This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

SAR Image Despeckling Employing a Recursive Deep CNN Prior

Huanfeng Shen, *Senior Member, IEEE*, Chenxia Zhou, Jie Li, *Member, IEEE*, and Qiangqiang Yuan, *Member, IEEE*

***Abstract*— Synthetic aperture radar (SAR) images are inherently affected by speckle noise, for which deep learning-based methods have shown good potential. However, the deep learning-based methods proposed until now directly map low-quality images to high-quality images, and they are unable to characterize the priors for all the kinds of speckle images. The variational method is a classic model optimization approach that establishes the relationship between the clean and noisy images from the perspective of a probability distribution. Therefore, in this article, we propose the recursive deep convolutional neural network (CNN) prior model for SAR image despeckling (SAR-RDCP). First, the data-fitting term and regularization term of the SAR variational model are decoupled into two subproblems, i.e., a data-fitting block and a deep CNN prior block. The gradient descent algorithm is then used to solve the data-fitting block, and a predenoising residual channel attention network based on dilated convolution is used for the deep CNN prior block, which combines an end-to-end iterative optimization training. In the experiments undertaken in this study, the proposed model was compared with several state-of-the-art despeckling methods, obtaining better results in both the quantitative and qualitative evaluations.**

***Index Terms*— Convolutional neural network (CNN), despeckling gain (DG) loss, residual channel attention, synthetic aperture radar (SAR) image despeckling, variation.**

I. Introduction

SYNTHETIC aperture radar (SAR) is an active remote sensing system, which has the advantages of all day and night and all-weather ability. It plays an important role in the earth observation field and provides source data for terrain classification, target detection, physical parameter extraction, and scene interpretation [1], [2]. Nevertheless, because of the coherent interference of radar waves reflected from many basic scatterers, the SAR images are inherently affected by speckle noise [3]. This has a great impact on the subsequent application accuracy. Therefore, despeckling for SAR images is essential to improve the quality before the application of SAR images. In the past few decades, many researchers have attempted to solve this ill-posed inverse problem, and many different methods have been proposed.

Generally, the existing despeckling methods fall into three broad categories: the filtering methods, the variational methods, and the deep learning-based methods. The early filtering works on despeckling were Bayesian methods in the spatial domain. One of the classical strategies was minimizing the root-mean-square error (MSE) between the real and estimated values based on the local statistics properties, such as the Lee filter [4], Kuan filter [5], Frost filter [6], and so on. For Bayesian methods in the transform domain, most filters exploit the discrete wavelet transform and have superior performance over conventional spatial filters, which are represented by homomorphic or nonhomomorphic filtering in the wavelet domain [7]–[9]. However, recently, the most popular methods do not follow a Bayesian method [10]–[17]. The sigma filter [8] is a conceptually simple noise smoothing algorithm, which averages the selected similar pixels in the local window with the center pixel according to the sigma probability of the speckle probability density function. The nonlocal filter is one of the most promising algorithms, which further demonstrates the importance of similar pixel selection [11]–[13]. Compared to the sigma filter, it defined the similar pixels and weights by measuring the similarity between a local patch centered at the reference pixel and the other local patch centered at a given nonlocal neighborhood pixel, such as the probabilistic patch-based filter (PPB) [11] and SAR block-matching 3-D (SAR-BM3D) [13] are excellent methods.

The variational method is another popular denoising method [18]–[20]. The difference with the filtering method is that it establishes and solves the energy functional between the real value and the observed value to solve the ill-posed inverse problem. A detailed introduction of the methods is given by Aubert and Aujol [18] (the AA model), Shi and Osher [19] (the SO model), and Ma *et al.* [20] (the adaptive nonlocal functional (ANLF) model). These variational methods have a high degree of stability and flexibility by adjusting the parameters of different images. However, the despeckling effects of the AA, SO, and ANLF models are different, mainly because of their different priors. The AA and SO models used a local total variation (TV)

Manuscript received November 14, 2019; revised January 8, 2020 and March 27, 2020; accepted May 2, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61671334 and Grant 41701400. *(Corresponding author: Jie Li.)*

Huanfeng Shen is with the School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China, and also with the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: shenhf@whu.edu.cn).

Chenxia Zhou is with the School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China (e-mail: zhoucx31@whu.edu.cn).

Jie Li is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China (e-mail: aaronleecool@whu.edu.cn).

Qiangqiang Yuan is with the School of Geodesy and Geomatics, Wuhan 430079, China, and also with the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: yqiang86@gmail.com).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2020.2993319

0196-2892 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

prior to the original or logarithmically transformed domain. The local TV prior assumes that the image is piecewise smooth, which leads to the staircase effect and oversmoothing in the restored image. The ANLF model introduces a nonlocal self-similarity prior, which achieves a certain balance between the original information maintenance and smoothness. However, since the existing SAR variational despeckling priors are estimated by a traditional approach, they are not accurate enough to describe the nonlinear characteristics of the SAR speckle, resulting in detailed loss. Therefore, to obtain better *a priori* knowledge, the deep mining of the relationship between the clean image and the contaminated SAR image is potential and promising.

With the in-depth study of deep learning in the field of image restoration, the deep learning-based methods for SAR image despeckling are also emerging. First, Chierchia *et al.* [21] proposed a homomorphic transform method, using logarithmic and exponential operations, to transform a multiplicative noise into an additive noise, and used a 17-layer fully convolutional residual network to learn the data characteristics (SAR-CNN). Wang *et al.* [22] proposed to use a componentwise division-residual layer with skip connection to estimate the denoised image in the network (IDCNN), which uses the combination of European loss and TV loss for an end-to-end training; and Zhang *et al.* [23] used a seven-layer lightweight network based on a dilated residual network (DRN) for the speckle removal of SAR images. The multilayer perceptron neural network [24] and a CNN combined with guided filtering [25] have also been proposed. These methods have a strong nonlinear fitting ability and fast speed. Nevertheless, the data-driven deep learning methods have difficulty in fully simulating all the noise conditions, which limits the ability of the noise suppression.

To address the shortcomings of the deep learning-based methods, a model-driven framework, constructing the relationship between the noisy and clean images from the perspective of a probability distribution, is one possibility. Thus, it would be interesting to investigate the integration of data-driven and model-driven methods, to leverage their respective merits and obtain a more stable, flexible, and efficient method. Therefore, in this article, the recursive deep CNN prior model for SAR image despeckling (SAR-RDCP) is proposed. The main contributions are summarized as follows:

1) A novel SAR image despeckling model employing an RDCP is proposed. The whole model can be divided into two recursive subblocks: the data-fitting (fidelity) block and the deep CNN prior block, which perform joint end-to-end training with a novel despeckling gain (DG) loss. The model adopts a joint iterative optimization, which is equivalent to a recursive neural network. We also analyze and summarize the optimal strategy for real data with different noise levels.
2) For the deep CNN prior block, a predenoising residual channel attention network (PRCA) based on dilated convolution is used. The before–after symmetric multiscale dilated convolutional unit is used to extract and reconstruct the multiscale information. In addition, the feature denoising block (FDB) and residual channel attention block (RCAB) are added to enhance the accuracy of information extraction.

The rest of this article is organized as follows. Section II introduces the AA variational model for SAR images and provides an analysis of the deep learning prior. In Section III, the proposed model and the related details are presented. Then, in Section IV, we describe the experiments conducted on both simulated and real SAR images to validate the performance of the proposed method. Finally, our conclusion is provided in Section V.

II. Related Work

A. AA Variational Model for SAR Images

The SAR data degradation model can be described as $f = uv$, where f and u are the observed image and noise-free image, respectively, and v is the multiplicative noise, which follows a gamma law with mean $= 1$. Its density function [18] is defined as

$$P(v) = \frac{L^L}{\Gamma(L)} v^{L-1} e^{-Lv} 1_{\{v \geq 0\}} \tag{1}$$

where L denotes the number of looks.

In order to obtain the optimal value of u, according to the maximum *a posteriori* (MAP) estimator, $P(u|f)$ needs to be maximized. Based on Bayes' rule, we have $P(u|f) = P(f|u)P(u)/P(f)$. Maximizing $P(u|f)$ is equivalent to minimizing the log-likelihood, which can be written as

$$\begin{aligned} u &= \arg\min\{-\log P(u|f)\}_u \\ &= \arg\min\{-\log P(f|u) - \log P(u)\} \end{aligned} \tag{2}$$

where $\log P(f|u)$ and $\log P(u)$ denote the data fidelity and prior terms, respectively. $P(u)$ is the *a priori* knowledge of u, and $P(f|u)$ can be formulated as [18]

$$P(f|u) = \frac{L^L}{u^L \Gamma(L)} f^L e^{-\frac{Lf}{u}}. \tag{3}$$

Therefore, (2) can be rewritten as

$$u = \arg\min\left\{\lambda\left(\log u + \frac{f}{u}\right) + \Theta(u)\right\} \tag{4}$$

where λ is the regularization parameter, $(\log u + (f/u))$ is the fidelity item, and $\Theta(u)$ is the prior term.

