IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Dual-Branch Structured De-Striping Convolution Network Using Parametric Noise Model

JONGHO. LEE[1,2], and YONGMAN. RO[1], Senior Member, IEEE

[1]Image and Video Systems Lab, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea
[2]Agency for Defense Development (ADD), Daejeon, South Korea

Corresponding author: Prof. Yong Man Ro (e-mail: ymro@ kaist.ac.kr).

ABSTRACT The stripe fixed pattern noise (FPN) of infrared images significantly corrupts image quality, so that most infrared imaging systems suffer from the degradation of visibility and detectability during operation. Therefore, the FPN de-striping method, which eliminates stripe patterns without substantial loss of image information, remains a core technology in the field of infrared image processing. In this paper, we propose the dual-branch structure based FPN de-striping deep convolutional neural network (DBS-DCN) to effectively extract structural features of FPN and preserve the image details in a single infrared image. In addition, we have established the parametric FPN model through the diagnostic experiments of infrared images based on the physical principle of an infrared detector and its signal response. We have optimized each parameter of the FPN model using measured data, which acquired on a wide range of detector temperatures. Further, we generate the training data using our FPN model to ensure stable learning performance against various stripe patterns. We performed comparative experiments with state-of-the-art methods using artificially corrupted infrared images and real corrupted infrared data, and our proposed method achieved outstanding de-striping results in both qualitative and quantitative evaluation compared to existing methods.

INDEX TERMS Infrared image, Fixed Pattern Noise, De-striping method, Convolution network

I. INTRODUCTION

Infrared detector creates infrared images by converting the radiant energy from a scene into an electrical signal [1]. Compared with visible images using the reflected energy from the Sun, infrared images are less affected by changes in illumination and can be utilized both daytime and nighttime with the capability to detect small differences in temperature. By leveraging these advantages, applications using infrared images have been extensively expanded to automotive night vision system, aviation weather observation system, industrial facility inspection system, and medical diagnosis, etc.

However, despite the advantages of infrared images that can distinguish temperature differences, infrared data have non-uniform response properties due to various factors. The first non-uniformity caused by geometrical-optics properties of the infrared imaging system appears as the low-frequency concentric patterns in infrared images [1]. The second is the nonuniformly smooth biased patterns, which occurs because the energy emitted from the internal optical system, not from the scene, reaches the focal plane array (FPA) of an infrared detector [2]. The last non-uniformity is the fixed pattern noise (FPN), which caused by the FPA signal readout mechanism of an infrared detector [3].

Among the aforementioned non-uniformities, due to the inherent property of infrared detector, FPN can appear in the raw infrared image even though uniform energy is provided to an infrared detector. FPN has vertical directionality and periodicity because the detector readout circuit (ROIC) process the signal of entire FPA in row by row, whereas the pixels in the same column share a column-wise multiplexer [3], [4]. As shown in Fig. 1, FPN can significantly degrade the quality of infrared image. The ultimate aim of the FPN de-striping method is to reconstruct a noise-free image with no substantial stripe patterns from a noisy infrared image. To this end, various researches have been conducted such as a statistics-based method [5], filtering-based method [6], etc. Although these methods remove FPN from noisy image data, they still have limitations such as over-smoothed images, unexpected ghosting artifacts, and parameters requiring manual adjustment [5]-[8].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access

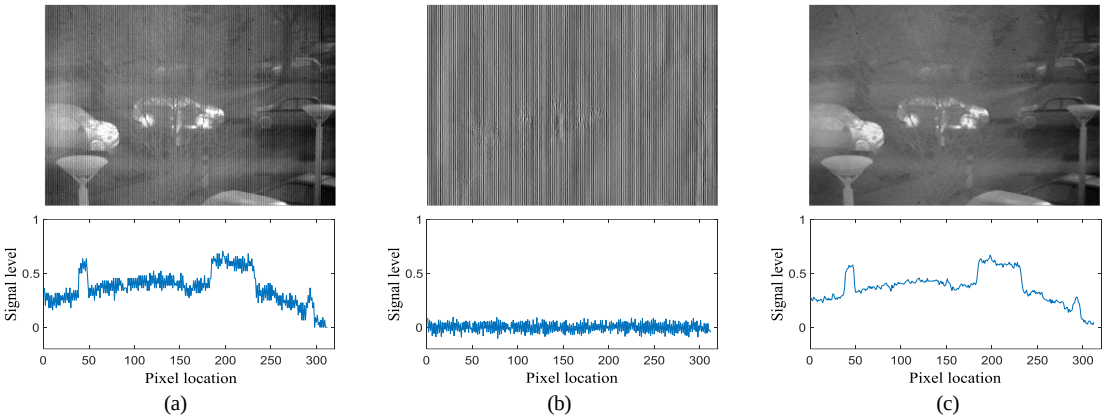


FIGURE 1. **FPN De-striping result and horizontal profile (a) Corrupted Image with FPN. (b) Estimated FPN. (c) Reconstructed Image.**

Recently, many researchers have introduced deep-learning based approaches to address quality enhancement in infrared images, and achieved progressive results [9], [10]. Motivated by the success of deep learning in the infrared domain, deep CNN (convolutional neural network) based FPN removal methods have been actively proposed and they accomplished remarkable performance improvement over the conventional methods. Kuang *et al*. [11] introduced a simple three-layer convolutional network for the stripe FPN removal from single infrared image. He *et al*. [4] implemented a residual deep network-based FPN de-striping method to compute residual information through learning. On the one hand, Xiao *et al*. [8] presented a deep convolutional network utilizing a local-global combination model for infrared cloud images with FPN. Chang *et al*. [7] proposed a multi-scale residual deep convolutional network for stripe noise and bias field removal. In [12], they applied wavelet transform to the deep neural network for strip noise removal.

However, despite the advanced performances of the above methods, FPN de-striping method is still a challenging task that needs to preserve image details while removing FPN from low-contrast infrared images [13], [14]. If infrared image has stronger FPN than the image detailed information or infrared data include complex structures, a large amount of information in infrared image can be lost in the FPN removal process. On the other hand, some FPN may remain in the reconstructed images even after the de-striping process. We call this the residual FPN.

To address the FPN removal completely without over-smoothing and residual FPN, in this paper, we propose a novel dual-branch structured de-striping deep convolutional network (DBS-DCN).

Our contributions are as follows.

- We establish a new parametric FPN model through the diagnostic experiments.
- We propose a novel dual-branch structured network that is effective and efficient to extract various scaled features.
- We have performed qualitative and quantitative evaluations to verify the proposed method with real infrared noisy images and artificially corrupted infrared images.

