



Single-image-based nonuniformity correction of uncooled long-wave infrared detectors: a deep-learning approach

ZEWI HE,^{1,2} YANPENG CAO,^{1,2,*} YAFEI DONG,^{1,2} JIANGXIN YANG,^{1,2} YANLONG CAO,^{1,2}
AND CHRISTEL-LÖIC TISSE³

¹State Key Laboratory of Fluid Power and Mechatronics Systems, School of Mechanical Engineering, Zhejiang University, Hangzhou 310027, China

²Key Laboratory of Advanced Manufacturing Technology of Zhejiang Province, School of Mechanical Engineering, Zhejiang University, Hangzhou 310027, China

³ULIS, ZI Les Iles Cordees BP27, 38113 Veurey-Voroize, France

*Corresponding author: caoy@zju.edu.cn

Received 1 February 2018; revised 4 April 2018; accepted 5 April 2018; posted 6 April 2018 (Doc. ID 320251); published 4 May 2018

Fixed-pattern noise (FPN), which is caused by the nonuniform opto-electronic responses of microbolometer focal-plane-array (FPA) optoelectronics, imposes a challenging problem in infrared imaging systems. In this paper, we successfully demonstrate that a better single-image-based non-uniformity correction (NUC) operator can be directly learned from a large number of simulated training images instead of being handcrafted as before. Our proposed training scheme, which is based on convolutional neural networks (CNNs) and a column FPN simulation module, gives rise to a powerful technique to reconstruct the noise-free infrared image from its corresponding noisy observation. Specifically, a comprehensive column FPN model is utilized to depict the nonlinear characteristics of column amplifiers in the readout circuit of FPA. A large number of high-fidelity training images are simulated based on this model and the end-to-end residual deep network is capable of learning the intrinsic difference between undesirable FPN and original image details. Therefore, column FPN can be accurately estimated and further subtracted from the raw infrared images to obtain NUC results. Comparative results with state-of-the-art single-image-based NUC methods, using real-captured noisy infrared images, demonstrate that our proposed deep-learning-based approach delivers better performances of FPN removal, detail preservation, and artifact suppression. © 2018 Optical Society of America

OCIS codes: (100.2550) Focal-plane-array image processors; (100.2980) Image enhancement; (110.4280) Noise in imaging systems; (110.3080) Infrared imaging; (100.4996) Pattern recognition, neural networks; (100.3010) Image reconstruction techniques.

<https://doi.org/10.1364/AO.57.00D155>

1. INTRODUCTION

Microbolometer focal plane arrays (FPAs) are the key component of modern uncooled infrared imaging systems. The operating principle of microbolometer detectors is that the temperature variation caused by the absorption of infrared radiation leads to a change in electrical resistance of the bolometer material. A two-dimensional (2D) thermal image is then generated by translating the changes in resistance of each microbolometer detector in FPA into a time-accumulated electrical signal [1]. However, the radiometric accuracy of FPA is adversely affected by the responsive nonuniformity of the detectors in the array. Such nonuniformity typically manifests itself as undesirable fixed-pattern noise (FPN) in the raw infrared data, which is particularly severe in uncooled long-wave infrared imaging systems. The existence of complex FPN significantly decreases performances of infrared applications

such as energy assessment, thermal analysis, and medical diagnosis.

Figure 1(a) illustrates the structures of FPA, which include a detector array, column-parallel blind bolometers, column-parallel accumulators, and column-parallel analog-to-digital converters (ADCs) [1,3,4]. The column-parallel blind bolometers provide references for extracting small thermal signals from severe background noise. The photo-induced electric signals are accumulated in the integrators, then read out through the corresponding column ADCs under the control of the sequential circuit [3]. It is noted that characteristics of reference bolometers and ADCs in different columns are not perfectly uniform, and such nonuniformity of the readout circuit will introduce vertical noise strips as shown in Fig. 1(b). Effective removal of this common type of FPN is a critical step to improve the radiometric accuracy of captured infrared data as shown in Fig. 1(c).

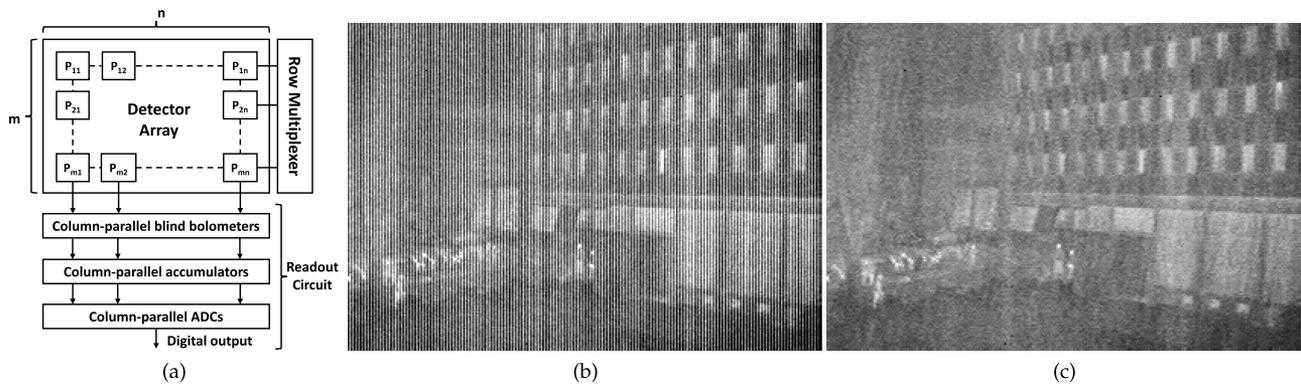


Fig. 1. Column-parallel reference bolometers and ADCs in infrared FPA have different characteristics and will cause obvious column FPN. (a) Block diagram of uncooled long-wave infrared FPA. (b) A raw infrared image with severe column FPN. (c) NUC result of our method. It is noted that without proper noise compensation, objects in a raw image are difficult to recognize. The raw image is available under the Creative Commons Attribution (CC-BY) license [2].

Accordingly, many strip nonuniformity correction (NUC) algorithms have been proposed to suppress column FPN while preserving thermal details [3–10]. However, the above-mentioned NUC methods exhibit a number of limitations as follows: (1) infrared images typically contain a substantial amount of low-magnitude signals, therefore it is difficult to remove strong strip noise without blurring thermal details; (2) the intensity adjustment operation (e.g., Midway Histogram Equalization [2,5]) usually introduces some obvious artifacts into the NUC results; (3) most of these methods contain many parameters that require fine-tuning to achieve satisfactory NUC performance for a particular input image.