B. Analysis of the Deep CNN Prior

The optimization of deep learning is characterized by a strong nonlinear fitting ability but has an excessive reliance on the data. The variational method has a strong mathematical basis and can be applied to different kinds of situations by adjusting the parameters and different priors. The prior term of the variational method plays a key role in image restoration performance. Embedding a deep CNN as a prior to the variational method has been used to tackle some inverse problems. For example, in natural image restoration, a number of successful models have been proposed. In [26], the deep learning prior was first plugged into the optimization procedure of a Gaussian white noise variational model. The variable splitting technique and half-quadratic splitting (HQS)

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

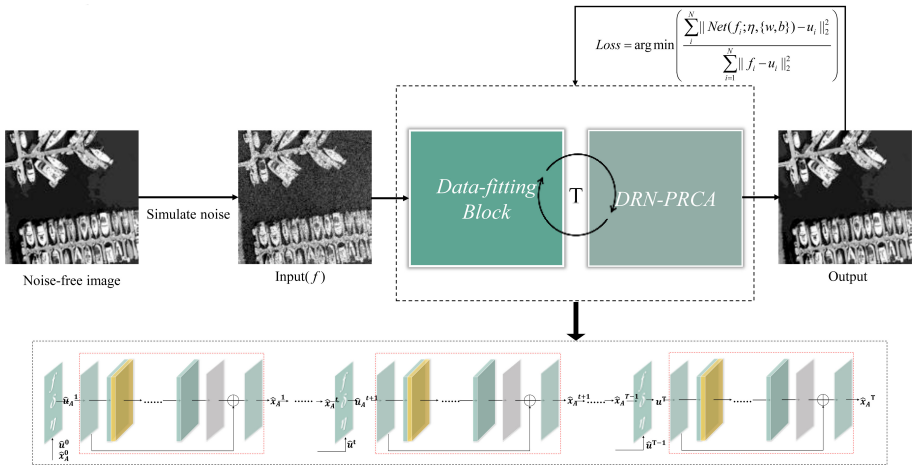

Fig. 1. Overall framework of the proposed SAR-RDCP method.

method are then usually adopted to decouple the optimization model. In this model, a pretrained deep learning model is used, which can then be plugged into different image restoration tasks. For the super-resolution and image enhancement tasks [27]–[29], Dong *et al.* [27] proposed a feed-forward recursive deep neural network, wherein the denoisers and the back-projection modules can be jointly optimized. For optical remote sensing image fusion, Shen *et al.* [30] combined a deep neural network with the variational model, wherein the difference is that the network outputs the gradient information of the high-resolution images. A unidirectional variation model and a deep CNN denoiser prior model have also been integrated for simultaneous destriping and denoising [31]. All the above-mentioned methods have shown that deep CNN priors are suitable for different image restoration tasks and have obtained positive results with optical images [26].

Recently, some deep CNN methods have played a certain role in SAR despeckling [21]–[25]. The common idea behind these methods is to directly train an end-to-end network, wherein the main difference lies in the network structure. Meanwhile, it also proves the validity of deep CNN in SAR despeckling. Therefore, based on the deep CNN idea with model-driven thought, it is reasonable and possible to apply a deep CNN prior to SAR image despeckling. In this article, an RDCP for SAR despeckling is proposed.

III. Methodology

A. Framework of SAR-RDCP

In SAR despeckling, because of the limitations of the prior term in the traditional variational methods and the training data in the deep learning methods, neither of these two approaches can maintain a good balance between detail preservation and noise reduction. Differing from the deep learning methods that completely rely on the data and the traditional variational methods that need to achieve a good effect by adjusting parameters, we propose SAR-RDCP. The proposed method takes into account the distribution characteristics of the SAR speckle to tap into the potential of the deep CNN prior. The overall framework of the proposed method is shown in Fig. 1. SAR-RDCP learns the nonlinear end-to-end mapping between the noisy data and the label to jointly optimize the data-fitting block and the deep CNN prior. Details of the joint optimization are provided below.

The deep CNN prior needs to be plugged into the variational model shown in (4). The first step here is to decouple the fidelity term and regularization term of (4), so that the auxiliary variable x is introduced, and (4) can be rewritten as

$$u = \arg\min\left\{\lambda\left(\log u + \frac{f}{u}\right) + \Theta(x)\right\}, \quad \text{s.t. } x = u \tag{5}$$

where $\Theta(x)$ is the deep CNN prior term. Then, in order to solve (5), a penalty term is introduced to transform the constrained optimization problem into a nonconstrained optimization problem. The HQS method is used to alternately solve the two variables, which are given by

$$(u, x) = \arg\min\left\{\lambda\left(\log u + \frac{f}{u}\right) + \gamma \|u - x\|_2^2 + \Theta(x)\right\} \tag{6}$$

$$\begin{cases} \hat{u}^t = \arg\min\left\{\lambda\left(\log u + \dfrac{f}{u}\right) + \dfrac{\gamma}{2}\left\|u - \hat{x}^{t-1}\right\|_2^2\right\} \\ \hat{x}^t = \arg\min\left\{\dfrac{\gamma}{2}\left\|\hat{u}^t - x\right\|_2^2 + \Phi(x)\right\} \end{cases} \tag{7}$$

where γ is the penalty coefficient and t is the iteration number.

Note that (7) alternates the minimization problems with respect to the data-fitting term (u) and deep CNN prior term (x). To solve the iterative optimization problem, a recursive network structure is used. The data-fitting term and deep CNN prior term constitute the two recursive units of the proposed method. Because the AA model is developed for the SAR intensity image, for the proposed SAR-RDCP model, the intensity image is used in the data-fitting term. However, the amplitude image is used in the deep CNN term with a PRCA because the nonlinear relationship is better fitted by using the amplitude image owing to the smaller difference of the clean image and the speckled image, which makes it easier to train. In addition, the speckle strength of the amplitude image is lower than the intensity image on the visual effects. Therefore, it should be converted into the following

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

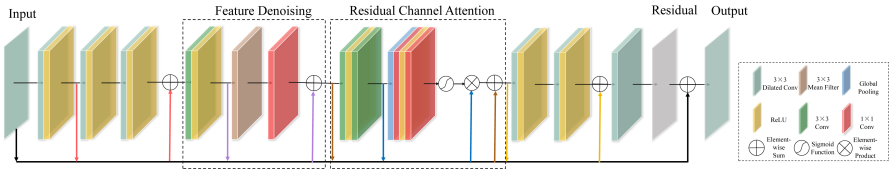

Fig. 2. Architecture of the proposed deep residual gradient CNN network.

equation:

$$\begin{cases} \hat{u}^t = \arg\min\left\{\lambda\left(\log u + \dfrac{f}{u}\right) + \dfrac{\gamma}{2}\left\|u - \left(\hat{x}_{\mathrm{A}}^{t-1}\right)^2\right\|_2^2\right\} \\ \hat{u}_{\mathrm{A}}^t = \mathrm{sqrt}\left(\hat{u}^t\right) \\ \hat{x}_{\mathrm{A}}^t = \mathrm{PRCA}\left(\hat{u}_{\mathrm{A}}^t\right) \end{cases} \tag{8}$$

where $\hat{u}_{\mathrm{A}}$ and $\hat{x}_{\mathrm{A}}$ represent the format of amplitude and are the sqrt of $\hat{u}$ and $\hat{x}$.

