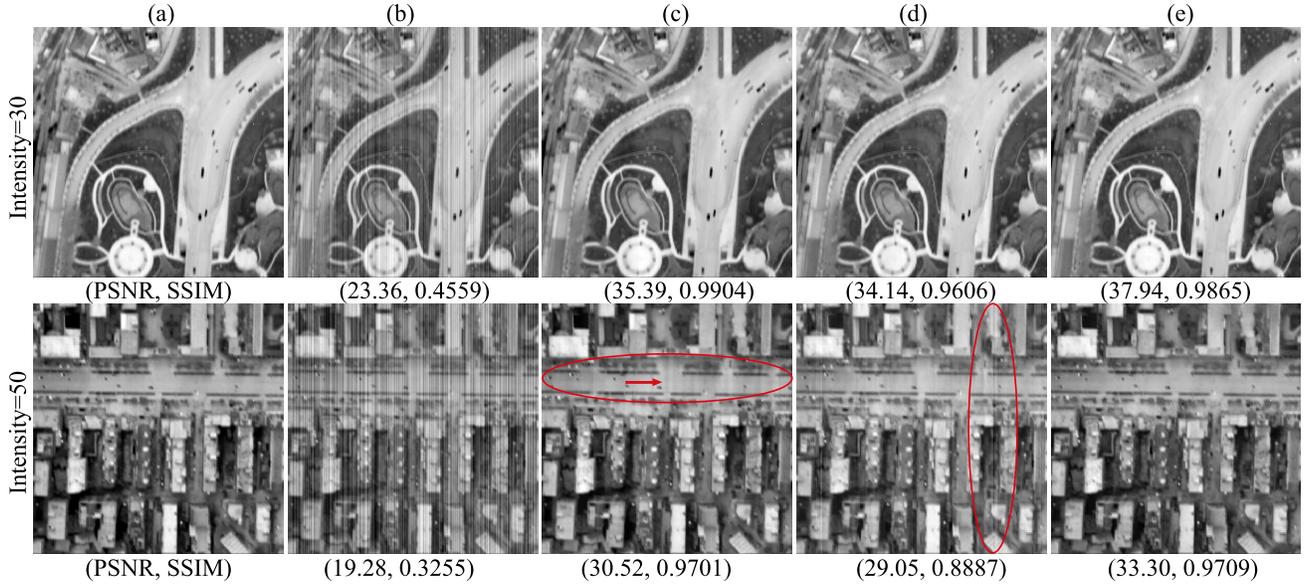
Fig. 4. Simulated nonuniform stripe noise removal results under different degradation levels. The first and second row show the light and heavy stripe noise, respectively. From the left to right columns (a) Original IR image. (b) Striped image. Correction results by (c) LRSID, (d) DLS-NUC, and (e) DMRN.

TABLE II
AVERAGE QUANTITATIVE RESULTS OF DIFFERENT METHODS UNDER SEVERAL STRIPE LEVELS ON 20 TEST IR IMAGES

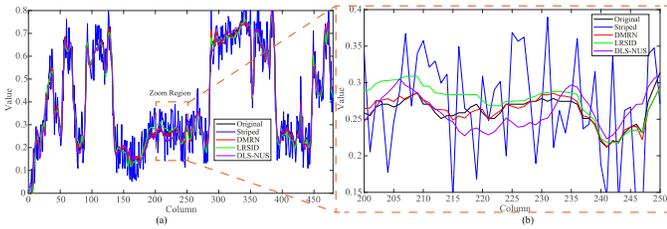|	Index	Stripe Noise	LRSID	DLS-NUS	DMRN
$S \subset \{-30, 30\}$	PSNR	23.43	35.13	34.31	37.67
	SSIM	0.5399	0.9864	0.9568	0.9844
$S \subset \{-50, 50\}$	PSNR	19.02	31.78	30.09	35.17
	SSIM	0.3486	0.9771	0.8955	0.9775

Fig. 5. Cross profile analysis for the nonuniform correction. Horizontal axis: column number of the IR image. Vertical axis: intensity value of the IR image. (a) Cross profile of a certain row. (b) Zoomed-in view of (a).