II. RELATED WORKS

A. FPN PROPERTIES

An appropriate information of FPN can play an important role in the stripe FPN de-striping discipline. According to [12], FPN has a vertical and periodic property and they demonstrated that FPN is densely extracted in the horizontal coefficients of the Haar discrete wavelet transform (HDWT). To the best of our knowledge, Cao *et al*. [15] derived the relationship between FPN and infrared data within a column through their thermal calibration experiments for the first time. From results of experimental analysis, they exploited a non-linear cubic FPN model to be used their training data [4]. Other approaches ([7], [11], [12], and [35]) define the stripe pattern noise as a random distribution model with a mean of zero and a small standard deviation.

B. EFFECTIVE FEATURE EXTRACTION

In the deep-learning-based approach, a receptive field that contributes to the output features of next layer is an important consideration to design an effective network [17]. Common methods to enlarge the receptive field are stacking additional layers, downsampling of feature size, and using the dilated convolutional filter, etc. DLS-NUC [4] employed the max-pooling layer to increase the receptive field in their feature extraction phase. Chang et al. [7] proposed the multi-scale residual deep convolutional network (DMRN) based on the encoding-decoding manner and they aimed to extract both large-scale features and fine textures. ICSRN [8] used the convolution filter of size 7×7 in their first four layers and attempted the receptive field to expand globally. In addition, they applied the structure called local-global combination to recover the detail information of image. On the one hand, in [35], Guan *et al*. introduced a mixed convolutional layer consisting of dilated convolution, sub-pixel convolution, and standard convolution to extract the multi-grained features in various scales.

C. RESIDUAL LEARNING STRATEGY

Kuang *et al* [11] trained to find the mapping relationships with the noise-free images from the corrupted input image data.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access

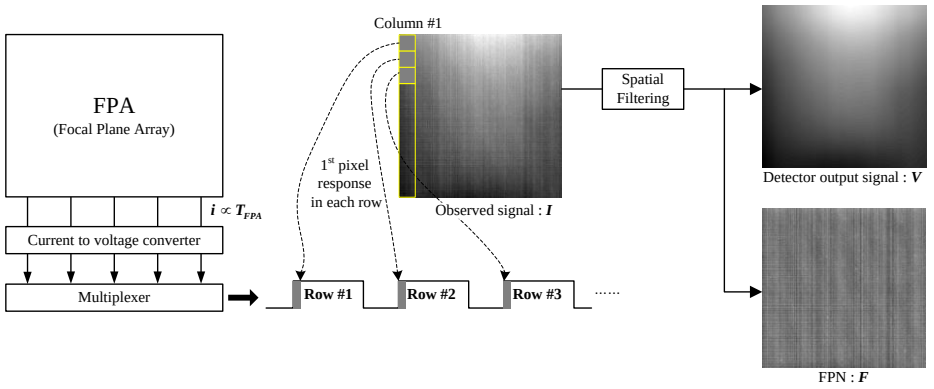


FIGURE 2. **Signal readout mechanism of Infrared detector. The observed image for the uniform blackbody source is corrupted by FPN.**

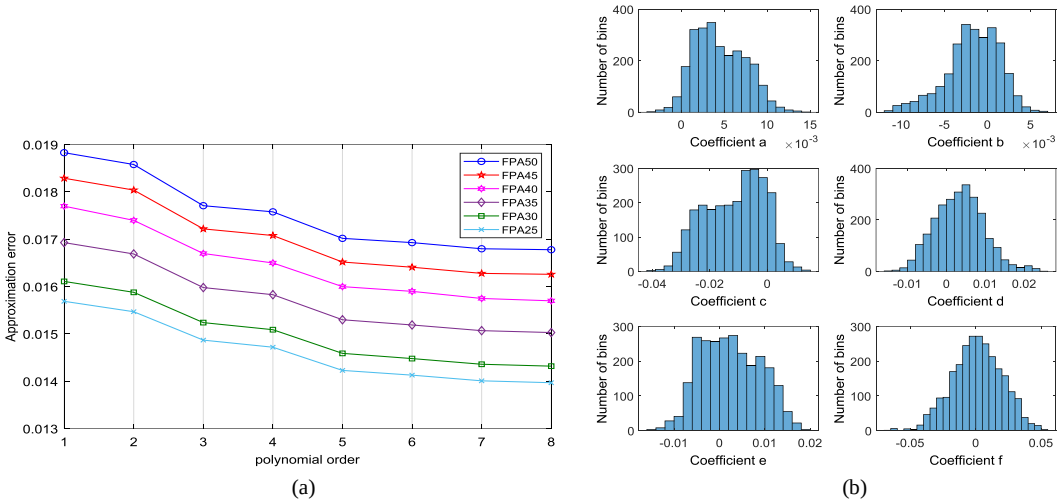


FIGURE 3. **The approximation errors and distributions of each coefficient values for the parametric FPN model. (a) Approximation errors between detector response and FPN for the order of polynomial of the acquired data at the FPA temperature range of 25℃ to 50℃, (b) Distributions of each coefficient values in the fifth-order polynomial model.**

On the other hand, instead of direct reconstructing a de-noised image, DnCNN [18] considered to estimate the difference between the noisy image and the noise of image. A corrupted images including FPN have highly correlated relationships with their ground truth images. Therefore, according to [19], a residual learning methodology for the residual information (*i.e.,* FPN itself) training could be an appropriate approach in terms of efficiency and effectiveness. In recent years, various residual learning-based FPN removal approaches have been proposed and achieved optimized performance with efficient convergence [4], [7].

III. PROPOSED APPROACH

A. PARAMETRIC FPN MODEL

The importance of training data in deep learning is well known. Besides, how well the FPN model was applied to the training data determines learning performance in the field of stripe pattern removal. As the first process of preparing training data, we derived the model of the corrupted infrared image due to FPN as follows. Ideally, the image signal corresponding to detector response should be uniform when uniform energy incident to infrared detector FPA. However, a single infrared image obtained for the uniform blackbody is significantly contaminated with FPN as shown in Fig. 2.

Normally, the observed image model of infrared signal is a linear composition [16].

$$\mathrm{I} = \mathrm{G} * \mathrm{X} + \mathrm{O} \tag{1}$$

where X is overall response of detector FPA at blackbody temperature corresponding to irradiance energy, and I is the observed image signal. G is the output signal gain associated with a temperature of the blackbody and O is the signal offset. The symbol $*$ stands for the element-wise product. When we represent the image I as a combination of columns, the signal difference between the two adjacent columns (i_1, i_2) of a single image without FPN should be almost zero. We can derive the difference between columns as (2) by substituting each column signal into (1).

$$\begin{aligned} \mathrm{i}_2 - \mathrm{i}_1 &= \mathrm{g}_2 * \mathrm{x}_2 + \mathrm{o}_2 - (\mathrm{g}_1 * \mathrm{x}_1 + \mathrm{o}_1) \\ &= (\mathrm{g}_2 * \mathrm{x}_2 - \mathrm{g}_1 * \mathrm{x}_1) + \mathrm{o}_2 - \mathrm{o}_1 \end{aligned} \tag{2}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access

Here, since X is a uniform response for the blackbody, x_1 and x_2 corresponding to column signal of the X are the same. On the other hand, through the analysis of acquired single image, we ascertained that each column has almost identical slope (*i.e.,* $g_1 \approx g_2$). Sometimes, G can be considered as a constant value, rather than a matrix in a single infrared image obtained at any blackbody temperature [16].