In recent years, the deep-learning technique has been successfully applied to solve image restoration problems (e.g., super-resolution [11], deblurring [12], and in-painting [13]) and has achieved breakthrough improvements. In this paper, we made the attempt to learn a better performing strip NUC operator from a large number of simulated training images. Compared with existing handcrafted NUC methods, our single-image-based method has many desirable advantages including artifact-free noise reduction, better detail preservation, and no user-specific parameters. A comprehensive column FPN model, which characterizes column-wise nonuniformity in the

readout circuit of FPA, is utilized to simulate a large number of training image pairs (infrared images with and without simulated column FPNs). The training images are then fed into our end-to-end deep network architecture to reconstruct the residual information between noisy and noise-free images. Instead of learning the mapping relationship between noisy and noise-free pairs [14], our network aims to reconstruct the residual information (i.e., column FPN) first and then subtracts it from the raw image to generate the NUC output. This residual learning strategy is proven effective for achieving better FPN removal performance. Moreover, we employ pooling and deconvolution layers in our architecture to enable both a large receptive field and computation cost reduction. The workflow of our proposed single-image deep-learning-based strip NUC method (DLS-NUC) is illustrated in Fig. 2. The contributions of our work include:

- We present a unified framework, which combines a column FPN simulation module with convolutional neural networks (CNNs), to train the optimal model for column FPN correction of infrared images. To the best of our knowledge, it is the first work revealing that NUC model trained using simulated data can be successfully employed to handle real-captured infrared images with column FPN.

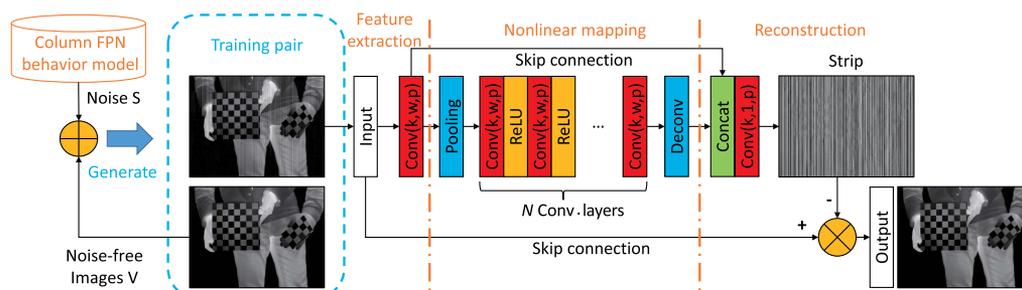


Fig. 2. Workflow of our proposed DLS-NUC method. A column FPN model is adopted to generate substantial training pairs. Our DLS-NUC method embeds end-to-end CNNs to learn the intrinsic difference between undesirable FPN and original image details. The estimated column FPN is subtracted from the noisy image to obtain NUC results. Conv, Deconv, and Concat denote convolutional, deconvolution, and concatenation layers, respectively. Conv(k, w, p) indicates that this convolutional operation takes w kernels of size $k \times k$ on the images/features with padding p . Please note that our residual network attempts to reconstruct FPN signals instead of the noise-free image. Such a residual learning strategy leads to better performances of noise reduction and detail preservation. All images are normalized to the 0–1 range for visualization.

- We design a residual network architecture with a large receptive field to accurately distinguish between undesirable FPNs and the original thermal details. Compared with state-of-the-state single-image-based NUC solutions, our DLS-NUC achieves artifact-free noise reduction, better detail—preservation, and has no user-specific parameters.

2. RELATED WORK

The traditional NUC methods are typically divided into two main categories: calibration-based and scene-based. The calibration-based techniques [15,16], which are based on uniform temperature references (e.g., shutter or blackbody), are widely applied in some commercial infrared cameras. However, they periodically freeze image capturing for a few seconds and are not suitable for dynamic infrared application [17]. To overcome this limitation, many scene-based NUC techniques, based on different working principles such as statistic analysis [18,19], algebraic computation [20–22], temporal registration [23], or Kalman filtering [24,25], have been proposed. A noticeable limitation of many scene-based methods is that they typically require multiple frames and are not applicable to single-frame images with FPNs. Moreover, “ghosting” artifacts will appear in the current frame if the image sequence does not contain enough scene motions. It is also worth mentioning that the above methods are developed for slowly drifting spatial nonuniformity and cannot achieve satisfactory results for column FPN removal [6,7,26].

Münch *et al.* [27] developed a de-stripping method based on wavelet decomposition and Fourier transform (WD-FT). Strip noise is filtered out in the Fourier domain of wavelet vertical components. Qian *et al.* [7] applied a threshold to distinguish the difference between image edges and column FPN. In the extracted non-edge area, the optimal NUC parameters are computed to minimize the energy of horizontal gradients. However, this threshold-based method will falsely remove weak thermal details while preserving strong strip noise in low-textured infrared images [28]. In [29], the NUC problem is translated into a gradient-constrained optimization problem, which aims to compute the optimal image in which the energy of horizontal gradients is as small as possible while the vertical gradients are similar to the ones of the original image. Tendero *et al.* [2,5] assumed that the difference between two adjacent columns is statistically small and then applied Midway Histogram Equalization (MHE) to adjust intensities of pixels within a column. MHE-based NUC does not need to separate column FPN from thermal details, and can therefore effectively remove FPN without blurring details. However, it is prone to generating false artifacts. In [9], a local linear model is used to depict the relationship between column FPN and thermal radiation. A 1D row guided filter is applied to perform edge-preserving image smoothing in the horizontal direction, while a 1D column guided filter is applied to separate column FPN from other high-frequency signals (edges and textures). Although this method can achieve a good balance between noise removal and detail preservation, still some vertical structures are falsely removed from its NUC results.

Stripe noise removal convolutional neural network (SNRCNN) presented by Kuang *et al.* [14] is by far the most

relevant method found in the literature to our research work. This deep-learning-based method involved three convolutional layers and performed well on images with simulated strip noise. However, SNRCNN was trained using visible images only and did not consider the natural differences between infrared and visible images (e.g., infrared images are low textured, and significant column FPN is visually more obvious than the weak thermal details), its performance drops significantly when applied to real-captured infrared images with column FPN. In comparison, our DLS-NUC model employs a comprehensive column FPN model to simulate high-fidelity training data and can successfully handle real-captured infrared images. Moreover, SNRCNN attempts to directly learn the complex mapping relationship between noisy and noise-free images, which is difficult to train, while our proposed residual deep network model aims to reconstruct the residual information (column FPN). The latter is proven effective for achieving higher image restoration accuracy. Finally, three-layer SNRCNN only considers information within a small receptive field, which is not sufficient to restore original image details. On the contrary, pooling and deconvolution layers are added into our DLS-NUC architecture to enable both a large receptive field and computation cost reduction.

3. METHODOLOGY

Our proposed single-image-based DLS-NUC method consists of a column FPN simulation module and a residual deep network. Its workflow is illustrated in Fig. 2. A comprehensive column FPN model is derived to characterize spatial nonuniformity in the readout circuit of FPA. Based on this model, a large number of high-fidelity training images are simulated and fed to a residual network model with a large receptive field to accurately distinguish between undesirable FPN and the original thermal details. The estimated residual information (i.e., column FPN) is subtracted from the raw infrared image to obtain NUC results.