The DMRN+ in Table I denotes the fine-tuned DMRN model. Under different conditions, the proposed DMRN consistently outperforms the SOTA bias field correction methods.

2) Stripe Noise Removal: We compare DMRN with SOTA methods for nonuniform stripe noise correction in the IR images. Here, we test them on different stripe noise levels and larger image size 480×480, as shown in Fig. 4. We can observe there exist obvious residual stripe noises in the correction results of LRSID and DLS-NUC as marked by the red ellipse, while in Fig. 4(e), the corrections results by DMRN are more visually pleasing. The quantitative comparison results are given in Table II. The DMRN is robust to different noise levels and image sizes and consistently outperforms the SOTA IR stripe correction methods.
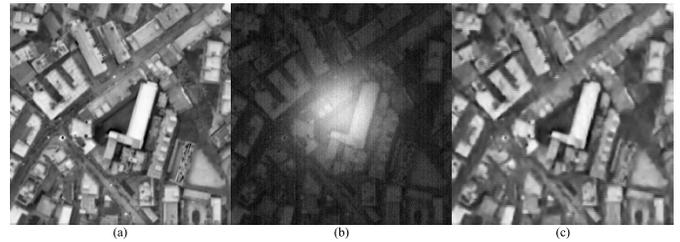

Fig. 6. Joint nonuniform correction. (a) Original IR image. (b) Degraded image. (c) Correction results of DMRN.

C. Discussion

1) Cross Profile Analysis: In this section, we analyze the cross profile of the correction result, as shown in Fig. 5. We take the stripe noise correction (first row in Fig. 4) as an example. From the zoomed-in view of Fig. 5(b), we can clearly observe that the correction result of DMRN (red curve) is much more closer to the original ground truth (black curve). Moreover, compared with the green curve of the striped image, we can infer from the smoothed red curve that the nonuniform stripe noise has been satisfactorily removed.

2) Joint Nonuniform Correction: We further demonstrate that our model is not only suitable for one specific nonuniform artifacts in IR but also works well for arbitrary mixed nonuniform artifacts, due to the universal approximation theory of the deep neural network [20]. We test the DMRN on the mixed stripe and bias field, which are commonly seen nonuniform artifacts in IR image, along with the Gaussian random noise, as shown in Fig. 6. As far as we know, there is no way to uniformly correct all of these artifacts. The DMRN [Fig. 6(c)] can well remove the artifacts and preserve the edge structure.

3) Benefit For Recognition: To validate the effectiveness of the proposed method for subsequent application, we employ Google Vision API[2] on the images before and after correction to perform the scene recognition. Here, we choose a natural

[2]https://cloud.google.com/vision/

Fig. 7. Effectiveness of DMRN for the recognition. (a) Google Vision API understanding result of the original image. (b) Bias field severely degrades the object recognition accuracy. (c) DMRN greatly improves the scene recognition accuracy.
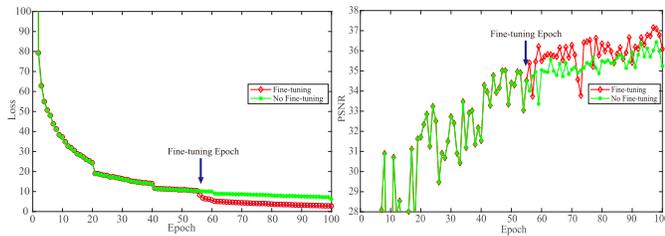

Fig. 8. Effectiveness of the fine-tuning. After the fine-tuning, the training loss drops drastically, while the PSNR value increase rapidly.

image as an example, since the API is mainly trained on natural images. As shown in Fig. 7, the bias field on the image puzzles the API with wrong labels such as light and sky. The recognition result of the correction by DMRN is almost the same as that of the original, such as the roof and house.