Meanwhile, even in a single infrared image, the offset of each column is not the same. As described in [3], infrared detector generates the FPA signal row by row. The pixels in the same position of each row are delivered through the column-wise multiplexer of the detector ROIC as shown in Fig.2. Therefore, temperature across the FPA is not consistent during sequential readout process and detector signal level will be affected by changes in the FPA temperature. As a result, signal difference due to the FPA temperature drift causes the offset difference between adjacent columns and induces vertical patterns in a single image.

From the above inferences, we cannot ignore the second term whereas the first term can be neglected in equation (2). If the image consist of n columns (*i.e.,* $I = [i_1, i_2, \dots, i_n] \in \mathbb{R}^n$), we can represent each column as a combination of the first column i_1 and offset difference.

$$\begin{aligned} i_2 &= i_1 + \Delta o_{2,1} \\ i_3 &= i_2 + \Delta o_{3,2} = i_1 + (\Delta o_{2,1} + \Delta o_{3,2}) \\ &\cdots \\ i_n &= i_1 + (\Delta o_{2,1} + \cdots + \Delta o_{n,n-1}) \end{aligned} \tag{3}$$

where Δo indicates the offset difference between the columns corresponding to the subscript number. From the perspective of single infrared image for the uniform blackbody source, we can regard the offset differences as FPN. Therefore, if we reconstructed the image I with a combination of columns, the image I of (1) can be represented by the uniform signal V and the FPN F as shown in equation (4).

$$\begin{aligned} I &= [i_1, i_2, \dots, i_n] \\ &= [i_1, i_1, \dots, i_1] + (\Delta o_{2,1} + \cdots + \Delta o_{n,n-1}) \\ &= V + F \end{aligned} \tag{4}$$

In (4), the first term of the second line can be regarded as the uniform because all components are the same column. This degradation model has been proven to be effective in many CNN-based FPN removal methods [4], [7], [11].

In order to examine the FPN properties against the FPA temperature variations, we have conducted the experiments to diagnose real infrared images referring to [15]. First, we adjusted the FPA temperature from 25 to 50 degrees Celsius while the blackbody temperature fixed (*i.e.,* incident energy from the scene is constant), and obtained the average image of 1000 frames at each FPA temperature to exclude temporal random noise. According to [6], infrared image for the uniform blackbody contains only signal bias corresponding to detector response and the stripe FPN, so that we divided the acquired infrared image into FPN and detector response by applying the mean filter with the large-sized kernel as shown in Fig. 2. Secondly, we performed the polynomial approximation to derive pixel-by-pixel correlations between FPN and detector response for the FPA temperature. By analysis results of the LMS (Least mean square) error for each polynomial order, we have ascertained that approximation errors at all FPA temperatures decrease with the order of the polynomial as shown in Fig. 3(a). Based on above analysis, we established the parametric FPN model with the fifth-order polynomial associated with each pixel response as shown in (5).

$$F_{i,j} = a_j V_{i,j}^5 + b_j V_{i,j}^4 + c_j V_{i,j}^3 + d_j V_{i,j}^2 + e_j V_{i,j}^1 + f_j \tag{5}$$

where $F_{i,j}$ and $V_{i,j}$ are FPN and pixel response of pixel (i,j), respectively. We determined each coefficient of our FPN model by analyzing the distribution of all measured data corresponding to each FPA temperature. Each coefficient generally shows a Gaussian distribution, but the coefficient values have an obviously different range as shown in Fig. 3(b). To improve the robustness against the diversity of stripe patterns and prevent over-fitting problems, we used a wide range of coefficient values when generating training data.

B. ARCHITECTURE

In order to extract the structural features of FPN in a single infrared image effectively, we propose the de-striping deep convolutional network based on the dual-branch structure that consists of the consecutive N-convolution filter-branch and combined filter-branch. Also, we employ a residual learning strategy with the skip connection not only to prevent the loss of detailed information but also to reconstruct the resulting image as close to the ground truth image as possible. In Fig. 4 we introduce the proposed dual-branch structured de-striping convolution network (DBS-DCN).

The receptive field of convolutional neural networks is the region of input feature map that affects the output features of next layer. In [17], they mentioned the concept of *effective receptive field* that pixels in the receptive field cannot equally affect the output features. In particular, they noticed that the boundary region of the receptive field could not provide a sufficient influence, compared with the central area of the receptive field. Therefore, to make a large enough receptive field, we reduced the feature map size to a third by using the stride convolution layer and series of convolution layers in the consecutive N-convolution filter-branch. After that, we attempted to extract global features in the forward-learning phase. All activation function were placed behind convolution layer to increase a nonlinearity and prevent gradient vanishing in the backward process. In our proposed, we employed the Leaky ReLU (rectified linear unit) as the activation function.

In [32], the inception module performs a few different sized convolutional operations in parallel to extract features comprehensively, and then combines features to improve learning efficiency. Inspired by the ensemble concept of inception

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access

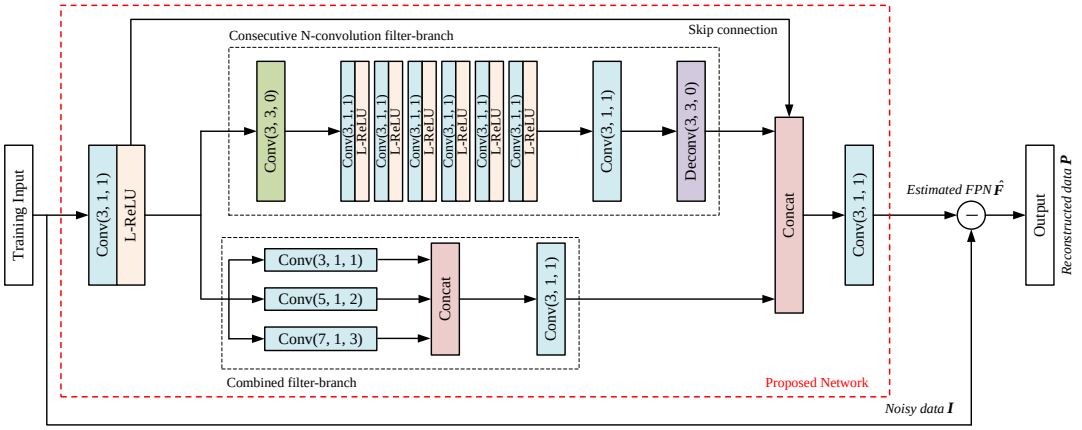


FIGURE 4. **Overall proposed network architecture. Conv and Deconv denote the convolution layer and the deconvolution layer, respectively. The numbers in parentheses indicate kernel size, stride, and padding, respectively. L-ReLU stands for the Leaky ReLU and Concat means the concatenation layer. We employed the transposed convolution as the deconvolution layer.**

module, we implemented the three-sized combined filter to extract various structural features of FPN in the spatial domain. The kernel sizes of the combined filter are 3 x 3, 5 x 5, and 7 x 7. In the combined filter-branch, we maintained the feature map size the same as the input size.