The recursive structure is unfolded as shown in Fig. 1 and Algorithm 1 shows the implementation steps, wherein the input of each recursion is the previous output. The reason for using a recursive structure is that the number of network parameters can be reduced while solving the problem of iterative optimization. Differing from [27], the positions of the two recursive units can be alternated. In the SAR-RDCP framework, the data-fitting block is the input unit and the deep CNN prior is the output unit. Then, for the subproblem of the data-fitting block (u), the traditional method is used, while for the subproblem of the deep CNN prior term (x_{A}), the PRCA is used to fit its nonlinear relationship, as shown in Fig. 2. The specific solutions for the two parts are given as follows:

B. Data-Fitting Block Subproblem

For the data-fitting block, we focus on the first subproblem of (7). In order to simplify the subproblem, parameters λ and γ are unified as one parameter η

$$\hat{u}^t = \arg\min\left\{\eta\left(\log u + \frac{f}{u}\right) + \frac{1}{2}\left\|u - \left(\hat{x}_{\mathrm{A}}^{t-1}\right)^2\right\|_2^2\right\}. \tag{9}$$

For this part, the simple and efficient method of gradient descent is used to solve the subproblem, which can then be converted into the following equation:

$$\hat{u}^t = \hat{u}^{t-1} - \delta\left(\eta\left(\frac{1}{\hat{u}^{t-1}} - \frac{f}{\left(\hat{u}^{t-1}\right)^2}\right) + \left(\hat{u}^{t-1} - \left(\hat{x}_{\mathrm{A}}^{t-1}\right)^2\right)\right) \tag{10}$$

where δ is the step length of the updating. Because it is embedded in the whole framework as a training block, the parameters of this block can be adjusted with the training of the network. In this article, η is set as a trainable parameter and δ as a fixed value.

C. PRCA Block Subproblem

For the deep CNN prior block, the second subproblem of (8) can be regarded as

$$x_{\mathrm{A}}^{t+1} = \mathrm{PRCA}\left(u_{\mathrm{A}}^{t+1}\right) \tag{11}$$

where PRCA refers to the predenoising residual channel attention network based on dilated convolution. The overall architecture of the PRCA is shown in Fig. 2. It consists of nine layers with four different blocks, i.e., "dilated convolution + rectified linear unit (ReLU)" in the first layers, "FDB" and "RCAB" in the middle layers, and "dilated convolution + ReLU" and "dilated convolution" in the last layers. A residual strategy and skip connection are also adopted, which can effectively help the network to converge faster and perform better in the training. The number of feature maps in each middle layer is set to 64. The detailed configuration of the PRCA block is provided in Table I.

1) Dilated Convolutional Layers: It is well known that contextual information facilitates the reconstruction of a degraded image. In addition, dilated convolution is an effective way to improve the receptive field size and capture fuller contextual

Algorithm 1 SAR Image Despeckling Employing a Recursive Deep CNN Prior (SAR-RDCP)

Train

Input: noisy amplitude image f_A
clean amplitude image u_A
Initialization:
(1) Set parameters T, δ, η *the parameters of PRCA*
(2) Initialize $\hat{u}$ as $\hat{u}^0 = f_A^2$, $\hat{x}_A$ as $\hat{x}_A^0 = f_A$
(3) Set $\hat{u}_A = sqrt(\hat{u})$, $f = f_A^2$
for t form 1 to T
$\hat{u}^t \leftarrow \hat{u}^{t-1} - \delta(\eta(\frac{1}{\hat{u}^{t-1}} - \frac{f}{(\hat{u}^{t-1})^2}) + (\hat{u}^t - (\hat{x}_A^{t-1})^2))$
$\hat{u}_A^t \leftarrow sqrt(\hat{u}^t)$
$\hat{x}_A^t \leftarrow PRCA(\hat{u}_A^t)$
return $\hat{x}_A \leftarrow \hat{x}_A^T$
$loss \leftarrow DG(\hat{x}_A, u_A, f_A)$
$loss.back()$ *update the parameters of PRCA and* η
Output: the SAR-RDCP model

Test

Input: test amplitude image f_A
Initialization: $\hat{u}$ as $\hat{u}^0 = f_A^2$, $\hat{x}_A$ as $\hat{x}_A^0 = f_A$
for t form 1 to T
$\hat{u}^t \leftarrow \hat{u}^{t-1} - \delta(\eta(\frac{1}{\hat{u}^{t-1}} - \frac{f}{(\hat{u}^{t-1})^2}) + (\hat{u}^t - (\hat{x}_A^{t-1})^2))$
$\hat{u}_A^t \leftarrow sqrt(\hat{u}^t)$
$\hat{x}_A^t \leftarrow PRCA(\hat{u}_A^t)$
return $\hat{x}_A \leftarrow \hat{x}_A^T$
Output: $\hat{x}_A$

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
DETAILED CONFIGURATION OF THE PRCA BLOCK

Main parts	Configuration	
First	64×3×3 1 Dilated Conv +ReLU	
	64×3×3 2 Dilated Conv +ReLU	
	64×3×3 3 Dilated Conv +ReLU	
Middle	Feature denoising	64×3×3 Conv +ReLU
		Mean filter (Average pool)
		64×1×1 Conv
	Residual channel attention	64×3×3 Conv +ReLU+Conv
		Global pooling
		64×1×1Conv +ReLU+Conv
Last	64×3×3 3 Dilated Conv +ReLU	
	64×3×3 2 Dilated Conv +ReLU	
	64×3×3 1 Dilated Conv	

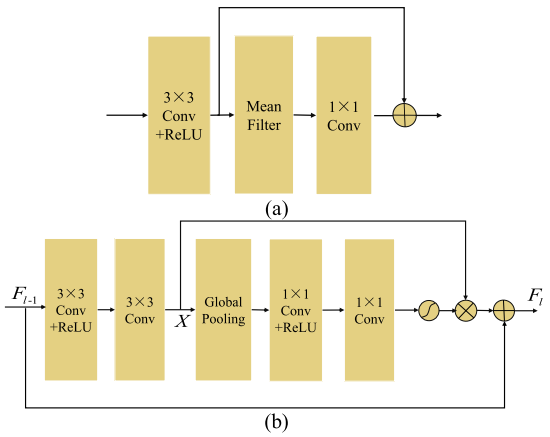

Fig. 3. (a) FDB. (b) RCAB.

information at the same parameter size. In PRCA, the symmetric dilated convolutional structure is used in the first and last three layers, as shown in Fig. 2. The first part is composed of three layers of dilated convolutions with ReLU nonlinear activation, and their dilation factors are 1, 2, and 3, respectively. It generates 64-channel feature maps with a filter size of 3×3. The dilated factors of the convolutional layers of the last part are 3, 2, and 1, respectively, and the ReLU nonlinearity is only used in the first two layers. The output image is reconstructed from the 64-channel feature maps with the 3×3 kernel.

2) Feature Denoising Block: The FDB can effectively remove the unnecessary "noise" in a feature map [32]. For noisy images, the noise can also be regarded as adding a kind of perturbation, which will also have an impact on the feature extraction, thus, generating feature maps with noise. Therefore, the FDB is added into the middle of the PRCA to correct the feature maps. The FDB wraps the denoising operation with a 1×1 convolution and skip connection, as shown in Fig. 3(a). For the denoising operation, the mean filter is used. Theoretically, the mean filter can smooth out details while removing noise, but it was proven in [32] that it can also improve the adversarial robustness, so the problem of the loss of some small details is not considered here. Moreover, the mean filter is simple and easy to implement (average pooling with a stride of 1) in the network, while not increasing the parameters and allowing fast training.

3) Residual Channel Attention Block: Images usually contain both high-frequency and low-frequency information. However, it is more difficult to reconstruct and recover the high-frequency information. In general, CNNs treat different information equally and lack some flexibility, which leads to them paying less attention to the high-frequency information. Recently, the attention mechanism has been used in image restoration, the key point of which is to distinguish between the different channel-wise and spatial-wise features [33]–[36]. In the process of SAR image despeckling, the recovery of high-frequency information, such as details, is one of the key points. In the proposed method, the RCAB [32] is used in the PRCA block, as shown in Fig. 3(b), for the details, which allows the network to attach importance to high-frequency channel-wise features. The RCAB is composed of a channel attention module and a residual block. If we let F_{l-1} and F_l be the input and output feature maps of the RCAB, then F_l can be formulated as

$$F_l = F_{l-1} + \mathrm{CA}(X) \cdot X \tag{12}$$

where CA denotes the channel attention function. In channel attention, the first thing to do is to transform the channel-wise global spatial information into the channel description factor by using global average pooling, and then use a gating mechanism (sigmoid gating) to capture the channel's statistical information from the aggregated information, and finally rescale the input features. X is mainly obtained by two stacked convolutional layers

$$X = W_l^2 \cdot \delta\left(W_l^1 \cdot F_{l-1} + b^1\right) + b^2 \tag{13}$$

where W and b are the weight and bias of the convolution, and δ is the ReLU activation function.