A. Training Data Simulation

It is important to simulate a large number of high-fidelity training image pairs (infrared image patches with and without simulated FPNs) to train the optimal deep network model for column FPN correction of real-captured infrared images. Previously, a number of noise models have been presented to characterize the relationship between column FPN and thermal radiation. A simple offset model is used to simulate column FPN in [14]. In [3,6], a linear correction model is utilized to normalize column-wise pixel outputs for strip NUC. Cao and Li [8] designed a series of thermal calibration experiments and revealed that thermal responses and strip noise terms of pixels within a column can be characterized by a quadratic curve model. In this paper, we employ a comprehensive column FPN model that incorporates the above-mentioned column FPN models.

Let $V(i, j)$ be the thermal response of a detector (i, j) on the FPA, and then the column FPN $S(i, j)$ caused by spatial nonuniformity of column-parallel readout circuits can be characterized by a nonlinear function G as

$$S(i, j) = G(V(i, j)), \quad (1)$$

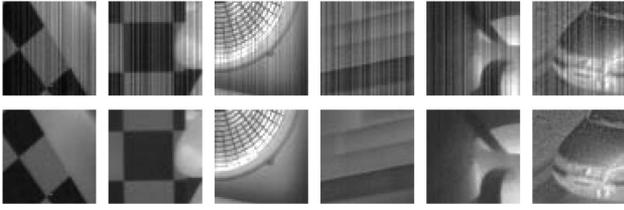


Fig. 3. Some training pairs (54×54 resolution). Top: image with simulated column FPN. Bottom: corresponding noise-free image.

where we make use of a polynomial model of degree M to approximate the nonlinear function G , and the FPN component of pixel (i, j) becomes

$$S(i, j) = a_j^M V^M(i, j) + a_j^{M-1} V^{M-1}(i, j) + \cdots + a_j^0, \quad (2)$$

where $a_j^M, a_j^{M-1}, \dots, a_j^0$ are the column-fixed polynomial coefficients. It is worth mentioning that our column FPN model adaptively incorporates other noise models mentioned above [3,6,8,14,30]. Specifically, when $a_j^M = a_j^{M-1} = \cdots = a_j^3 = 0$, Eq. (2) becomes the second-order polynomial model adopted in [8]. When $a_j^M = a_j^{M-1} = \cdots = a_j^2 = 0$, it becomes a linear model [3,6]. When $a_j^M = a_j^{M-1} = \cdots = a_j^1 = 0$, it presents a simple offset model used in [14,30]. Through adjusting the values of polynomial coefficients, our model can simulate different types of column FPNs. In our implementation, $a_j^M, a_j^{M-1}, \dots, a_j^0$ are randomly assigned to values in the interval between -0.1 and 0.1 . A critical parameter of this column FPN model is the degree number M . The comparative results using FPN models of different degree numbers are provided and discussed in Section 3.C. The experimental results demonstrate that setting $M = 3$ provides a good balance between complexity and performance.

Based on the derived column FPN model, we generate a large number of training image patches. We make use of a commercial uncooled long-wave infrared camera (Xenics Gobi-640-GigE) to capture 100 high-resolution (640×480) infrared images for training and another 20 images for testing. Then the images are cropped into sub-images with a size of 54×54 , and data augmentation methods (left-to-right flip, rotation, and scale) are applied to expand the number of training samples. For each noise-free image patch V , we add simulated column FPN S computed based on Eq. (2) to generate its noisy version $I = V + S$. Please note that each simulated noisy image patch has a unique pattern of column FPNs. In total, 192,384 pairs of image patches with and without simulated column FPNs are collected to build up the training dataset for our residual deep network. Some sample training image pairs are shown in Fig. 3. Using such a large training dataset covering different FPN patterns to train a deep network model will significantly boost the robustness of image restoration and enable its application on real-captured infrared images.

B. Deep Network Architecture

As illustrated in Fig. 2, our residual network consists of three consecutive operations including feature extraction, nonlinear mapping, and image reconstruction. The feature extraction operation takes noisy image patches as input and then outputs a

high-dimensional matrix which contains a set of feature maps. The nonlinear mapping operation transfers the high-dimensional matrix to another matrix through nonlinear functions. In this step, a pooling layer is employed to simultaneously enlarge the receptive field and reduce the computational cost. To increase the nonlinearity, N convolutional layers are followed by $N - 1$ rectified linear unit (ReLU) activation functions. Finally, the reconstruction operation utilizes a concatenation layer to fuse the feature maps before and after pooling layers to alleviate information loss during the downsampling. The fused feature maps are aggregated to generate the final output.

In each training image pair (I, V) , I denotes the input noisy image and V denotes the noise-free reference image, as illustrated in Fig. 2, the feature extraction operation employs a convolutional layer to extract high-dimensional features as

$$F_0 = W_0 * I + B_0, \quad (3)$$

where W_0 and B_0 indicate the filtering weights and biases, respectively, and $*$ denotes the convolution operation. Here, W_0 corresponds to w filtering kernels of size $k \times k$, and B_0 corresponds to w biases. Intuitively, W_0 applies w convolutions on the image, and each convolution has a kernel size $k \times k$ with padding p . Each convolution result is added with corresponding bias. Note that the output F_0 consists of w feature maps. In our implementation, we set $p = \frac{1}{2}(k - 1)$ to avoid size reduction of feature maps. All convolutional layers in our DLS-NUC share the same setting of p . In addition, the stride of convolution operation is set to 1.

The output of feature extraction operation is then fed to a max pooling layer. The downsampling operation resizes extracted feature maps with a resizing ratio r as

$$F'_0 = \downarrow r(F_0), \quad (4)$$

where F'_0 denotes the resized feature maps after the pooling layer, and $\downarrow r$ means a downsampling operation with a scale factor r . There are two main advantages in employing this downsampling operation. First, it accelerates the restoration process, since the processing cost of a smaller-size image in subsequent convolutional layers will be reduced by a factor of r^2 . Second, it enlarges the receptive fields of the subsequent layers by r^2 times and incorporates more neighborhood information to achieve high-accuracy restoration. We will evaluate the impacts of this pooling layer design later in Section 3.C.

