SAR Image Restoration From Spectrum Aliasing by Deep Learning

ZHE LIU, NING WU, AND XINGXING LIAO
School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
Corresponding author: Zhe Liu (liuzhe@uestc.edu.cn)

ABSTRACT The diversity and the space-variance of the spectrum-aliasing effects in SAR bring challenges to the available model-driven restoration methods. In this paper, a hybrid data-driven and model-driven deep learning scheme is innovatively proposed to deal with this problem. In this scheme, the SAR image restoration is realized via a U-shaped deep neural network (DNN). Meanwhile, the model of the SAR echo data is employed to generate the training data to learn the weighting parameters of the network. The DNN in this method is designed to suit the SAR spectrum-aliasing application. Due to the general SAR echo mathematical model employed for training data construction, the spectrum aliasing features in different SAR systems can be accommodated and learned. Hence, the proposed method can work well for various SAR configurations. Simulation results using the real Radarsat-1 data as well as the synthetic data prove the effectiveness as well as the robustness of the proposed method.

INDEX TERMS Synthetic aperture radar (SAR), spectrum aliasing, restoration, deep learning, U-net.

I. INTRODUCTION Synthetic aperture radar (SAR) is a remote sensing system that collects the echo data from the interest scene and focuses them into a two-dimensional image. In the SAR system, spectrum aliasing often occurs and leads to false detection or incorrect recognition of target, especially in the maritime applications, e.g., the ship detection, the coast surveillance and so on [1], [2]. Thus, restoring a high quality image from the spectrum aliasing is critical for SAR imaging processing.

In the conventional SAR systems where the Nyquist criteria is satisfied, the spectrum aliasing is usually introduced by the non-ideal azimuth antenna pattern (AAP). In many recently-developed SAR systems, e.g., the high-resolution wide-swath (HRWS) SAR systems, in order to enlarge the swath extent while preserving the azimuth resolution or to decrease the downlink data requirement, subNyquist sampling is employed in either/both of the azimuth and range direction of the echo data [3]. The spectrum aliasing in the conventional SAR can lead to the azimuth artifacts; while that in the subNyquist SAR causes the artifacts and the resolution loss in the image.

Most available restoration methods to remedy the SAR spectrum aliasing effect are model-driven, because they are derived from the analytic formation of the echo or the imaging result. Due to the specific constraints on the derivation model, these methods lack of generality. For example, matching filtering (MF) methods [4], [5], selective filtering (SF) methods [6]–[8] and the search method [9] cannot work for subNyquist SAR because they are all based on the spectrum aliasing model of the conventional SAR. Method in [10] cannot work for the squint-mode SAR as it utilizes the location change of the artifacts in different sub-frequency bands under broad-sight condition. Methods in [3], [11]–[13] employ compressive sensing (CS) technique to deal with the aliasing in subNyquist SAR. However, the performance of the CS methods highly depends on the sparsity of the image scene and the restricted isometry property (RIP) of the measurement matrix, which can hardly be satisfied in conventional SAR. Moreover, CS methods are sensitive to noise and this makes it unsuitable for SAR application. Meanwhile, CS methods suffer from huge memory storage requirement and large computation complexity during the construction of the measurement matrix and the vectorization operation of the imaging reflectivity matrix. In [13], weighted lasso is proposed to address the bad RIP problem with an increased storage and computation burden. In [3], fast iterative shrinkage thresholding (FISTA) [14] is extended to

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Feng.
support the 2-D matrix recovery which fits the SAR problem without the usage of the vectorization. However, this method is not flexible due to the specific sampling scheme and additional analogue preprocessing device requirement in the SAR system.

Recently, the data-driven deep-learning methods, which are realized by learning from the training data through the deep neural network (DNN), have shown their superiority over the state-of-the-art model-driven methods in many SAR research fields, such as SAR imaging [15], SAR target recognition [16]–[18], SAR image segmentation [19]–[21] and so on. However, to the authors’ knowledge, no related work about the research of the deep learning in the restoration of the SAR spectrum aliasing problem has been reported yet. In addition, various DNNs have been applied in the biomedical imaging, i.e. magnetic resonance imaging (MRI) and the X-ray computed tomography (CT), to mitigate the streak artifacts [22]–[24]. Nevertheless, they cannot be used in the spectrum-aliasing problem of SAR due to the following two reasons. First, the DNNs are application-specific and there are distinct differences in the image and the artifacts of the two applications. Second, unlike the bio-medical imaging, there is no available training data for image reconstruction purpose in the SAR application.