D. Overall Network Training

Because of the recursive network structure, the PRCA does not need to be pretrained. The whole neural network training procedure, as shown in Fig. 1, is end-to-end. Therefore, all the data-fitting blocks use the same network parameters, and the same applies to the PRCA. In the presence of a reference image, the obvious measure of choice for noise rejection is the MSE. However, in order to meet the multiplicative characteristics of speckle, a logarithmic scale by means of the related DG measure is used to improve readability [38]. The DG measure is defined as

$$\mathrm{DG} = 10 \log_{10}\left(\frac{\mathrm{MSE}(u, f)}{\mathrm{MSE}(u, \hat{u})}\right) \tag{14}$$

where u, $\hat{u}$, and f are, respectively, the clean image, the filtered image, and the speckled image. In order to make the DG indicator more consistent with the convergence characteristic of the loss function, the DG loss function is formulated as

$$\mathrm{Loss} = \arg\min\left(\frac{\sum_i^N \|\mathrm{Net}(f_i; \eta, \{w, b\}) - u_i\|_2^2}{\sum_{i=1}^N \|f_i - u_i\|_2^2}\right). \tag{15}$$

DG loss is equivalent to dividing MSE loss by a fixed value. Usually, data are normalized before the network training.

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

As a result, the loss value is small and the convergence is slow. However, DG loss can solve the problem. In Section IV-C2, we discuss the practicality of DG loss.

IV. Experimental Results and Discussion

In this section, to verify the performance of the proposed SAR-RDCP method, the results obtained with both real and simulated SAR images are described. The proposed method was compared with three traditional methods and one deep learning method: PPB [11], patch ordering-based SAR image despeckling via transform-domain filtering (SAR-POTDF) [14], SAR-BM3D [12], and the DRN for SAR despeckling (SAR-DRN, training with the same data set as the proposed SAR-RDCP method) [23]. For the simulated experiments, the following classic evaluation indices were used: the peak signal-to-noise ratio (PSNR, as high as possible) and the structural similarity index (SSIM, as closer to 1 as possible). PSNR and SSIM are defined as follows:

$$\mathrm{PSNR}(x, y) = 10 \log_{10} \frac{2^{k-1}}{\mathrm{MSE}(x, y)} \tag{16}$$

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{17}$$

where x and y are, respectively, the despeckled image and the reference image; μ_x and μ_y represent the mean of x and y; σ_x and σ_y represent the variance of x and y; σ_{xy} is the covariance between x and y; and c_1 and c_2 are constants. The equivalent number of looks (ENL) and the edge-preservation degree based on the ratio of average (EPD-ROA) are listed for comparison purposes in the real-data experiments [1]. The larger the ENL, the better the quality of the speckle reduction. The ENL for SAR amplitude images is defined as

$$\mathrm{ENL}(A) = 0.5227^2 \cdot \frac{\bar{x}^2}{\mathrm{var}(x)} \tag{18}$$

where A stands for the amplitude image, and $\bar{x}$ and $\mathrm{var}(x)$, respectively, represent the image mean and variance in a homogeneous region. EPD-ROA should be as close to 1 as possible and is defined as

$$\text{EPD-ROA} = \frac{\sum_i^m |I_{D1}(i)/I_{D2}(i)|}{\sum_i^m |I_{O1}(i)/I_{O2}(i)|} \tag{19}$$

where i is the index set of the SAR image; $I_{D1}(i)$ and $I_{D2}(i)$, respectively, represent the adjacent pixel values in the horizontal and vertical directions of the despeckled image; and $I_{O1}(i)$ and $I_{O2}(i)$ represent the adjacent pixel values in the horizontal and vertical directions of the noisy image, respectively. EPD-ROA clearly indicates the edges for multiplicative speckle noise.

1) Parameter Settings and Network Training: The regularization parameter η and the step length δ were initialized as 0.55 and 0.001, respectively. The iteration number T was fixed as 6. The Adam solver [39] was adopted as the gradient descent optimization method, with momentum $\beta_1 = 0.9$, momentum $\beta_2 = 0.999$, and $\varepsilon = 10^{-8}$, where the learning rate γ was initialized to 0.001. The proposed network was trained for 20 epochs with a batch size of 64, and after five epochs, the learning rate was reduced by being multiplied by a descending factor of gamma = 0.1. We implemented the different models in the PyTorch framework and trained the models with an NVIDIA Quadro P4000 GPU.

TABLE II
Average Quantitative Evaluation Results for the Test Data Set

Looks	Index	PPB	SAR-BM3D	SAR-POTDF	SAR-DRN	SAR-VDCNN
10	PSNR	20.0074	21.3854	20.0579	<u>23.1157</u>	**25.0508**
	SSIM	0.7683	<u>0.8249</u>	0.7916	0.8031	**0.8441**
8	PSNR	18.4895	21.2388	19.2465	<u>22.0150</u>	**23.6797**
	SSIM	0.7385	<u>0.8030</u>	0.7856	0.7850	**0.8277**
4	PSNR	18.2656	19.2601	17.8693	<u>21.3697</u>	**22.5394**
	SSIM	0.6627	<u>0.7334</u>	0.6863	0.7291	**0.7767**
2	PSNR	16.3238	17.5382	16.1843	<u>20.4138</u>	**21.1902**
	SSIM	0.5449	0.6427	0.5496	<u>0.6601</u>	**0.7134**
1	PSNR	15.4585	15.3082	14.4934	<u>19.8488</u>	**20.1921**
	SSIM	0.4507	0.5485	0.4494	<u>0.5885</u>	**0.6437**

2) Training and Test Data Sets: Because it is difficult to acquire real noiseless SAR images, we used the UC Merced land-use data set as the training data. In this data set, some images were uniformly selected from each scene to form a 400-image training data set [40], and the data set was evenly divided into five parts of 80 images each, simulating the SAR speckled amplitude images with equivalent look numbers (L) of 1, 2, 4, 8, and 10, which were used as input for the network training. The training data were then cut into patch sizes of 40×40, with the stride equal to 10, and the rotation and turnover operations were randomly carried out to expand the data. The test data set was divided into two parts: simulated data and real data. For the simulated test data set, we randomly selected 21 images of different scenes from the UC Merced land-use data set, which were not in the training data set, and then five different levels of speckle ($L = 1, 2, 4, 8,$ and 10) were added, respectively. For the real test data set, we selected the Flevoland and Death Valley four-look Airborne SAR (AIRSAR) amplitude images, cropped to 500×500.

A. Simulated Data Experiments

For the simulated SAR image despeckling experiments, in order to allow a fair and comprehensive comparison, a combination of quantitative and visual comparisons is used to analyze the effects of the different methods.

Table II lists the average quantitative evaluation results for the test data set, with the best performance marked in bold and the second-best underlined. It can be seen from Table II that, compared to the traditional methods, the deep learning-based methods show a great advantage in quantitative evaluation. However, the proposed SAR-RDCP method shows a more obvious improvement in the degree of speckle removal (PSNR) and the preservation of edge details (SSIM). Generally speaking, detail preservation and smoothing are contradictory. For example, when $L = 10$, 8, and 4, the SSIM of SAR-DRN is lower than that of SAR-BM3D, while the PSNR indices are higher. It is worth mentioning that the proposed method

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

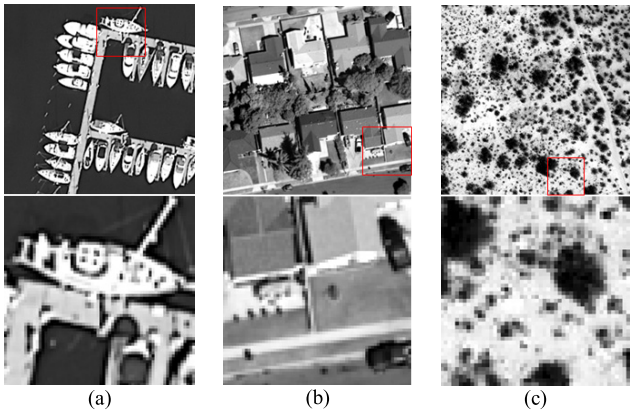

Fig. 4. Original images used in the test experiments. (a) Harbor. (b) Medium-residential. (c) Chaparral.

performs the best in all the different looks in the PSNR and SSIM, which indicates that it maintains a balance between loss of details and speckle removal. Compared to the second-best method, the PSNR of the proposed method shows a gain of about 0.3–2 dB when L goes from 1 to 10. This proves that the larger L, the greater the improvement of the quantitative result. Therefore, the proposed method has more advantages with regard to low noise. We will discuss the situation of strong noise in Section IV-C1.