4) Effectiveness of the Fine-Tuning: To alleviate the lack of IR image, we first train our model on the RGB image. Then, we fine-tune our model on the collected IR images. We show this transfer learning strategy is very effective for nonuniform correction. In Fig. 8, the training loss and PSNR value of each epoch are shown before and after the fine-tuning. It can be seen that the training loss drops suddenly at the epoch we fine-tune the model, and the loss of the fine-tuned model is significantly lower than the no fine-tuned model. On the contrary, the PSNR value increases after we fine-tune the model.

IV. CONCLUSION

In this letter, we propose a DMRN for infrared image nonuniform correction. We utilize the deep feature that is more representative and robust to various nonuniform artifacts in IR image. The multiscale information is used in our network to better represent the IR images. The residual learning and fine-tuning strategy are introduced for better training the network. Experimental results show that the proposed DMRN is superior to competing deep and nondeep methods by a large margin.

REFERENCES

[1] Y. Zheng and J. C. Gee, "Estimation of image bias field with sparsity constraints," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 255–262.

[2] Y. Cao and C.-L. Tisse, "Single-image-based solution for optics temperature-dependent nonuniformity correction in an uncooled long-wave infrared camera," *Opt. Lett.*, vol. 39, no. 3, pp. 646–648, 2014.

[3] L. Liu, L. Yan, H. Zhao, X. Dai, and T. Zhang, "Correction of aeroheating-induced intensity nonuniformity in infrared images," *Infr. Phys. Technol.*, vol. 76, pp. 235–241, May 2016.

[4] L. Li, L. Yan, N. Sang, C. Gao, and J. Hu, "Aero-thermal radiation correction via multi-scale bias field estimation," in *Proc. IEEE Conf. Asian Conf. Pattern Recognit.*, Nov. 2016, pp. 246–250.

[5] L. Liu and T. Zhang, "Intensity non-uniformity correction of aerothermal images via ℓ_p-regularized minimization," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 33, no. 11, pp. 2206–2212, 2016.

[6] U. Vovk, F. Pernus, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI," *IEEE Trans. Med. Imag.*, vol. 26, no. 3, pp. 405–421, Mar. 2007.

[7] B. M. Ratliff, M. M. Hayat, and R. C. Hardie, "An algebraic algorithm for nonuniformity correction in focal-plane arrays," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 19, no. 9, pp. 1737–1747, 2002.

[8] X. Sui, Q. Chen, and G. Gu, "Adaptive grayscale adjustment-based stripe noise removal method of single image," *Infr. Phys. Technol.*, vol. 60, pp. 121–128, Sep. 2013.

[9] Y. Cao, M. Y. Yang, and C.-L. Tisse, "Effective strip noise removal for low-textured infrared images based on 1-D guided filtering," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 12, pp. 2176–2188, Dec. 2016.

[10] E. Vera, P. Meza, and S. Torres, "Total variation approach for adaptive nonuniformity correction in focal-plane arrays," *Opt. Lett.*, vol. 36, no. 2, pp. 172–174, 2011.

[11] J. Zhao *et al.*, "Single image stripe nonuniformity correction with gradient-constrained optimization model for infrared focal plane arrays," *Opt. Commun.*, vol. 296, pp. 47–52, Jun. 2013.

[12] 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.

[13] 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.

[14] Z. He, Y. Cao, Y. Dong, J. Yang, Y. Cao, and C.-L. Tisse, "Single-image-based nonuniformity correction of uncooled long-wave infrared detectors: A deep-learning approach," *Appl. Opt.*, vol. 57, no. 18, pp. D155–D164, 2018.

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

[16] W. S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep laplacian pyramid networks for fast and accurate superresolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 5835–5843.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

[18] A. Vedaldi and K. Lenc, "MatConvNet: Convolutional neural networks for MATLAB," in *Proc. ACM Int. Conf. Multimedia*, 2014, pp. 689–692.

[19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, 2015, pp. 1–15.

[20] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.