N convolutional layers are embedded in the architecture to perform nonlinear mapping of feature maps as

$$F'_i = \max(0, W_i * F'_{i-1} + B_i), \quad i \in \{1, 2, \dots, N-1\}, \quad (5)$$

$$F'_N = W_N * F'_{N-1} + B_N, \quad (6)$$

where F'_i denotes the output feature maps of i -th convolution, and W_i and B_i indicate the filtering weights and biases of the i th convolution, respectively. We apply a ReLU activation function (i.e., $\max(0, x)$) [31] after each convolution operation to increase the nonlinearity of the decision function. Then we utilize a deconvolution layer to upsample the feature maps with the same factor r as

$$F_N = \uparrow r(F'_N), \quad (7)$$

where F_N denotes the resized feature maps after the deconvolution layer, and $\uparrow r$ means an upsampling operation with a scale

factor r . Due to the deployment of the pooling layer, some detailed information is unavoidably lost in the feature map F_N . As an effective remedy, we combine the feature maps from low-resolution space F_N with the feature maps from original space F_0 in the reconstruction step as

$$F_{\text{fusion}} = [F_0; F_N], \quad (8)$$

where the concatenated output F_{fusion} contains $2 * w$ feature maps. Finally, we apply a convolution layer $\text{Conv}(k, 1, p)$ with $w = 1$ to reconstruct the residual information between I and V as

$$S = W_c * F_{\text{fusion}} + B_c, \quad (9)$$

where W_c and B_c indicate the filtering weights and biases of the reconstruction convolution, respectively. This residual learning strategy has been successfully applied to reduce the convergence time and improve the performance [32]. More evaluation results are presented in Section 3.C. The final predicted noise-free image P is calculated by subtracting the reconstructed column FPN S from the input noisy image I ,

$$P = I - S, \quad (10)$$

and our objective function is defined as

$$L = \|P - V\|_1, \quad (11)$$

where $\|\cdot\|_1$ denotes the L_1 norm. Training is carried out by minimizing the above-mentioned loss function using a mini-batch gradient descent based on backpropagation. Our goal is to learn a model with the optimal parameters (W and B) that can estimate a P as close as possible to its corresponding V . The mean squared error (MSE) or L_2 is the most widely used loss function for image restoration problems. However, L_2 correlated poorly with human observations for image quality assessment [33]. In DLS-NUC, we make use of a L_1 loss function to drive our learning.

C. More Analysis

In this section, we evaluate how each significant component of our DLS-NUC method contributes to performance. Specifically, we examine the impact of three components, including the column FPN model, pooling/deconvolution layers, and residual learning strategy.

1. Column FPN Model

The degree number M of the column FPN model is a critical parameter that determines the complexity of a simulated training dataset. We experimentally evaluate the performances of five alternatives using different degree numbers. Specifically, we generate five different training datasets using (a) an offset model ($M = 0$) [14,30], (b) a linear model ($M = 1$) [3,6,9],

(c) a quadratic model ($M = 2$) [8], (d) a cubic model ($M = 3$), and (e) a quartic model ($M = 4$), respectively. Then we train five network models and evaluate them on 20 testing images with column FPN simulated using these five different models. For a fair comparison, we adopt the same deep network architecture and make the size of the training dataset equal. The comparative results in terms of peak signal-to-noise ratio (PSNR) values are shown in Table 1. It is noted that a network model trained using simulated data of the noise model with larger M always achieves a higher PSNR value. A more comprehensive model provides more complex column FPN patterns, and thus its trained network model has a better expressive/generalization ability. However, the improvement becomes insignificant when M is larger than 3. In Fig. 4 we also show some comparative NUC results of using different column FPN models. In our implementation, we set $M = 3$ to achieve a good balance between model complexity and good performance.

2. Pooling/Deconvolution Layers

In deep neural networks, the receptive field determines how many neighboring pixels in the input layer are considered to restore lost information. For the task of strip NUC, a reasonably large receptive field is utilized to capture global information for accurate reconstruction of column FPN. Two common solutions to enlarge the receptive field include using large kernels and stacking more layers; however, both need to employ more parameters and adversely increase the runtime of deep network models.

Another feasible solution is to add pooling layers, which was originally presented by Simonyan and Zisserman [34] and He *et al.* [32]. In this paper, we employ a pooling layer to enlarge the receptive field and a deconvolution layer to restore the original spatial resolution, respectively. The feature maps extracted in the low-resolution space are fused with ones extracted in the original space to preserve image details. Moreover, performing feature extraction on low-resolution spaces will increase the processing speed. In Fig. 5, we show network architectures with (DLS-NUC) and without (DLS-NUC Plain) pooling/deconvolution layers. Their architecture settings are listed in Table 2. It is noted that our DLS-NUC can achieve a significantly larger receptive field with the same number of convolutional layers.

As illustrated in Fig. 6, DLS-NUC effectively removes significant column FPN. In comparison, some obvious strips are still visible in the result of the DLS-NUC plain. For runtime comparison, we execute both models 10 times and compute their average processing times. These experiments are performed in Matlab R2015b without optimizations or

Table 1. PSNR Values of Different Trained Models Applied to Different Simulated Datasets (Unit: dB)^a

Models/Datasets	Sim. with A	Sim. with B	Sim. with C	Sim. with D	Sim. with E	Average
Trained with A	39.71	39.92	39.89	40.23	39.90	39.93
Trained with B	39.97	40.07	40.24	40.53	40.31	40.22
Trained with C	39.98	40.08	40.25	40.55	40.33	40.24
Trained with D	40.00	40.11	40.30	40.62	40.38	40.28
Trained with E	40.01	40.12	40.31	40.62	40.40	40.29

^aSim. with A denotes a dataset containing noisy testing images simulated with noise model A. The average of these 5 datasets are also provided.

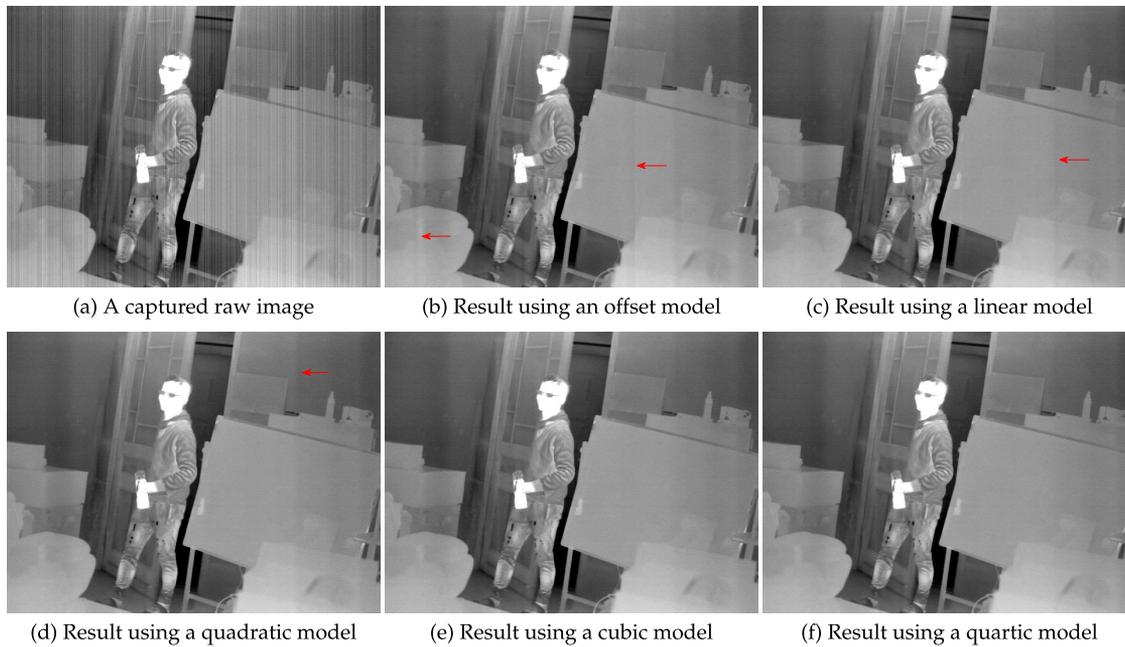


Fig. 4. NUC results using different column FPN models. (a) A raw image with obvious column FPN and comparative results of DLS-NUC trained using (b) an offset model, (c) a linear model, (d) a quadratic model, (e) a cubic model, and (f) a quartic model. It is noted that network trained using simulated data of the offset ($M = 0$) or linear ($M = 1$) models cannot produce satisfactory NUC results. The NUC results are very similar when M is larger than 3. In our implementation, we set $M = 3$ to achieve a good balance between model complexity and good performance.