In this paper, a hybrid data-driven & model-driven method is innovatively proposed to reconstruct the aliasing-free SAR image. The method is data-driven because it restores the image via a DNN, while it is also model-driven due to that the model of the SAR echo is used to generate the training data. The DNN in this method is inspired by the network architecture in [22] and is modified to fit the SAR spectrum-aliasing application. Due to a general SAR echo mathematical model being employed, the model-based training-data construction can ensure the different spectrum aliasing features in various systems being accommodated and learned. Hence, the proposed method can break the constraint of the traditional model-driven methods and can work well for various SAR configurations. The validity of the proposed method is verified by the synthetic data as well as the experiment data.

The organization of this paper is as follows. In Section II, the signal model of the SAR echo data is presented. In Section III, the proposed method is shown in detail. In Section IV, the results with synthetic data as well as experimental data are shown. Section V draws the conclusions.

II. SIGNAL MODEL OF SPECTRUM ALIASING IN SAR

In the SAR system, the echo data after frequency demodulation and 2-D discrete sampling can be expressed by

\[ \hat{S}[m, n] = w(m\Delta \hat{\tau})h(n\Delta \hat{\tau}) \text{rect}(\frac{n\Delta \hat{\tau}}{T_r}) \]

(1)

in which \( w \) is the AAP function, \( h \) is the system transfer function (STF) of the point target, \( T_r \) is the transmitting pulse duration. \( \Delta \hat{\tau} = \frac{1}{F_{\xi}} \), \( \Delta \hat{\tau} = \frac{1}{F_{\eta}} \) are the sampling interval in the 2-D direction determined by the range and azimuth sampling rate \( F_{\xi}, F_{\eta} \), respectively. \text{rect}(z) is the rectangular function, whose value is 1 for \( |z| \leq 1/2 \) and is 0 elsewhere. \( m = -M/2, \cdots, M/2 - 1 \), \( n = -N/2, \cdots, N/2 - 1 \). \( M \) and \( N \) are the 2-D sampling number of echo data. In (1), the target’s scattering coefficient is assumed to be 1 for simplicity; and the range antem function term is ignored as it has little impact on the spectrum aliasing discussed in this paper.

The 2-D spectrum of the echo data can be expressed by

\[ \hat{S}(\xi, \eta) = \sum_{p,q} \hat{S}_{p,q}(\xi, \eta) \]

(2)

where \( \xi, \eta \) are the baseband frequency along the azimuth and the range direction, respectively. \( p, q = \cdots, -1, 0, 1, \cdots \).

The spectrum segment \( \Delta_0,0 \) is the baseband spectrum segment; while \( \hat{S}_{p,q} \) with \( p \neq 0 \) or \( q \neq 0 \) denote the folded ones outside of the baseband. The analytical formula of \( \hat{S}_{p,q} \) is

\[ \hat{S}_{p,q}(\xi, \eta) = W(\xi_p)H(\eta_q)\text{rect}(\frac{\eta_q}{B_{\eta}})\text{rect}(\frac{\xi - \xi_{DC}}{F_{\xi}})\text{rect}(\frac{\eta}{F_{\eta}}) \]

(3)

where \( W \) is the spectrum of the AAP. \( \xi_{DC} \) is the center Doppler frequency, \( \xi_p = \xi - \xi_{DC} + pF_{\xi} \), \( \eta_q = \eta + qF_{\eta} \), \( B_{\eta} \) is the bandwidth of the transmitting signal. The first rect function in (3) is introduced by the limited duration of the transmitting signal and the other two are by the finite sampling rates.

When the Nyquist criteria is satisfied and the AAP is ideal, i.e. \( F_{\xi} > B_{\xi}, F_{\eta} > B_{\eta} \) and \( W(\xi) = \text{rect}(\frac{\xi - \xi_{DC}}{B_{\xi}}) \) with \( B_{\xi} \) being the Doppler bandwidth, the echo is aliasing-free and (2) becomes \( \text{rect}(\frac{\xi - \xi_{DC}}{B_{\xi}})\text{rect}(\frac{\eta}{B_{\eta}})H(\eta, \xi) \). By performing focusing operation on the aliasing-free echo, we can achieve the desired imaging result whose spectrum is

\[ X(\xi, \eta) = \text{rect}(\frac{\xi - \xi_{DC}}{B_{\xi}})\text{rect}(\frac{\eta}{B_{\eta}})\exp(-j\phi) \]

(4)

in which \( \phi = 2\pi(\xi/a + \eta/c) \). \( a, r \) is the target’s azimuth and range position. \( c, v \) is the velocity of light and the SAR platform, respectively.