In the FPN removal, noisy observed image and noise-free reconstructed image are highly correlated. In addition, our training aims to seek the estimated FPN, not the predicted noise-free image from noisy input image. To this end, we designed our network to find the residual mapping relationship using the residual learning strategy with the skip connection structure. Through the training of network, we can obtain the reconstructed FPN-free infrared image by subtracting the estimated FPN from noisy input image. We attempted to minimize the training loss as much as possible, and for this reason, we applied the L_1-based distance function as our loss. It is known that the L_1 loss outperforms the L_2 loss in the fields of image restoration [21], [22].

$$\begin{aligned} L &= |\, P - V \,|_1 \\ &= |\, (I - \hat{F}) - V \,|_1 \\ &= |\, (V + F - \hat{F}) - V \,|_1 \\ &= |\, F - \hat{F} \,|_1 \end{aligned} \tag{6}$$

In this paper, we have trained our proposed DBS-DCN by optimizing using Adam [20] with modified L_1 loss [21] in 256-sized mini-batch. The weights, which are training parameters, are initialized by using the 'He' method [23].The leaky value and weight decay are 0.2, 0.0001, respectively. In addition, we set N to 6 for the consecutive convolution filter. The initial learning rate set to 0.0001 and is reduced by a factor of 10 for every 40 epochs.

IV. EXPERIMENTS

A. TRAINING DATA

For the training that can achieve generalized performance, we collected noise-free ground truth images from several public infrared datasets [13], [24]-[29] and we captured some clean infrared images using our imaging system. Fig. 5 shows some examples used in training data.

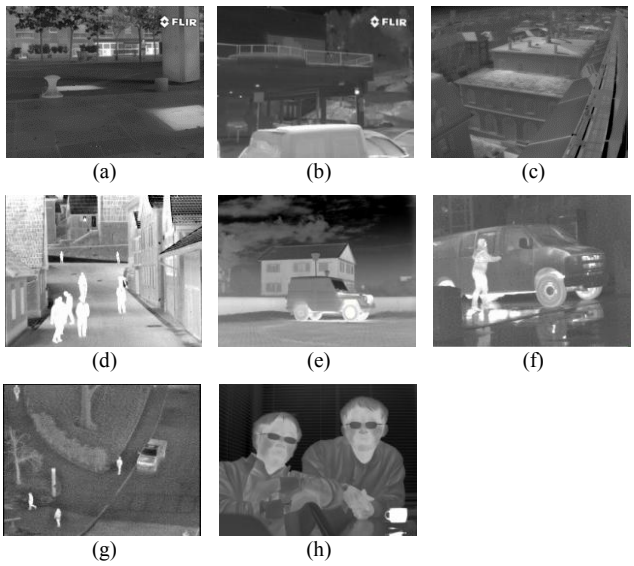


FIGURE 5. **Noise-free image samples (a)-(b) CVC-15 DB [25], (c) ASL DB [26], (d) LTIR DB [28], (e) TNO DB [24], (f) Morris DB [13], (g) OSU DB [27], (h) Our DB.**

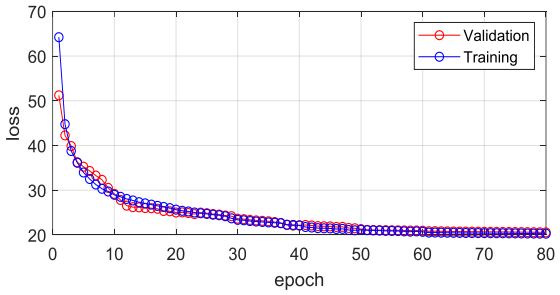


FIGURE 6. **Training and validation results.**

We transformed the 130 ground truth infrared images into a patch image with a size of 54 x 54 by cropping with stride 100. Then, we increased the number of patch up to about 4,500 by