To give detailed contrasting results, three images (harbor, medium-residential, and chaparral, as shown in Fig. 4) were selected from the test data set to show their visual results in Figs. 5–7, which show the despeckled images and zoomed despeckled images. As can be seen from Fig. 5, with the increase in noise level, the loss of image details becomes stronger and stronger, and the difficulty of despeckling becomes greater and greater. Therefore, the detailed recovery of the image after denoising becomes worse and worse. However, it can be seen from the despeckled results that the proposed method performs the best among all the methods in detailed preservation. For example, in the enlarged image, the detailed part of the ship shows the best restoration effect, and the edge boundary of the ship is the clearest, even with regard to one look with the strongest speckle. For the SAR-DRN method, the biggest problem is that the edges are blurry and not clear. In Fig. 5(b)–(d), the traditional methods produce artificial effects while despeckling. The PPB method shows obvious ringing effects at the edges, while the SAR-BM3D method generates lattice artifacts in the homogeneous area, in addition to oversmoothing. SAR-POTDF has the problem of noise residue, particularly the pointwise artifacts generated at the low look level.

Visual comparisons of the despeckling results for the medium-residential and chaparral images are shown in Figs. 6 and 7, respectively. The medium-residential image covers an area with dense buildings mixed with a small amount of vegetation. In such an area, which is very different from water, the details and textures are more complex. Compared to the other methods, the effect of the proposed method is superior. The chaparral image covers a sparse vegetation area with more complex details. It can be seen from Figs. 6 and 7 that some of the small nodular vegetation is oversmoothed by the other methods, whereas the proposed method is able to restore these areas, to a certain extent. This further demonstrates that the proposed SAR-RDCP method can adapt to a variety of different scenarios.

B. Real-Data Experiments

In order to further illustrate the practicability of this method, the real data of two different scenes were used for the experiments.

It is worth taking a closer look at the despeckled SAR images of the different methods. In Fig. 8, which shows the despeckled images for the Death Valley, oversmoothing is apparent with some of the traditional models, such as PPB and SAR-POTDF, with many strong points and linear targets lost in the background. In addition, the SAR-DRN provides a blurry result with some residual speckle. Compared to the other traditional methods, the SAR-BM3D method performs better, providing an acceptable balance between smoothing and detail preservation. However, the proposed SAR-RDCP method provides some advantages over the SAR-BM3D approach. For example, in the zoomed-in figures, some point structures are well preserved in the results of SAR-RDCP, whereas there is some blurring in the SAR-BM3D results. The results shown in Fig. 9 are consistent with those in Fig. 8, again indicating that the proposed method can obtain the best despeckling effect and has the best detail-preserving ability among the five methods. In addition to oversmoothing, the PPB method also results in the so-called ringing phenomenon near sharp edges. The results obtained with SAR-DRN contain more residual speckles than the previous image. The SAR-BM3D performs well in suppressing the speckle, but it also introduces linear structures. Overall, the proposed method achieves the best balance between controlling for artifacts, noise removal, and detail retention.

In order to make a more accurate comparison, Table III lists the quantitative evaluation indices of ENL and EPD-ROA for the two images. The areas chosen for computing the ENL are indicated by the red rectangles marked as I and II in the original images of Figs. 8 and 9. The best and second-best results are indicated by bold and underline, respectively. From the quantitative results, it can be seen that the proposed method performs well in EPD-ROA. The improvement in ENL values is evident in all the methods, whereas the PPB algorithm shows the largest increase. However, because ENL only reflects the noise removal level of homogeneous regions, it is unfair to blindly rely on the ENL for the evaluation. When combined with the visual results, the PPB algorithm is not the best.

Generally, the ENL can reflect the effectiveness of the algorithm, to some extent, but perfectly homogeneous regions are rare in an image. Therefore, to further illustrate the effectiveness of the algorithm, an unsupervised estimated approach is used, which is called the "ENL map," which involves calculating small sample ENLs in a sliding window (set to 3×3) until the whole SAR image is covered [41]. Figs. 10 and 11 show

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Fig. 5. Despeckled images for the harbor image. (a) Noisy image. (b) SAR-PPB. (c) SAR-BM3D. (d) SAR-POTDF. (e) SAR-DRN. (f) Proposed method.

the ENL maps for the Flevoland and Death Valley images. Because of the small ENL values in heterogeneous regions, such as marginal details, and the large values in homogeneous regions, the ENL values for homogeneous regions should show a large increase, whereas the increase in the heterogeneous regions should be small or even zero. This point is proven

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Fig. 6. Despeckled images for the medium-residential image. (a) Noisy image. (b) SAR-PPB. (c) SAR-BM3D. (d) SAR-POTDF. (e) SAR-DRN. (f) Proposed method.

and shown in Figs. 10 and 11. The maps also show that the despeckling ability of the SAR-DRN needs to be improved, especially for the Flevoland image. Detail losses are apparent in some of the methods, such as PPB, SAR-POTDF, and SAR-BM3D. To sum up, the despeckling ability of the proposed method is the best in the AIRSAR amplitude data.

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

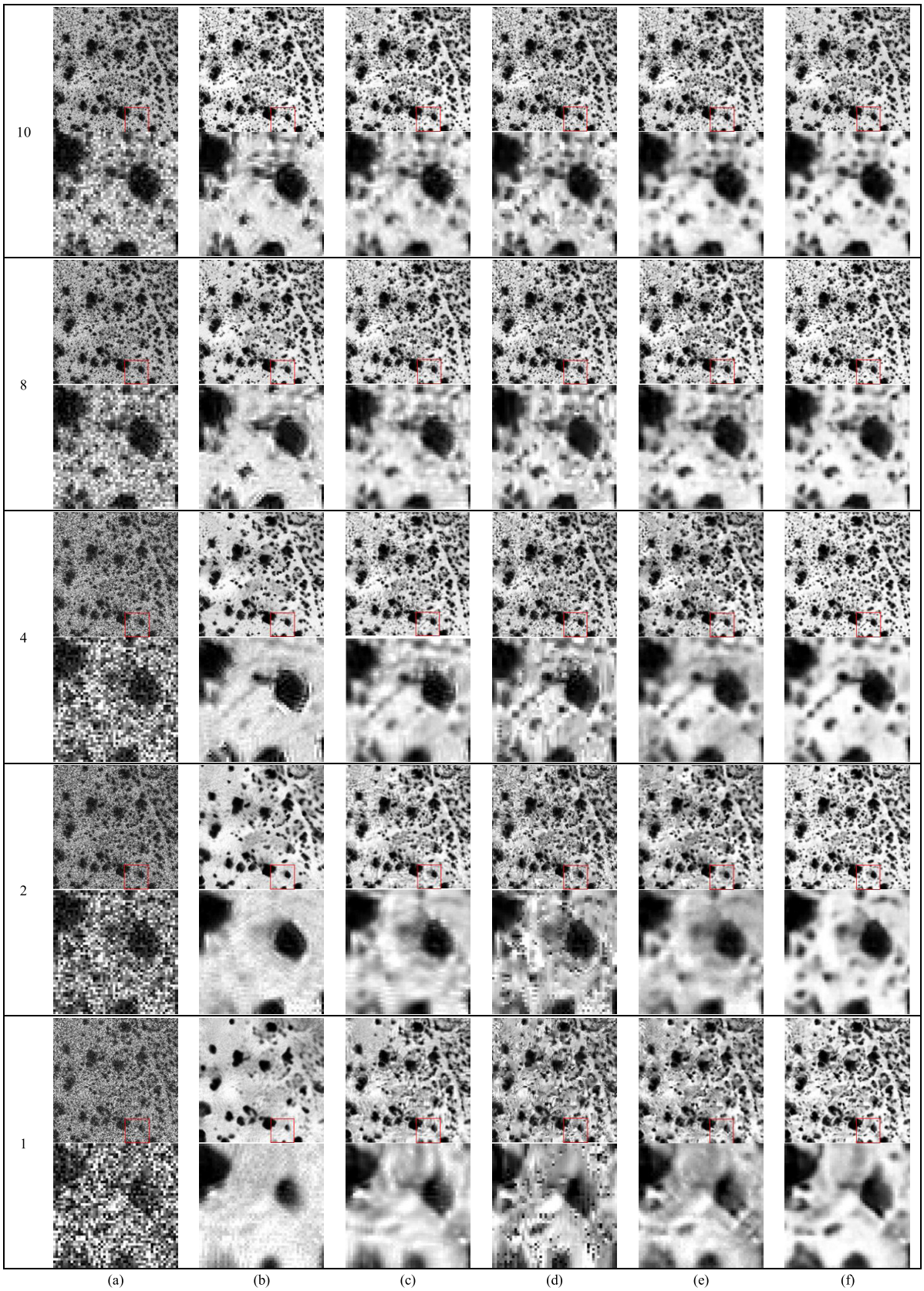

Fig. 7. Despeckled images for the chaparral image. (a) Noisy image. (b) SAR-PPB. (c) SAR-BM3D. (d) SAR-POTDF. (e) SAR-DRN. (f) Proposed method.