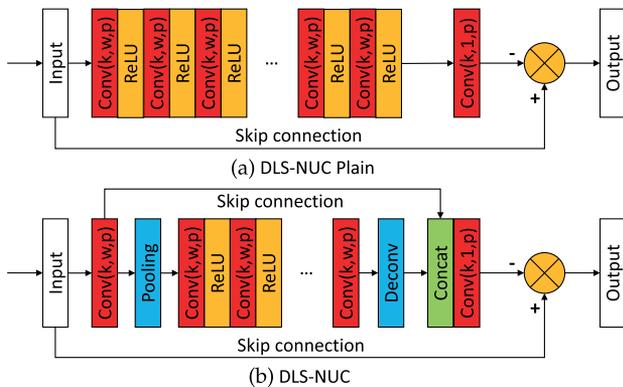


Fig. 5. Architectures of two deep network models. (a) The DLS-NUC plain model without pooling/deconvolution layers. (b) The DLS-NUC model with pooling/deconvolution layers.

parallel-computing implementation on a PC equipped with an Intel Core i7-6820HK CPU (2.70 GHz) and 16 GB memory. It is worth mentioning that DLS-NUC still achieves a faster running speed, although it contains three extra layers (pooling, deconvolution, and concatenation). This is because the processing times of a smaller-size image in subsequent convolutional layers are significantly reduced by a factor of r^2 .

3. Residual Learning

With the increase of the network depth, the training process becomes more difficult due to the gradient exploring/vanishing problem [32]. In this paper, we first introduce the residual learning strategy to the context of deep-learning-based NUC. Instead of reconstructing the noise-free image [14],

we attempt to compute the residual information and subtract it from the raw image to remove column FPN. Performance curves with/without residual learning are shown in Fig. 7. First, we observe that the residual-learning network converges much faster than non-residual-learning one. Second, a residual-learning network obtains a more accurate restoration when the training process is converged. The PSNR values are calculated on our 20 testing images with simulated column FPN.

4. EXPERIMENTAL RESULTS

A. Implementation Details

In our DLS-NUC method, a cubic column FPN model is utilized for generating training pairs, and we set N and r to 8 and 3, respectively, to form our network architecture. Our network comprises 10 convolutional layers, 1 pooling layer, 1 deconvolution layer, and 1 concatenation layer. For each convolutional layer, (k, w, p) is set to $(3, 32, 1)$, except the final convolutional layer in the reconstruction part, where we set (k, w, p) to $(3, 1, 1)$ to reconstruct the residual information between paired training images. Training is carried out by optimizing the loss function using the “Adam” optimizer with a mini-batch of 64 sub-images. The weights are initialized according to the method described in [35], which is proved effective for networks employing ReLUs. The initial learning rate is set to 0.0001 and decreases by a factor of 10 every 40 epochs, and training is regularized by weight-decay (penalty item multiplied by 0.0001). We empirically train our model by 80 epochs on a single GPU of NVIDIA GTX 1080Ti.

Our DLS-NUC model is trained using simulated noisy images, and therefore we hope to investigate its effectiveness on real-captured infrared images. For this purpose, we make use of

Table 2. Architecture Settings of DLS-NUC Plain and DLS-NUC^a

Network	# Convolution	Pooling	Resizing Factor	Deconvolution	Concatenation	Receptive Field
DLS-NUC Plain	10	×	—	×	×	21 × 21
DLS-NUC	10	✓	3	✓	✓	53 × 53

^aPooling/convolution layers are employed in DLS-NUC to enlarge the receptive field.

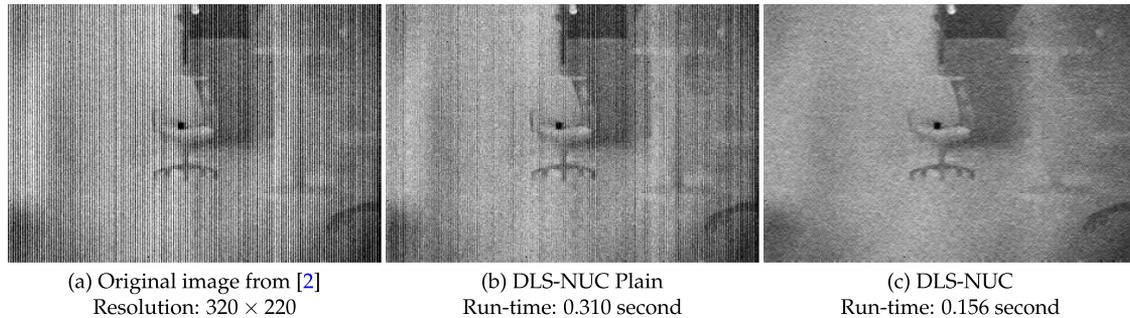


Fig. 6. Comparative NUC results. (a) A raw image with obvious column FPN, (b) NUC result of DLS-NUC plain model, (c) NUC result of DLS-NUC. DLS-NUC employs pooling/deconvolution layers to enlarge the receptive field and can thus better eliminate strip noise using less runtime.

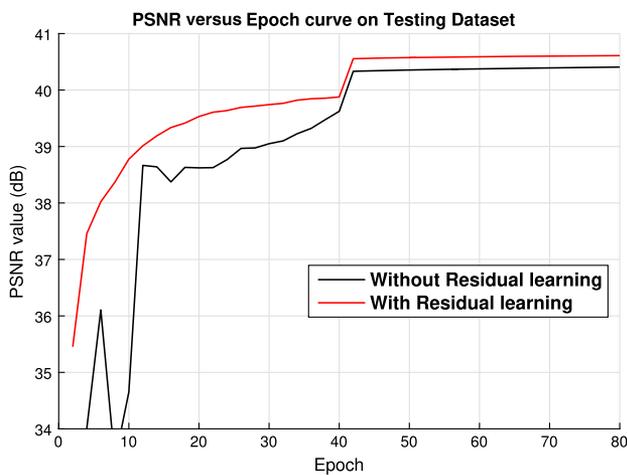


Fig. 7. PSNR versus epoch curves of different configurations.

a 20-image infrared dataset (Data1), which is available under the Creative Commons Attribution (CC-BY) license [2] (http://demo.ipol.im/demo/glmt_mire/). In addition, we turn off the on-chip calibration of our long-wave infrared camera and capture 10 infrared images with obvious strip noise (Data2). These testing images will be made publicly available in the future.