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access

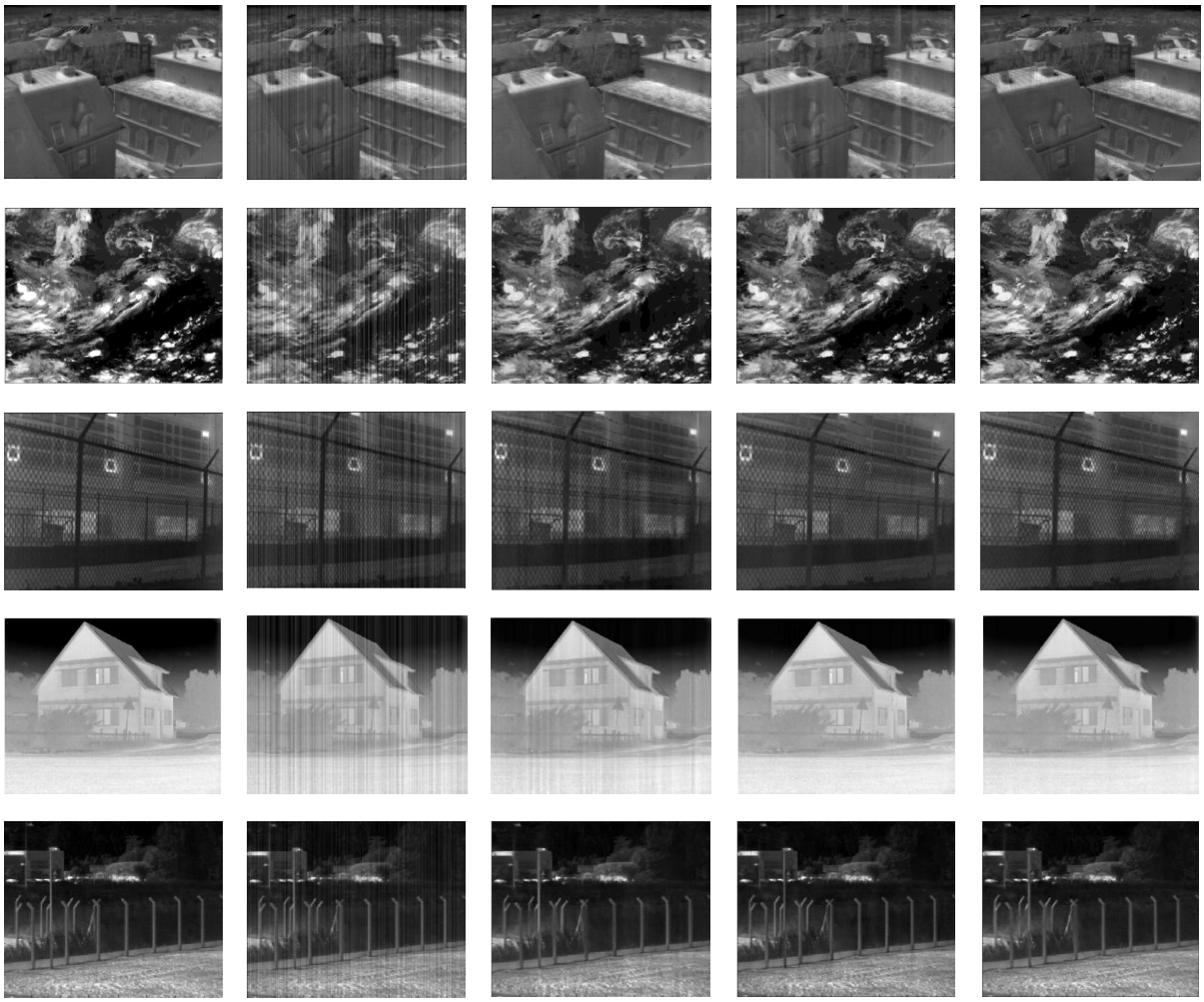


FIGURE 7. Comparative FPN de-striping results using each their own FPN model. In order from the first row, ASL DB [26], DB [8], Morris DB [13], TNO DB [24], and Our DB (a) Ground truth image, (b) Corrupted synthetic image, (c) DLS-NUC [4], (d) DMRN [7], (e) Proposed.

applying data augmentation [30] such as rotating 90 degrees and up-and-down flipping. We generated the training data by applying various scaled FPN to patches. Through this process, the total number of patches has been expanded to about 18,000. In our experiments, we employed 15,360 training images and 2,560 validation data. Fig. 6 shows the convergence results of training and validation loss. The total number of epoch is 80, and the result shows that loss is converged around 60 epoch without over-fitting problem.

B. COMPARATIVE EXPERIMENTS USING SYNTHETIC INFRARED IMAGES

We conducted comparative experiments with state-of-the-art stripe FPN removal methods [4], [7], [11] to evaluate the effectiveness of our proposed method. Using synthetic noisy images, we evaluated the robustness of our proposed method against FPN diversity and noise strength. For a convincing comparative evaluation, we employed 50 ground truth images that not used for training data, and we generated artificially corrupted images using three FPN models. A cubic model of DLS-NUC with randomly assigned coefficients [4], a random distribution model of DMRN [7], and our parametric model. Besides, all FPN simulations were followed as mentioned in their papers.

In Fig. 7, we demonstrate the de-striping performance for the experiments applying their own FPN model. Overall, most of the methods effectively removed the stripe FPN against synthetic infrared images. However, DLS-NUC contained the residual FPN in some test images despite using their FPN model used in both training and validation phases.

To evaluate the robustness against different stripe patterns, we compared de-striping performance of each method using synthetic noisy images generated by three different FPN models (*i.e.,* cubic model, random distribution model, and our fifth-order polynomial parametric model). As you can see, DLS-NUC produced an unexpected vertical artifact (Fig. 8(b)) and obvious residual FPN (Fig. 8(c)) in result images. DMRN also was not able to remove FPN completely as shown in Fig. 9(a), 9(b). On the other hand, our proposed DBS-DCN completely removed most of FPN without residual FPN and artifacts in the reconstructed image.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access

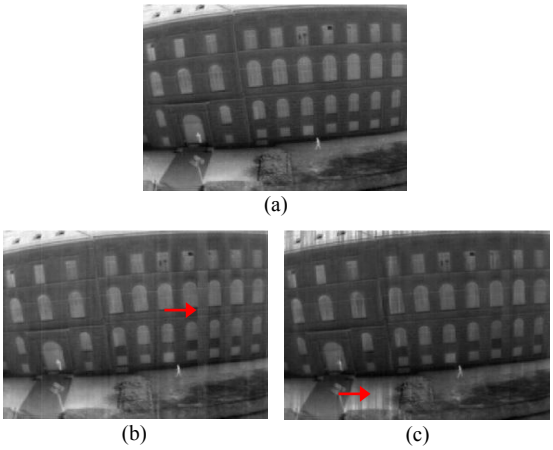


FIGURE 8. De-striping results of DLS-NUC. (a) Ground truth, (b) Result for the random distribution FPN, (c) Result for the parametric FPN.

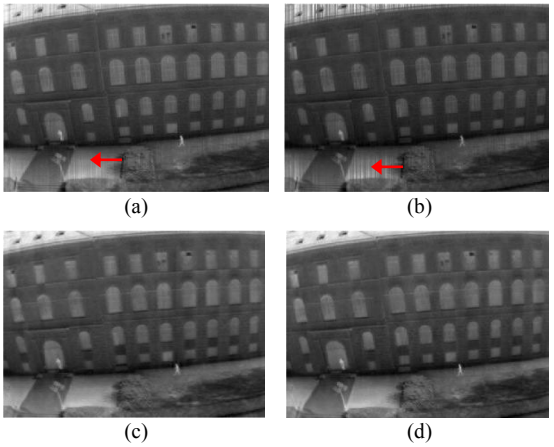


FIGURE 9. De-striping results comparison. The first and second rows are DMRN and proposed results, respectively. (a) Result for the cubic FPN, (b) Result for the parametric FPN, (c) Result for the cubic FPN, and (d) Result for the random distribution FPN.

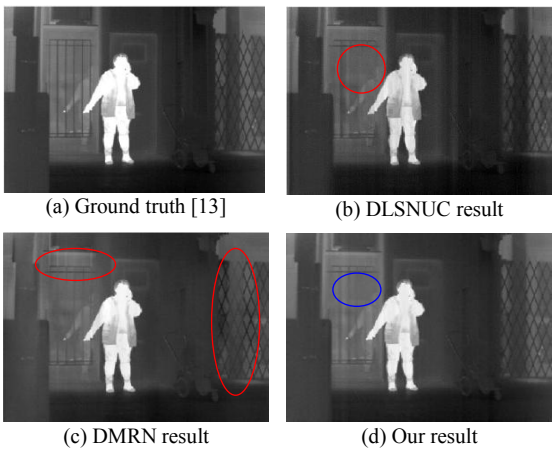


FIGURE 10. Qualitative assessment against information preservation.

Fig. 10 shows qualitative results against the preservation of vertical directional information, which are similar to FPN. DLSNUC and DMRN appeared unexpected artifacts on the left side and the blurred region on the right side in the FPN suppression process as shown in red in Fig. 10. In our result, the proposed method demonstrated that some vertical patterns were blurred as marked in blue (Fig. 10(d)).

For the quantitative comparison against synthetic image data, we have performed a quality assessment using two kinds of representative full-reference quality measures. Peak signal-to-noise ratio (PSNR), structural similarity (SSIM). Table 1 demonstrates the quantitative assessment of the comparative methods on two quality indices. In experiments using each FPN model, the best results pointed in bold letters.

TABLE 1. Average SSIM and PSNR for 50 synthetic infrared data.