C. Further Discussion

1) Strong Noise Situation: In the simulated experiments, we mentioned that the improvement of the proposed SAR-RDCP method is not obvious with regard to strong noise. Therefore, based on the overall framework, we fine-tuned the SAR-RDCP method by adjusting the position of the

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

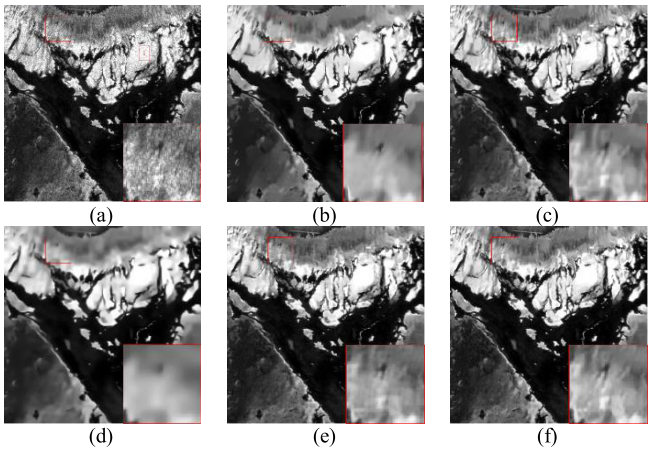

Fig. 8. Despeckled images for the Death Valley image. (a) Noisy image. (b) SAR-PPB. (c) SAR-BM3D. (d) SAR-POTDF. (e) SAR-DRN. (f) Proposed method.

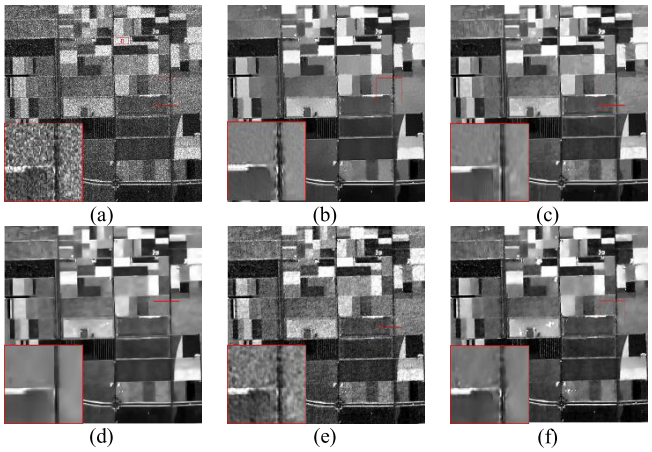

Fig. 9. Despeckled images for the Flevoland image. (a) Noisy image. (b) SAR-PPB. (c) SAR-BM3D. (d) SAR-POTDF. (e) SAR-DRN. (f) Proposed method.

TABLE III

QUANTITATIVE EVALUATION RESULTS FOR THE AIRSAR DATA SET

Method	name	ENL	EPD-ROA-H	EPD-ROA-V
noise	Death Valley	23.54	1.0	1.0
	Flevoland	7.14	1.0	1.0
PPB	Death Valley	**712.60**	0.9541	0.9653
	Flevoland	**493.77**	0.8263	0.8815
SAR-BM3D	Death Valley	215.76	0.9548	0.9671
	Flevoland	366.11	0.8180	0.8789
SAR-POTDF	Death Valley	285.22	0.9420	0.9544
	Flevoland	483.06	0.8134	0.8749
SAR-DRN	Death Valley	191.48	0.9630	**0.9758**
	Flevoland	83.94	0.8260	0.8909
Proposed	Death Valley	446.73	**0.9637**	0.9752
	Flevoland	405.91	**0.8496**	**0.8922**

data-fitting block and the PRCA block to make the data-fitting block the output block, which we call Switched-SAR-RDCP, as shown in Fig. 12. In this regard, the regularization parameter η can be fine-tuned to obtain better despeckling results. To facilitate the display of the results, we selected

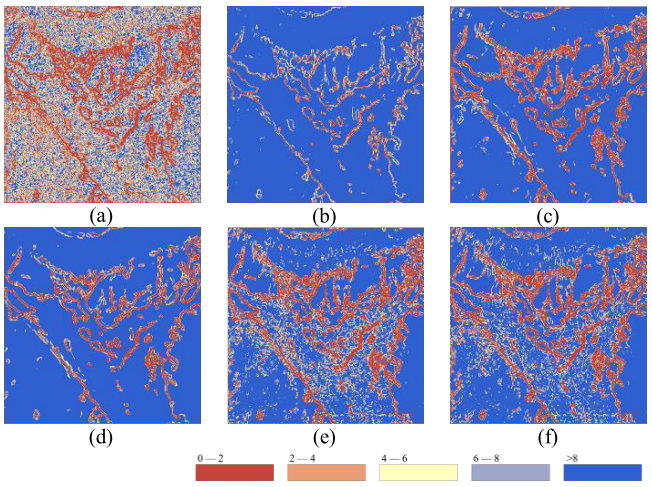

Fig. 10. ENL maps for the Death Valley image. (a) Noisy image. (b) SAR-PPB. (c) SAR-BM3D. (d) SAR-POTDF. (e) SAR-DRN. (f) Proposed method.

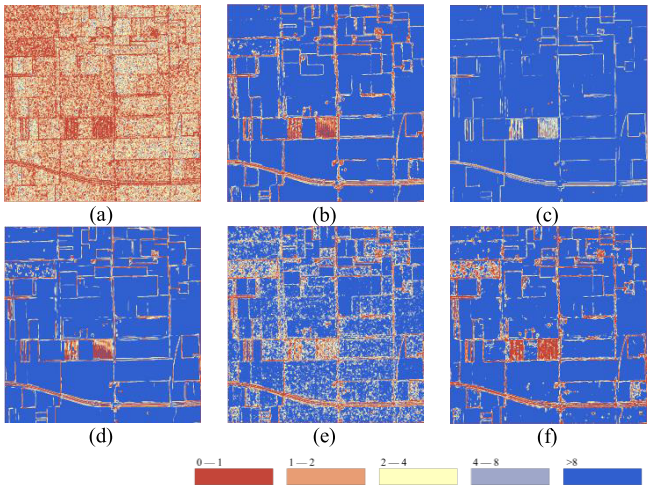

Fig. 11. ENL maps for the Flevoland image. (a) Noisy image. (b) SAR-PPB. (c) SAR-BM3D. (d) SAR-POTDF. (e) SAR-DRN. (f) Proposed method.

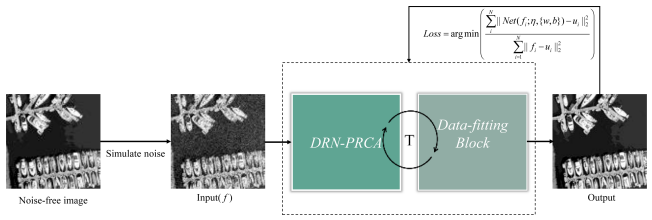

Fig. 12. Overall framework of Switched-SAR-RDCP.

single-look RADARSAT-2 amplitude data of the city of Quebec in Canada and Sentinel-1 HV channel amplitude data of the city of Wuhan in China to compare the detail preservation and speckle suppression shown in Figs. 13 and 14, respectively. Fig. 13(c) and (d) shows that SAR-DRN and SAR-RDCP achieve good performances in detail preservation, but residual speckle is still visible in the images. Fig. 13(b) and (e) shows that SAR-BM3D and Switched-SAR-RDCP perform well in speckle suppression, but in strong point preservation, the result of Switched-SAR-RDCP is better than that of the SAR-BM3D. Fig. 14 shows the despeckling results

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

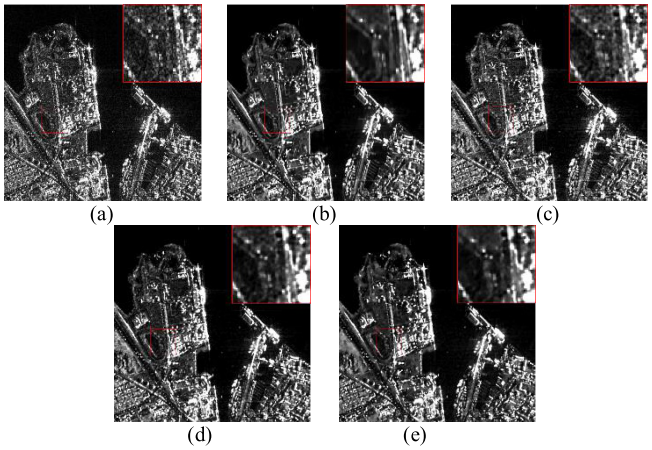

Fig. 13. Despeckled RADARSAT-2 images. (a) Noisy image. (b) SAR-BM3D. (c) SAR-DRN. (d) SAR-RDCP. (e) Switched-SAR-RDCP ($\eta = 0.35$).