B. Qualitative Evaluation

We compare our DLS-NUC method with state-of-the-art single-image-based strip NUC solutions. The handcrafted NUC methods include the WD-FT method [27], the Midway Histogram Equalization (MHE)-based method [2], and the 1D guided-filter-based method (GF) [9]. We also consider the deep-learning-based NUC method (SNRCNN) [14]. The source codes of the WD-FT (<ftp://ftp.empa.ch/>

pub/empa/outgoing/BeatsRamsch/stripFilter/xStripes.jar), MHE (<http://www.ipol.im/pub/art/2012/glmt-mire/>), and SNRCNN (https://github.com/Kuangxd/Train_SNRCNN) methods are publicly available, and the implementation of the GF method [9] is provided by the authors.

In Fig. 8, we show NUC results of an infrared image with obvious vertical edges. It is observed that some obvious strip noise remains visible in the correction results of the WD-FT, MHE, and SNRCNN methods. Moreover, the WD-FT and MHE methods falsely generate some artifacts, which significantly degrade image quality. Overall, the GF method achieves satisfactory correction results; however, it over-smooths the vertical edge (building boundary) and adds some blurry visual effects to its output. In comparison, our proposed DLS-NUC method can effectively remove FPN in infrared images without losing original information or causing artifacts. In Fig. 9, we show NUC results of a raw infrared image with severe column FPN using different methods. It is observed that our proposed DLS-NUC method significantly outperforms other alternative solutions since it can better remove column FPN and preserve original thermal details. As a result, human targets can be easily identified in the output of our DLS-NUC method.

We also evaluate the DLS-NUC method on our own captured infrared images. It is noted that column FPNs in our captured raw images present different visual characteristics. Comparative NUC results are illustrated in Fig. 10, and it is visually observed that our DLS-NUC achieves the best noise correction performance. It completely removes strips without blurring details and generates no artifacts. The underlying reason is that DLS-NUC model is trained using a large amount of high-fidelity simulated data; therefore, it can successfully handle images captured using different types of infrared devices. Figure 11 shows an example that the proposed DLS-NUC model favorably preserves original information when applied

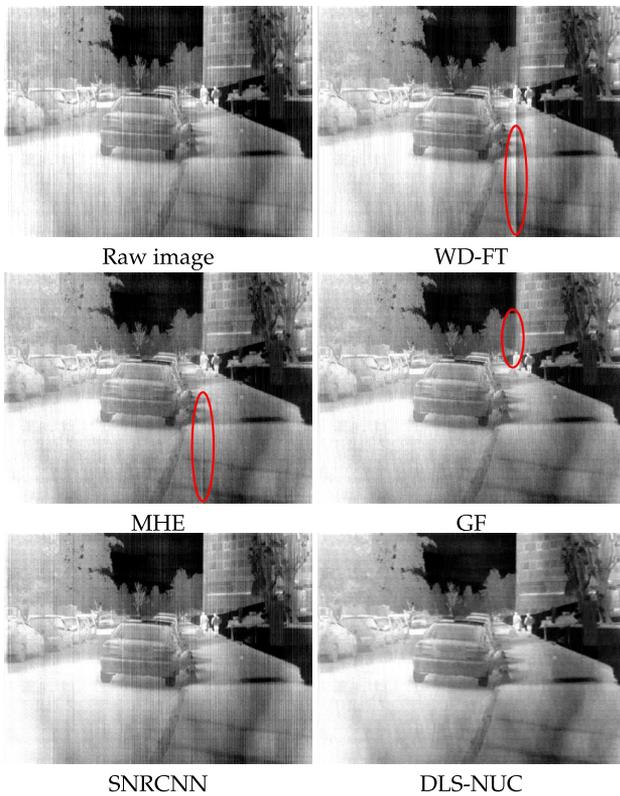


Fig. 8. Strip NUC results of WD-FT [27], MHE [2], SNRCNN [14], GF [9], and our DLS-NUC methods on an infrared image with obvious vertical edges.

to a noise-free image, while other non-deep-learning NUC methods cause obvious vertical artifacts. The SNRCNN method falsely removes some image details. Through deep learning of big high-fidelity data, the DLS-NUC method can accurately distinguish between undesirable FPN and original image details. Therefore, it effectively removes strip noise and preserves original information.

C. Quantitative Evaluation

We make use of two reference-free metrics (roughness index ρ [19,29,36] and smoothing difference index D_{sT}^{sF} [37]) to quantitatively evaluate performances of different column FPN correction methods. Roughness index ρ , which is used to measure the high-pass content of an image, was first introduced by Hayat *et al.* [19] to evaluate the performance of NUC methods, and its definition is

$$\rho = \frac{\|b_1 * P\|_1 + \|b_2 * P\|_1}{\|P\|_1}, \quad (12)$$

where P denotes the restored image to evaluate, $b_1 = [-1, 1]$ is a horizontal mask, $b_2 = [-1, 1]$ is a vertical mask, $\|\cdot\|_1$ indicates the L_1 norm, and $*$ denotes the convolution operation. A NUC result with a smaller ρ value indicates column FPN is better suppressed. D_{sT}^{sF} evaluates how horizontal smoothing is performed in structure and non-structure regions as follows:

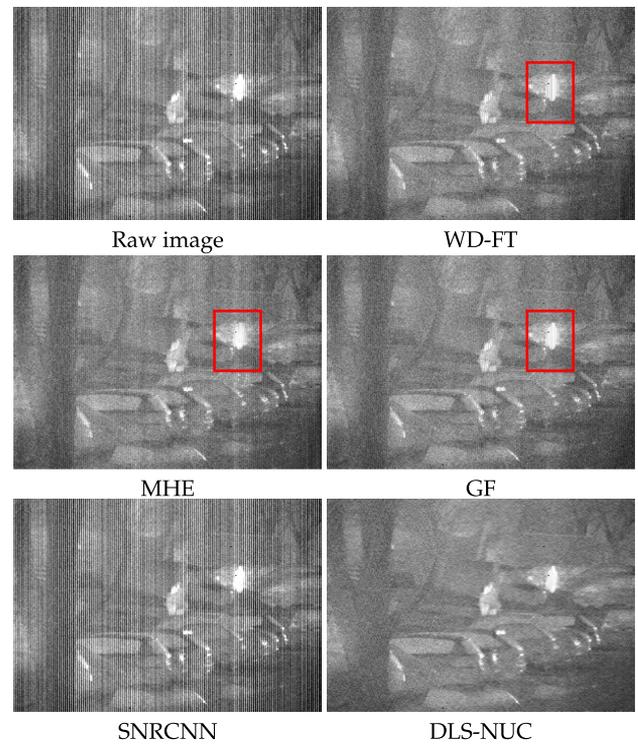


Fig. 9. Correction results of a raw infrared image with severe column FPN.