	DLS-NUC [4]	DMRN [7]	Proposed
	SSIM / PSNR	SSIM / PSNR	SSIM / PSNR
FPN in [4]	0.945 / 34.646	0.922 / 33.377	**0.958** / **35.807**
FPN in [7]	0.905 / 31.504	**0.981** / **38.465**	0.945 / 35.248
Our FPN	0.924 / 32.988	0.918 / 31.372	**0.962** / **39.284**
Average	0.925 / 33.046	0.940 / 34.738	**0.955** / **36.780**

As you can see in Fig. 7 and 8, DLS-NUC could not completely suppress FPN from synthetic images, and these results lead to a relatively low assessment index for all FPN models. On the other hand, DMRN achieved the highest results using its own FPN model on both indices. However, experiments using the other two models showed significant degradation of the performance indices. In contrast to the comparative methods, our DBS-DCN achieved consistently notable results across all the FPN models in both evaluation indices. This is worth noting that our method accomplished robustness against the various stripe patterns whereas the other two comparative methods are highly dependent on the FPN model.

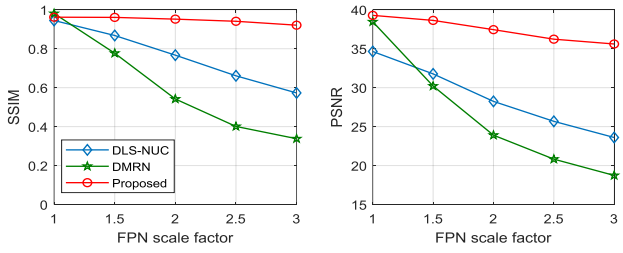
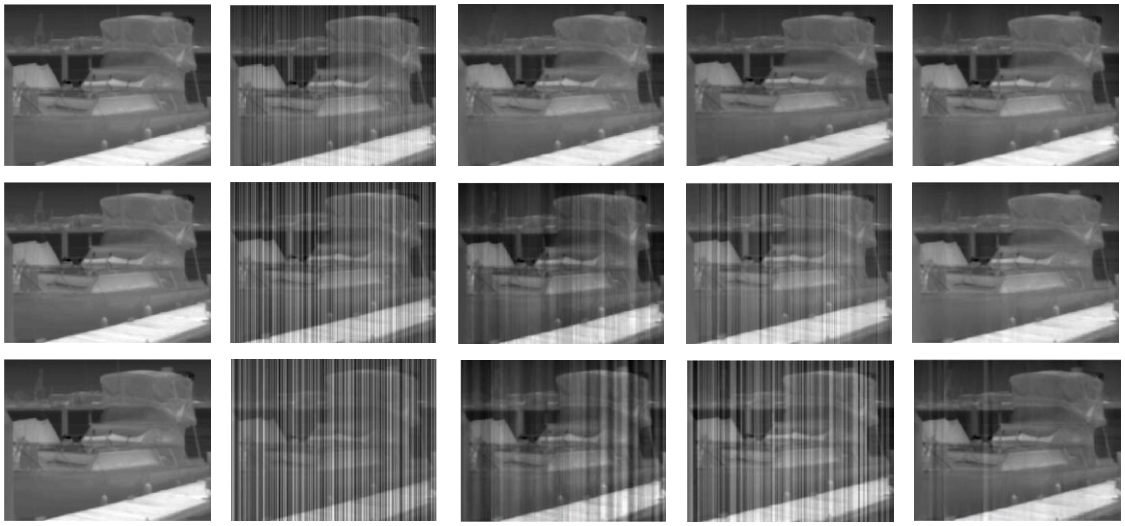


FIGURE 11. De-striping results comparison for the FPN scale. The left and right subfigures are average SSIM and PSNR, respectively.

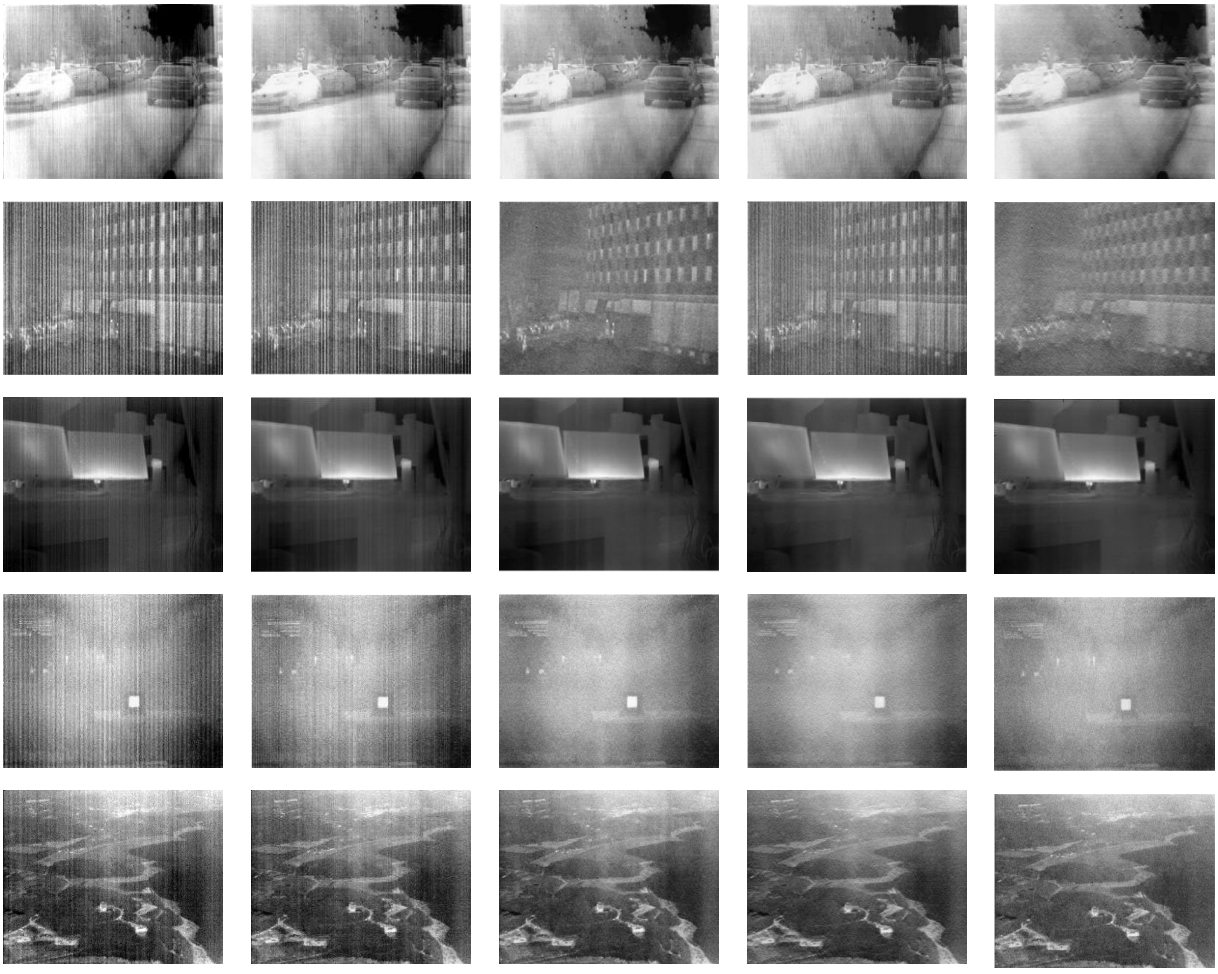
To evaluate the de-striping ability against the FPN scale, we generated the corrupted synthetic infrared data by increasing the scale of FPN up to three times. As seen in fig. 11, we demonstrated the de-striping results for the FPN scales as the average SSIM and PSNR. DLS-NUC and DMRN generated significant residual FPN. Besides, these two methods showed that the assessment indices are significantly degraded when the FPN scale is increased. On the other hand, our proposed method achieved stable de-striping performance against the pattern noise strength. Fig. 12 shows experimental results for the FPN scale 1, 2, and 3.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access