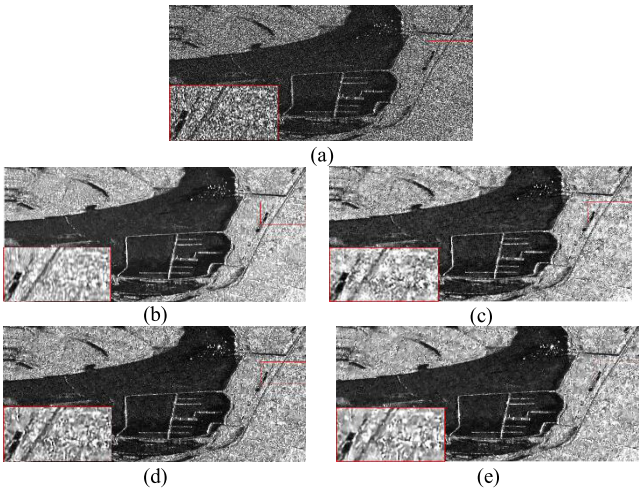

Fig. 14. Despeckled Sentinel-1 images. (a) Noisy image. (b) SAR-BM3D. (c) SAR-DRN. (d) SAR-RDCP. (e) Switched-SAR-RDCP ($\eta = 0.20$).

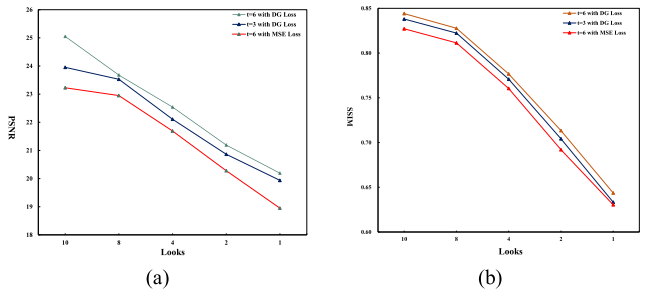

Fig. 15. (a) and (b) PSNR and SSIM, respectively, for the SAR image despeckling results under different looks, iteration numbers, and loss functions.

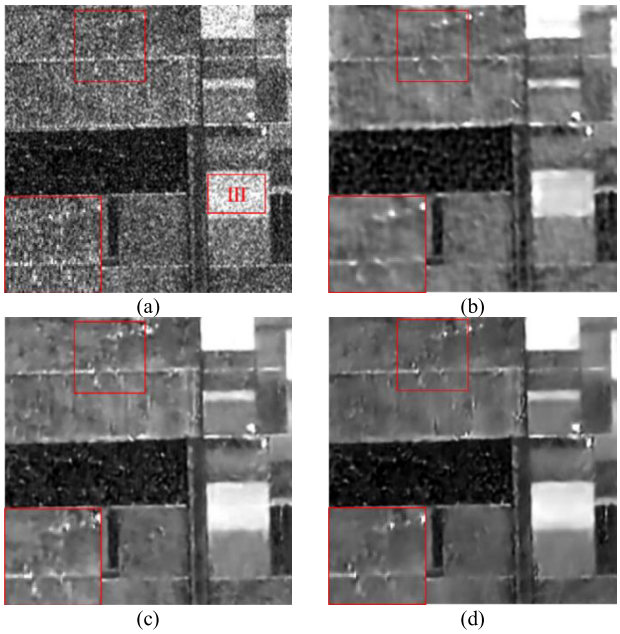

Fig. 16. With/without FDB and the RCAB for the AIRSAR Flevoland image. (a) Original (marked III ENL: 6.46). (b) Without FDB and RCAB (ENL: 67.17). (c) With residual attention block (ENL: 166.98). (d) With FDB and RCAB (ENL: 177.78).

of the four methods on the Sentinel-1 data. From Fig. 14(b), we can see that there is a vagueness phenomenon in the results of SAR-BM3D. Fig. 14(c) shows an obvious loss of detail and blurring. Compared to SAR-BM3D and SAR-DRN, SAR-RDCP and Switched-SAR-RDCP perform better, showing more clear texture and structural information. From the zoomed images in Fig. 14(d) and (e), the Switch-SAR-PDCP method not only preserves details but also removes noise to the maximum extent. Thus, overall, the despeckling effect of Switched-SAR-RDCP is the best among all the methods.

2) Iteration Number t and DG Loss: As mentioned earlier, in the proposed method, the iteration number t is set to 6 and the DG loss is used. In order to verify the reliability of these settings, we designed three sets of comparative experiments in the proposed framework, i.e., $t = 6$ with DG loss, $t = 3$ with DG loss, and $t = 6$ with MSE loss. For the number of iterations t, only 3 and 6 iterations were compared because when the number of iterations is too large, the network training efficiency and test efficiency rapidly fall and the stable performance is reduced. Fig. 15 shows the quantitative evaluation results (PSNR and SSIM) with different speckle levels. It can be clearly seen that the results using DG loss are better, especially when t is equal to 6. This further illustrates that the DG loss can effectively improve the accuracy of the network. Moreover, we find that training is more stable and less prone to gradient explosion when DG loss is used as the optimization loss function in the process of network training. This is because the DG loss increases the overall loss value compared with MSE loss, to make the gradient update more robust.

3) FDB and the RCAB: In deep learning image restoration, feature extraction and enhancement are very important, and are directly related to the learning ability of the network. However, some noise features and weak features will inevitably appear in the process of network feature extraction, which is not conducive for feature reconstruction. Furthermore, SAR data features are more complex, and extraction and recovery are more difficult. In order to verify the impact of the FDB and RCABs, three comparison experiments were implemented with the AIRSAR Flevoland data, as shown in Fig. 16. In the

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

absence of FDB and RCAB, the details are blurred and the residual noise is serious, whereas the structural details are clearly sharpened in Fig. 16(c) with RCAB. Fig. 16(d) shows that the proposed method can suppress the image noise while effectively highlighting the image details. What is more, a homogeneous area marked III with the red rectangle in Fig. 16(a) was selected for the calculation of the ENL, wherein the highest ENL value is obtained by the proposed method. The proposed method also achieves the best visual effect, which demonstrates the validity of the predenoising residual channel attention module.

V. Conclusion

In this article, we have proposed an RDCP despeckling model, by learning a nonlinear end-to-end mapping between the speckled and noise-free SAR images. A number of innovations have also been proposed in terms of the network structure and loss function, wherein FDB and RCAB can enhance the features extracted from the network, while the new DG loss function has also been proposed, the effectiveness of which was verified in the experiments. The simulated- and real-data experiments proved that the proposed SAR-RDCP method outperforms the many mainstream methods, and with regard to strong noise, we suggest using the switch position for the SAR-RDCP method.

In our future work, we will devote ourselves to developing a data-driven method combining imaging mechanisms to achieve more complex and important full-polarization SAR despeckling and consider the utilization of the phase information. Due to the complexity of the full-polarization SAR data, it is difficult to achieve nonlinear mapping at present, and the data are also a problem [42]–[43].

References

[1] J.-S. Lee, M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2363–2373, Sep. 1999.

[2] X. Ma, H. Shen, J. Yang, L. Zhang, and P. Li, "Polarimetric-spatial classification of SAR images based on the fusion of multiple classifiers," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 3, pp. 961–971, Mar. 2014.

[3] J. W. Goodman, "Some fundamental properties of speckle," *J. Opt. Soc. Amer.*, vol. 66, no. 11, p. 1145, Nov. 1976.

[4] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 2, pp. 165–168, Mar. 1980.

[5] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-7, no. 2, pp. 165–177, Mar. 1985.

[6] V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman, "A model for radar images and its application to adaptive digital filtering of multiplicative noise," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, no. 2, pp. 157–166, Mar. 1982.

[7] S. Solbo and T. Eltoft, "Homomorphic wavelet-based statistical despeckling of SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 4, pp. 711–721, Apr. 2004.

[8] F. Argenti, T. Bianchi, A. Lapini, and L. Alparone, "Fast MAP despeckling based on Laplacian–Gaussian modeling of wavelet coefficients," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 1, pp. 13–17, Jan. 2012.