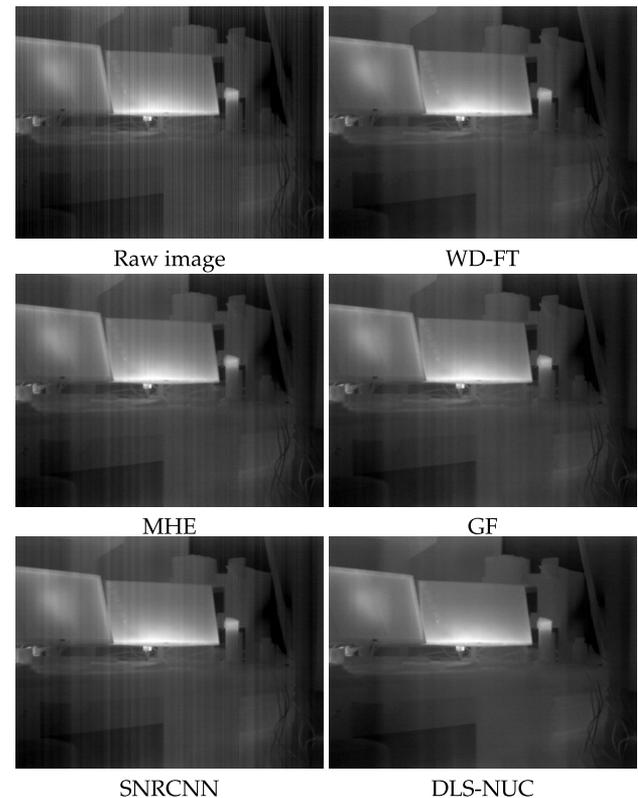


Fig. 10. Correction results of a raw infrared image captured using our own uncooled long-wave infrared camera.

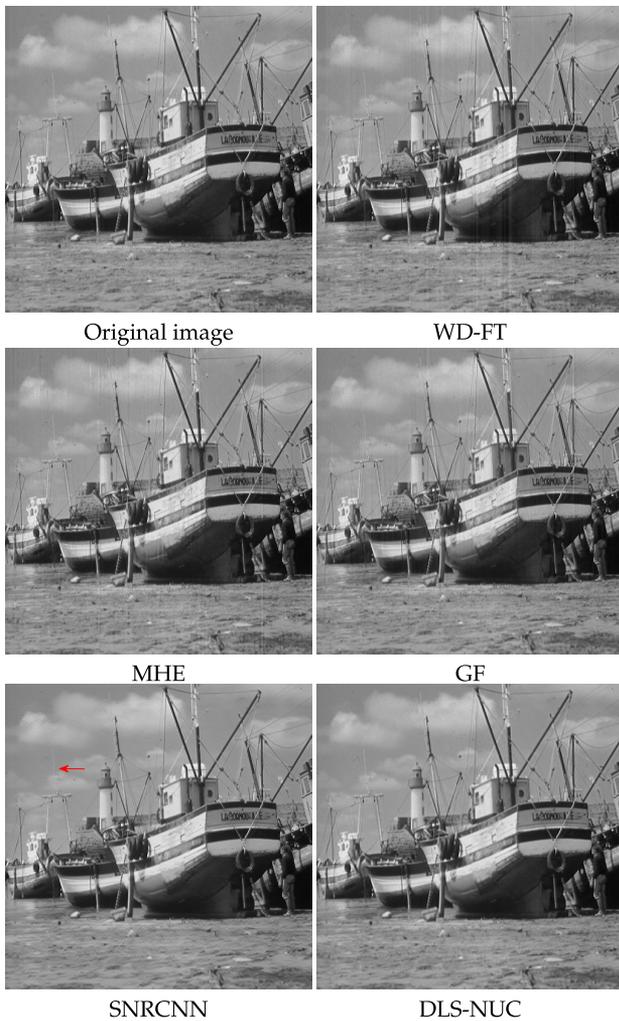


Fig. 11. Processing results of a noise-free image using different NUC methods. It is noticed that our DLS-NUC well preserves original information and causes no obvious vertical artifacts.

$$D_{s_T}^{s_F}(G, I) = \frac{\sum_{i \in s_T} \partial_x G(I(i))}{\sum_{i \in s_T} \partial_x I(i)} - \frac{\sum_{i \in s_F} \partial_x G(I(i))}{\sum_{i \in s_F} \partial_x I(i)}, \quad (13)$$

where $I(i)$ is the value of a pixel i in raw image I , $G(\cdot)$ denotes a proposed column FPN correction method, s_T and s_F define pixels in structure and non-structure regions, respectively. A higher $D_{s_T}^{s_F}(G, I)$ value indicates the proposed FPN correction method G has a better ability to suppress column FPN while preserving original information.

Table 3. Quantitative Evaluation of WD-FT [27], MHE [7], GF [9], SNRCNN [14], and Our DLS-NUC Methods

Methods	Mean ρ		Mean $D_{s_T}^{s_F}$	
	Data1 [2]	Data2	Data1 [2]	Data2
WD-FT [27]	0.3393	0.1360	0.37	1.45
MHE [7]	0.2924	0.1396	0.34	1.33
GF [9]	0.2498	0.1166	0.30	1.31
SNRCNN [14]	0.5229	0.1452	0.15	1.53
DLS-NUC	0.1963	0.1134	0.55	1.58

The quantitative evaluation results (mean $D_{s_T}^{s_F}$ and mean ρ) are shown in Table 3. It is observed that our DLS-NUC method yields the lowest ρ and the highest $D_{s_T}^{s_F}$ values for both Data1 and Data2. The experimental results illustrate that our proposed method not only better eliminates column FPN but also preserves original thermal details.

5. CONCLUSION

In recent years, many changing image restoration problems have been successfully solved through deep learning of big data. In this paper, we made the attempt to learn a better-performing strip NUC operator from a large number of simulated training images. We present a unified framework, DLS-NUC, which combines a column FPN simulation module with convolutional neural networks (CNNs) to train the optimal model for column FPN correction of infrared images. Compared with existing handcrafted NUC methods, our proposed single-image-based method has many desirable advantages including superior noise reduction, better detail preservation, being artifact-free, and having no user-specific parameters. This is the first research work revealing that a NUC model trained using simulated data can be successfully employed to handle real-captured raw infrared images. In the future we plan to extend this deep-learning approach by incorporating more FPN models to compensate for other types of FPNs presented in infrared FPA (e.g., periodical FPNs presented between two adjacent periods [3], optics temperature-dependent FPNs [38,39], and FPNs of individual detectors [18,22,25]). We will also make efforts to optimize the processing speed of this deep network model for real-time implementation.

Funding. National Natural Science Foundation of China (NSFC) (51575486, 51605428, U1664264); Fundamental Research Funds for the Central Universities.