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

FIGURE 12. **Comparative FPN de-striping results against the FPN scales. In order from the first row, the results using scales 1, 2, and 3. (a) Ground truth image, (b) Corrupted image by the corresponding scale factor, (c) DLS-NUC, (d) DMRN [7], (e) Proposed.**



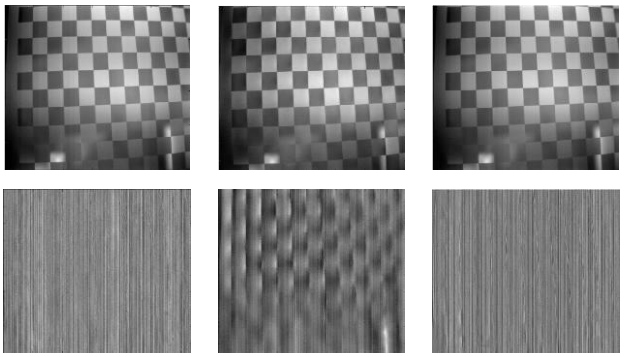
(a) (b) (c) (d) (e)

FIGURE 13. **Comparative experimental results using the corrupted real infrared data (a) Original noisy image, (b) SNRCNN [11], (c) DLS-NUC [4], (d) DMRN [7], and (e) Our proposed.**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access

C. COMPARATIVE EXPERIMENTS USING THE REAL INFRARED IMAGES

For the de-striping comparison using corrupted real images, we collected 32 images from publicly released real infrared data [4], [5] and we acquired 20 infrared raw images using our infrared imaging system. Fig. 13 shows qualitative assessment results of the four comparative methods using corrupted real infrared images. The first two rows are outdoor images in [5] and third-row data is indoor data from [4]. The last two rows are indoor and outdoor images of our data. As can be seen in Fig. 13(b), SNRCNN [11] contains significant residual stripe patterns. Compared with the relatively shallow network such as SNRCNN, three methods based on the deep convolutional network achieved better results in real infrared image data. In particular, DLSNUC achieved a remarkable de-striping result in corrupted real infrared test images and showed notable performance comparable to our proposed method. Although DMRN suppressed most of FPN effectively, the reconstructed image contains a significant amount of residual FPN as shown in the second row of Fig. 13(d).



(a) DLSNUC [4] (b) DMRN [7] (c) Proposed

FIGURE 14. Qualitative assessment against image details preservation. The first row shows the reconstructed noise-free image. The second row demonstrates the estimated FPN, and DMRN includes some checkerboard pattern corresponding to image details.

Overall, comparative methods suppressed the stripe pattern effectively. However, as shown in the second row of Fig. 14, three estimated FPN obtained from the training result indicate different performances. DMRN estimated substantial image details as FPN, and this result led to significant information loss in image. On the other hand, our proposed DBS-DCN and DLSNUC removed FPN comprehensively, and the estimated FPN excluded almost all image details.

To compare the quantitative performance of our proposed method against the comparative FPN removal methods, we employed three kinds of reference-free image quality metrics. Roughness [34], root-mean-square error of the horizontal adjacent pixel (RMSE-AP) [33], and average vertical gradient error (AVGE) [36]. The roughness index can measure the high-frequency components in both horizontal and vertical directions. Therefore, the roughness is commonly used to quantitatively comparison of the residual FPN in infrared images [4], [16]. Nevertheless, the way to evaluate how clear the FPN has been removed while the image details are well preserved is quite crucial work. Cao *et al*. [33] attempted to evaluate the preserving ability of image details by measuring the gradients between adjacent pixels in the horizontal direction, so that they employed the RMSE-AP index.

$$\mathrm{RMSE}_{\mathrm{AP}} = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n-1}(P(i,j)-P(i,j+1))^2}{m\times(n-1)}} \tag{7}$$

where P is the reconstructed image with m rows and n columns, $P(i, j)$ is a pixel signal in the i-th row and the j-th column. On the other hand, Zeng *et al*. [36] introduced the AVGE that calculates the difference of gradients between real corrupted image and its reconstructed noise-free image in the vertical direction. AVGE can assess the preserving ability on the vertical details of the image, so that we applied the AVGE as a complement to the RMSE-AP index from the perspective of quantitative evaluation of information loss.

$$\mathrm{AVEG} = \frac{1}{K}\sum_{k=1}^{K}\left|\,|\nabla_y(P_k)| - |\nabla_y(I_k)|\,\right| \tag{8}$$

where P_k and I_k are the reconstructed image and noisy image at pixel k, respectively. K is the total pixel number and ∇_y indicates the vertical gradient operator.

Table 2 shows the quantitative evaluation results of the comparative FPN de-striping methods using the average value of roughness, RMSE-AP, and AVGE. As you can see, our proposed method acquired notable results in all quality assessment metrics. Through comparative experiments for corrupted real data, we showed that our proposed method achieved the ability to suppress the real FPN and preserve the detail information of infrared images.

TABLE 2. Comparative results on corrupted 52 real infrared data.

Methods	Roughness	RMSE_AP	AVGE
Noisy	0.247	27.981	-
SNRCNN	0.201	22.946	0.059
DLS-NUC	0.116	9.176	0.019
DMRN	0.128	14.775	0.030
Proposed	**0.113**	**8.744**	**0.012**

V. CONCLUSION

The FPN de-striping without image detail loss is quite hard work due to a relatively low-contrast infrared image property. To eliminate FPN and prevent information corruption, we proposed a new de-striping deep convolutional network based on the dual-branch structure and residual learning. In addition, we established a parametric FPN model through the diagnostic researches to generate effective training data. Compared to previously existing methods, our proposed DBS-DCN showed remarkable de-striping performance on both qualitative and quantitative evaluation for both real infrared data and corrupted synthetic image data. In a future study, we are going to expand

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access

our deep learning network for correcting different types of infrared image non-uniformity.