[9] F. Argenti, A. Lapini, T. Bianchi, and L. Alparone, "A tutorial on speckle reduction in synthetic aperture radar images," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 3, pp. 6–35, Sep. 2013.

[10] J.-S. Lee, J.-S. Lee, J.-H. Wen, T. L. Ainsworth, K.-S. Chen, and A. J. Chen, "Improved sigma filter for speckle filtering of SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 202–213, Jan. 2009.

[11] C.-A. Deledalle, L. Denis, and F. Tupin, "Iterative weighted maximum likelihood denoising with probabilistic patch-based weights," *IEEE Trans. Image Process.*, vol. 18, no. 12, pp. 2661–2672, Dec. 2009.

[12] C.-A. Deledalle, L. Denis, G. Poggi, F. Tupin, and L. Verdoliva, "Exploiting patch similarity for SAR image processing: The nonlocal paradigm," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 69–78, Jul. 2014.

[13] S. Parrilli, M. Poderico, C. V. Angelino, and L. Verdoliva, "A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 606–616, Feb. 2012.

[14] B. Xu, Y. Cui, Z. Li, B. Zuo, J. Yang, and J. Song, "Patch ordering-based SAR image despeckling via transform-domain filtering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 4, pp. 1682–1695, Apr. 2015.

[15] S. Chen, L. Gao, and Q. Li, "SAR image despeckling by using nonlocal sparse coding model," *Circuits, Syst., Signal Process.*, vol. 37, no. 7, pp. 3023–3045, Jul. 2018.

[16] W. Ni and X. Gao, "Despeckling of SAR image using generalized guided filter with Bayesian nonlocal means," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 567–579, Jan. 2016.

[17] X. Ma, P. Wu, and H. Shen, "A nonlinear guided filter for polarimetric SAR image despeckling," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1918–1927, Apr. 2019.

[18] G. Aubert and J.-F. Aujol, "A variational approach to removing multiplicative noise," *SIAM J. Appl. Math.*, vol. 68, no. 4, pp. 925–946, Jan. 2008.

[19] J. Shi and S. Osher, "A nonlinear inverse scale space method for a convex multiplicative noise model," *SIAM J. Imag. Sci.*, vol. 1, no. 3, pp. 294–321, Jan. 2008.

[20] X. Ma, H. Shen, X. Zhao, and L. Zhang, "SAR image despeckling by the use of variational methods with adaptive nonlocal functionals," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3421–3435, Jun. 2016.

[21] G. Chierchia, D. Cozzolino, G. Poggi, and L. Verdoliva, "SAR image despeckling through convolutional neural networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 5438–5441.

[22] P. Wang, H. Zhang, and V. M. Patel, "SAR image despeckling using a convolutional neural network," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1763–1767, Dec. 2017.

[23] Q. Zhang, Q. Yuan, J. Li, Z. Yang, and X. Ma, "Learning a dilated residual network for SAR image despeckling," *Remote Sens.*, vol. 10, no. 2, p. 196, 2018.

[24] X. Tang, L. Zhang, and X. Ding, "SAR image despeckling with a multilayer perceptron neural network," *Int. J. Digit. Earth*, vol. 12, no. 3, pp. 354–374, Mar. 2019.

[25] S. Liu *et al.*, "Convolutional neural network and guided filtering for SAR image denoising," *Remote Sens.*, vol. 11, no. 6, p. 702, 2019.

[26] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2808–2817.

[27] W. Dong, P. Wang, W. Yin, G. Shi, F. Wu, and X. Lu, "Denoising prior driven deep neural network for image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 10, pp. 2305–2318, Oct. 2019.

[28] K. Zhang, W. Zuo, and L. Zhang, "Deep plug-and-play super-resolution for arbitrary blur kernels," 2019, *arXiv:1903.12529*. [Online]. Available: http://arxiv.org/abs/1903.12529

[29] R. Liu, L. Ma, Y. Wang, and L. Zhang, "Learning converged propagations with deep prior ensemble for image enhancement," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1528–1543, Mar. 2019.

[30] H. Shen, M. Jiang, J. Li, Q. Yuan, Y. Wei, and L. Zhang, "Spatial–spectral fusion by combining deep learning and variational model," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6169–6181, Aug. 2019.

[31] Z. Huang *et al.*, "Unidirectional variation and deep CNN denoiser priors for simultaneously destriping and denoising optical remote sensing images," *Int. J. Remote Sens.*, vol. 40, no. 15, pp. 5737–5748, Aug. 2019.

[32] C. Xie, Y. Wu, L. V. D. Maaten, A. L. Yuille, and K. He, "Feature denoising for improving adversarial robustness," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 501–509.

[33] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 286–301.

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[34] J.-H. Kim, J.-H. Choi, M. Cheon, and J.-S. Lee, "MAMNet: Multi-path adaptive modulation network for image super-resolution," 2018, *arXiv:1811.12043*. [Online]. Available: http://arxiv.org/abs/1811.12043
[35] Y. Hu, J. Li, Y. Huang, and X. Gao, "Channel-wise and spatial feature modulation network for single image super-resolution," 2018, *arXiv:1809.11130*. [Online]. Available: http://arxiv.org/abs/1809.11130
[36] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.
[37] G. Di Martino, M. Poderico, G. Poggi, D. Riccio, and L. Verdoliva, "Benchmarking framework for SAR despeckling," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1596–1615, Mar. 2014.
[38] H. Feng, B. Hou, and M. Gong, "SAR image despeckling based on local homogeneous-region segmentation by using pixel-relativity measurement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 7, pp. 2724–2737, Jul. 2011.
[39] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[40] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst. (GIS)*, 2010, pp. 270–279.
[41] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft, "Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3795–3809, Nov. 2009.
[42] X. Ma, P. Wu, Y. Wu, and H. Shen, "A review on recent developments in fully polarimetric SAR image despeckling," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 743–758, Mar. 2018.
[43] H. Shen, L. Lin, J. Li, Q. Yuan, and L. Zhao, "A residual convolutional neural network for polarimetric SAR image super-resolution," *ISPRS J. Photogramm. Remote Sens.*, vol. 161, pp. 90–108, Mar. 2020.

Huanfeng Shen (Senior Member, IEEE) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

In 2007, he joined the School of Resource and Environmental Sciences (SRES), Wuhan University, where he is a Luojia Distinguished Professor and an Associate Dean of SRES. He was or is the PI of two projects supported by the National Key Research and Development Program of China and six projects supported by the National Natural Science Foundation of China. He has authored over 100 research articles in peer-reviewed international journals. His research interests include remote sensing image processing, multisource data fusion, and intelligent environmental sensing.

Dr. Shen is a Council Member of the China Association of Remote Sensing Application, an Education Committee Member of the Chinese Society for Geodesy Photogrammetry and Cartography, and a Theory Committee Member of the Chinese Society for Geospatial Information Society. He is a member of the Editorial Board of the *Journal of Applied Remote Sensing and Geography and Geo-Information Science*.

Chenxia Zhou received the B.S. degree in geographical information science from Anhui University, Hefei, China, in 2018. She is pursuing the M.S. degree with the School of Resource and Environmental Sciences, Wuhan University, Wuhan, China.

Her research interests include synthetic aperture radar image despeckling, quality improvement, remote sensing image processing, and deep learning.

Jie Li (Member, IEEE) received the B.S. degree in sciences and techniques of remote sensing and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2011 and 2016, respectively.

He is an Associate Professor with the School of Geodesy and Geomatics, Wuhan University. His research interests include image quality improvement, image super-resolution reconstruction, data fusion, remote sensing image processing, sparse representation, and deep learning.

Qiangqiang Yuan (Member, IEEE) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2006 and 2012, respectively.

In 2012, he joined the School of Geodesy and Geomatics, Wuhan University, where he is a Professor. He has authored more than 70 research articles, including more than 50 peer-reviewed articles in international journals, such as the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. His research interests include image reconstruction, remote sensing image processing and application, and data fusion.

Dr. Yuan was a recipient of the Youth Talent Support Program of China in 2019, the Top-Ten Academic Star of Wuhan University in 2011, and the Hong Kong Scholar Award from the Society of Hong Kong Scholars and the China National Postdoctoral Council in 2014. He is an Associate Editor of the IEEE ACCESS, and has frequently served as a referee for more than 40 international journals for remote sensing and image processing.

Authorized licensed use limited to: Wuhan University. Downloaded on July 15,2020 at 02:44:31 UTC from IEEE Xplore. Restrictions apply.