REFERENCES

1. C.-C. Hsieh, C.-Y. Wu, F.-W. Jih, and T.-P. Sun, "Focal-plane-arrays and CMOS-readout techniques of infrared imaging systems," *IEEE Trans. Circuits Syst. Video Technol.* **7**, 594–605 (1997).
2. Y. Tendero, S. Landeau, and J. Gilles, "Non-uniformity correction of infrared images by midway equalization," *Image Process. Line* **2**, 134–146 (2012).
3. Z. Liu, J. Xu, X. Wang, K. Nie, and W. Jin, "A fixed-pattern noise correction method based on gray value compensation for TDI CMOS image sensor," *Sensors* **15**, 23496–23513 (2015).
4. C. Minassian, J. C. Tissot, M. Vilain, O. Legras, S. Tinnes, B. Fieque, J. M. Chiappa, and P. Robert, "Uncooled amorphous silicon TEC-less 1/4 VGA IRFPA with 25 μm pixel-pitch for high volume applications," *Proc. SPIE* **6940**, 69401Z (2008).
5. Y. Tendero and J. Gilles, "Admire a locally adaptive single image non-uniformity correction and denoising algorithm: application to uncooled IR camera," *Proc. SPIE* **8353**, 83531O (2012).
6. B. Narayanan, R. C. Hardie, and R. A. Muse, "Scene-based nonuniformity correction technique that exploits knowledge of the focal-plane array readout architecture," *Appl. Opt.* **44**, 3482–3491 (2005).
7. W. Qian, Q. Chen, G. Gu, and Z. Guan, "Correction method for stripe nonuniformity," *Appl. Opt.* **49**, 1764–1773 (2010).
8. Y. Cao and Y. Li, "Strip non-uniformity correction in uncooled long-wave infrared focal plane array based on noise source characterization," *Opt. Commun.* **339**, 236–242 (2015).

9. Y. Cao, M. Y. Yang, and C.-L. Tisse, "Effective strip noise removal for low-textured infrared images based on 1D guided filtering," *IEEE Trans. Circuits Syst. Video Technol.* **26**, 2176–2188 (2016).
10. Y. Cao, Z. He, J. Yang, X. Ye, and Y. Cao, "A multi-scale non-uniformity correction method based on wavelet decomposition and guided filtering for uncooled long wave infrared camera," *Signal Process.* **60**, 13–21 (2018).
11. C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.* **38**, 295–307 (2015).
12. S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Computer Vision and Pattern Recognition (CVPR)* (2017).
13. C. Yang, X. Lu, Z. Lin, E. Shechtman, O. Wang, and H. Li, "High-resolution image inpainting using multi-scale neural patch synthesis," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017).
14. X. Kuang, X. Sui, Q. Chen, and G. Gu, "Single infrared image stripe noise removal using deep convolutional networks," *IEEE Photon. J.* **9**, 3900913 (2017).
15. J. M. Mooney, F. D. Shepherd, W. S. Ewing, J. E. Murguia, and J. Silverman, "Responsivity nonuniformity limited performance of infrared staring cameras," *Opt. Eng.* **28**, 1151–1161 (1989).
16. M. Schulz and L. Caldwell, "Nonuniformity correction and correctability of infrared focal plane arrays," *Infrared Phys. Technol.* **36**, 763–777 (1995).
17. X. Sui, Q. Chen, and G. Gu, "Adaptive grayscale adjustment-based stripe noise removal method of single image," *Infrared Phys. Technol.* **60**, 121–128 (2013).
18. J. G. Harris and Y.-M. Chiang, "Nonuniformity correction of infrared image sequences using the constant-statistics constraint," *IEEE Trans. Image Process.* **8**, 1148–1151 (1999).
19. M. M. Hayat, S. N. Torres, E. Armstrong, S. C. Cain, and B. Yasuda, "Statistical algorithm for nonuniformity correction in focal-plane arrays," *Appl. Opt.* **38**, 772–780 (1999).
20. B. M. Ratliff, M. M. Hayat, and R. C. Hardie, "An algebraic algorithm for nonuniformity correction in focal-plane arrays," *J. Opt. Soc. Am. A* **19**, 1737–1747 (2002).
21. B. M. Ratliff, M. M. Hayat, and J. S. Tyo, "Radiometrically accurate scene-based nonuniformity correction for array sensors," *J. Opt. Soc. Am. A* **20**, 1890–1899 (2003).
22. B. M. Ratliff, M. M. Hayat, and J. S. Tyo, "Generalized algebraic scene-based nonuniformity correction algorithm," *J. Opt. Soc. Am. A* **22**, 239–249 (2005).
23. R. C. Hardie, M. M. Hayat, E. Armstrong, and B. Yasuda, "Scene-based nonuniformity correction with video sequences and registration," *Appl. Opt.* **39**, 1241–1250 (2000).
24. S. N. Torres and M. M. Hayat, "Kalman filtering for adaptive nonuniformity correction in infrared focal-plane arrays," *J. Opt. Soc. Am. A* **20**, 470–480 (2003).
25. A. Averbuch, G. Liron, and B. Z. Bobrovsky, "Scene based non-uniformity correction in thermal images using Kalman filter," *Image Vis. Comput.* **25**, 833–851 (2007).
26. W. Qian, Q. Chen, and G. Gu, "Minimum mean square error method for stripe nonuniformity correction," *Chin. Opt. Lett.* **9**, 051003 (2011).
27. B. Münch, P. Trtik, F. Marone, and M. Stipanoni, "Stripe and ring artifact removal with combined wavelet-Fourier filtering," *Opt. Express* **17**, 8567–8591 (2009).
28. H. Pfister, W. Matusik, N. J. W. Morris, and S. Avidan, "Statistics of infrared images," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2007), pp. 1–7.
29. J. Zhao, Q. Zhou, Y. Chen, T. Liu, H. Feng, Z. Xu, and Q. Li, "Single image stripe nonuniformity correction with gradient-constrained optimization model for infrared focal plane arrays," *Opt. Commun.* **296**, 47–52 (2013).
30. N. Högasten and M. Ingerhed, "Systems and methods for processing infrared images," U.S. patent 8,780,208 (15 July 2014).
31. V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *International Conference on Machine Learning* (2010), pp. 807–814.
32. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition* (2016).
33. H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Trans. Comput. Imaging* **3**, 47–57 (2017).
34. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations (ICLR)* (2015), pp. 1–14.
35. K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on imagenet classification," in *IEEE International Conference on Computer Vision* (2015).
36. R. Lai, Y. T. Yang, Q. Li, and H. X. Zhou, "Improvement in adaptive nonuniformity correction method with nonlinear model for infrared focal plane arrays," *Opt. Commun.* **282**, 3444–3447 (2009).
37. Y. Cao, Z. He, J. Yang, Y. Cao, and M. Y. Yang, "Spatially adaptive column fixed-pattern noise correction in infrared imaging system using 1D horizontal differential statistics," *IEEE Photon. J.* **9**, 7803513 (2017).
38. A. Wolf, J. E. Pezoa, and M. Figueroa, "Modeling and compensating temperature-dependent non-uniformity noise in IR microbolometer cameras," *Sensors* **16**, 1121–1134 (2016).
39. 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.* **39**, 646–648 (2014).