REFERENCES

[1] L. Shkedy et al., "Temperature dependence of spatial noise in InSb focal plane arrays." Infrared Detectors and Focal Plane Arrays VI. Vol. 4028. International Society for Optics and Photonics, 2000.
[2] X. Kuang et al., "Single Infrared Image Optical Noise Removal Using a Deep Convolutional Neural Network", IEEE photonics Journal, vol. 10, n. 2, April 2018.
[3] C. Minassian et al., "Uncooled amorphous silicon TEC-less 1/4 VGA IRFPA with 25 µm pixel-pitch for high volume applications." Infrared Technology and Applications XXXIV. Vol. 6940. International Society for Optics and Photonics, 2008.
[4] Z. He, Y. Cao, Y. Dong, J. Yang, Y. Cao, and C.-L. Tisse, "Single image-based non-uniformity correction of uncooled long-wave infrared detectors: A deep-learning approach," Appl. Opt., vol. 57, no. 18, pp. D155–D164, 2018.
[5] Y. Tendero, S.Landeau, J.Gilles, "Non-uniformity correction of infrared images by midway equalization." Image Processing On Line 2 (2012): 134-146.
[6] 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.
[7] Yi. Chang et al., "Infrared Aero-thermal Non-uniform Correction via Deep Multi-scale Residual Network", IEEE geoscience and remote sensing letters. Vol. 16, no. 7, July 2019.
[8] P. Xiao, Y. Guo, and P. Zhuang, "Removing stripe noise from infrared cloud images via deep convolutional networks," IEEE Photon. J., vol. 10, no. 4, Aug. 2018, Art. no. 7801114.
[9] Y. Choi, N. Kim, S. Hwang, and I. S. Kweon, "Thermal image enhancement using convolutional neural network," in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), Oct. 2016, pp. 223–230.
[10] K. Lee et al., "Brightness-based convolutional neural network for thermal image enhancement." IEEE Access 5 (2017): 26867-26879.
[11] 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.
[12] J. Guan, L. Rui, and X. Ai, "Wavelet deep neural network for stripe noise removal." IEEE Access 7 (2019): 44544-44554.
[13] N. J. Morris et al., "Statistics of infrared images," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), May/Jun. 2007, pp. 1–7.
[14] T. R. Goodall, A. C. Bovik, and Nicholas G. Paulter. "Tasking on natural statistics of infrared images." IEEE Transactions on Image Processing 25.1 (2015): 65-79.
[15] Y. Cao and Y. Li, "Strip non-uniformity correction in uncooled longwave infrared focal plane array based on noise source characterization," Opt. Commun. 339, 236–242 (2015).
[16] L. Geng, Q. Chen, and W. Qian. "An adjacent differential statistics method for IRFPA nonuniformity correction." IEEE Photonics Journal 5.6 (2013): 6801615-6801615.
[17] W. Luo, et al., "Understanding the effective receptive field in deep convolutional neural networks." Advances in neural information processing systems. 2016. p. 4898-4906.
[18] K. Zhang et al., "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising." IEEE Transactions on Image Processing 26.7 (2017): 3142-3155.
[19] 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.
[20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. ICLR, 2015, pp. 1–15.
[21] 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.
[22] 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).
[23] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in Proc. IEEE Int. Conf. Comput. Vis., Dec. 2015, pp. 1026–1034.
[24] A. Toet et al., "TNO Image fusion dataset." Figshare. data (2014).
[25] C. Aguilera, F. Barrera, F. Lumbreras, A. D. Sappa, and R. Toledo, "Multispectral image feature points," Sensors, vol. 12, no. 9, pp. 12661–12672, 2012.
[26] J. Portmann, S. Lynen, M. Chli, and R. Siegwart, "People detection and tracking from aerial thermal views," in Proc. IEEE Int. Conf. Robot. Autom. (ICRA), Jun. 2014, pp. 1794–1800.
[27] J. Davis and M. Keck, "A two-stage template approach to person detection in thermal imagery." 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05)-Volume 1. Vol. 1. IEEE, 2005.
[28] A. Berg, J. Ahlberg, and M. Felsberg, "A thermal object tracking benchmark," in Proc. 12th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS), Aug. 2015, pp. 1–6.
[29] M. M. Zhang et al., "VAIS: A Dataset for Recognizing Maritime Imagery in the Visible and Infrared Spectrums." in Proc. IEEE Conf. Comput. Vis. Pattern Recognit Workshops. 2015, pp. 10-16.
[30] C. Shorten and T. M. Khoshgoftaar. "A survey on image data augmentation for deep learning." Journal of Big Data 6.1 (2019): 60.
[31] 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).
[32] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015.
[33] Y. Cao et al., "Spatially adaptive column fixed-pattern noise correction in infrared imaging system using 1D horizontal differential statistics." IEEE photonics journal 9.5 (2017): 1-13.
[34] M. M. Hayat, S. N. Torres, E. Armstrong, S. C. Cain, and Brian Yasuda, "Statistical algorithm for nonuniformity correction in focal-plane arrays," Appl. Opt. 38, 772-780 (1999)
[35] J. Guan et al., "Fixed pattern noise reduction for infrared images based on cascade residual attention CNN." Neurocomputing 377 (2020): 301-313.
[36] Q. Zeng, et al. "Single infrared image-based stripe nonuniformity correction via a two-stage filtering method." Sensors 18.12 (2018): 4299.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019057, IEEE Access

JONGHO. LEE received the B.S. and M.S. degrees in electrical engineering from Kwangwoon University, Seoul, Korea, in 2000 and 2002, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea. His research interest includes infrared image analysis, image quality assessment, and infrared image restoration using deep learning.

YONGMAN. RO (S'85–M'92–SM'98) received the B.S. degree from Yonsei University, Seoul, South Korea, and the M.S. and Ph.D. degrees from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea. He was a Researcher at Columbia University, a Visiting Researcher at the University of California at Irvine, Irvine, CA, USA, and a Research Fellow at the University of California at Berkeley, CA, USA. He was a Visiting Professor with the Department of Electrical and Computer Engineering, University of Toronto, Canada. He established Image and Video Systems Lab, KAIST, in 1997. He is currently a Professor with the Department of Electrical Engineering, KAIST. During the years, he has been conducting research in a wide spectrum of image and video systems research topics. His recent research interests are deep learning, machine learning in computer vision and image processing (2D, 3D, and VR), medical imaging, visual recognition, and visual quality assessment.

He has received the Young Investigator Finalist Award from ISMRM in 1992 and the Young Scientist Award of the year, South Korea, in 2003. He has served as a TPC member of many international conferences including as the program chair, and has organized special sessions. He has served as an Associate Editor for IEEE SIGNAL PROCESSING LETTERS. He currently serves as an Associate Editor for Transitions on Data Hiding and Multimedia Security (Springer–Verlag).