In practice, however, due to the nonideal AAP and/or the subNyquist sampling, SAR echo spectrum is aliased by \( \hat{S}_{p,q}, p \neq 0, q \neq 0 \) and/or by the truncation in \( \Delta_0,0 \). Therefore, the alias-present echo leads to the degraded image whose spectrum is given by

\[ \hat{X}(\xi, \eta) = \sum_{p,q} \hat{X}_{p,q}(\xi, \eta) \]

(5)

in which \( \hat{X}_{p,q}(\xi, \eta) = W(\xi_p)H(\eta_q)\text{rect}(\frac{\eta_q}{B_{\eta}})\text{rect}(\frac{\xi - \xi_{DC}}{F_{\xi}}) \). According to [6], we have

\[ \Delta H_{p,q}(\xi, \eta) = \exp(-j\frac{\pi}{\mu}(p^2F_{\xi}^2 + 2pF_{\xi}\xi - 2qF_{\eta}\eta)) \times \exp(-j\frac{\pi c r}{2v f_{0}}(p^2F_{\xi}^2 + 2pF_{\xi}\xi(\eta + qF_{\eta})) \]

(6)

where \( f_{0} \) is the central frequency of the transmitting signal. \( \mu = \frac{B_{\xi}}{F_{r}} \).
According to (5) and (6), spectrum segments $\hat{X}_{p,q}$ are combined together to contribute to the degradation of the desired image. Among them, $\hat{X}_{p,q}$ with $p \neq 0$ or $q \neq 0$ generate artifact replicas. $\hat{X}_{0,0}$ produces true target imaging but with the resolution loss under the subNyquist sampling.

### III. PROPOSED METHOD

#### A. INVERSE-PROBLEM MODEL

From (4) and (5), we could find the relationship between the spectrum of the ideal image and the corresponding degraded image is

$$\hat{X}(\xi, \eta) \simeq \Gamma(\xi, \eta)X(\xi, \eta)$$

(7)

where $\Gamma(\xi, \eta) = \sum_{p,q} \Gamma_{p,q}(\xi, \eta)$, $\Gamma_{p,q}(\xi, \eta) = \hat{X}_{p,q}(\xi, \eta)/\hat{X}(\xi, \eta)$.

In (7), we assume that the main energy of the aliasing spectrum concentrate within the Doppler bandwidth in case of the Nyquist criteria being satisfied along the azimuth direction, i.e., $G_{p,q}(\xi, \eta) \simeq G_{p,q}(\xi, \eta) \text{rect}(\xi - \xi_{DC})/B_\xi$ s.t. $F_\xi > B_\xi$.

Let $\Delta G_{p,q}(\xi, \eta) = G_{p,q}(\xi, \eta) - \hat{X}(\xi, \eta)$, we have

$$\Gamma_{p,q}(\xi, \eta) \simeq \Delta H_{p,q}(\xi, \eta) \Delta G_{p,q}(\xi, \eta)$$

(8)

in which the term $\Delta G_{p,q}$ is

$$\Delta G_{p,q}(\xi, \eta) = \begin{cases} 
W(\xi_p); & \frac{F_\eta}{B_\eta} < 1, \\
W(\xi_p)\text{rect}(\frac{F_\eta}{B_\eta}); & \frac{F_\eta}{B_\eta} > 1, \\
W(\xi_p)\text{rect}(\frac{\xi - \xi_{DC}}{B_\xi}); & \frac{F_\xi}{B_\xi} < 1, \\
W(\xi_p)\text{rect}(\frac{\xi - \xi_{DC}}{B_\xi})\text{rect}(\frac{F_\eta}{B_\eta}); & \frac{F_\xi}{B_\xi} < 1.
\end{cases}$$

(9)

Performing 2-D inverse Fourier transform (IFT) on (7), we can get

$$\hat{x}(t, \tau) = \gamma(t, \tau) \otimes x(t, \tau)$$

(10)

where $x$ and $\hat{x}$ are the desired alias-free image and the degraded aliasing-present image, respectively. $\gamma$ is the 2-D IFT of $\Gamma$ and it is the transformation function from the $x$ to $\hat{x}$. $\otimes$ is the convolution operator. Given $\hat{x}$ and $\gamma$, our task is to restore $x$. According to [25] and [26], it can be considered as an inverse problem.

From (8) and (9), we can find that, $\gamma$ varies with the relationship between the bandwidth and the sampling rate. Meanwhile, (6) shows that $\gamma$ is invariant with the target’s range position $r$. Furthermore, owing to the finite size of $\hat{x}$, the limit of $p$, $q$ is dependent on the target’s azimuth and range position and it leads to the 2-D space-variance of the transformation function.

Equations from (7) to (10) show that, restoration the desired SAR image from the degraded one can be uniformly represented as an inverse problem with respect to the transformation function $\gamma$. However, the diversity and the space-variance of the transformation function bring challenges to the conventional model-driven methods and it is hard to solve such problem in a unified way.

#### B. RESTORATION FRAMEWORK

We propose to utilize a DNN to recover the degraded image caused by different kinds of spectrum aliasing in SAR. Through the layer-by-layer nonlinear operation, the deep DNN is used to realize a reconstruction function which is expressive enough to handle the inverse problem confronting diverse and space-variant $\gamma$.

The general model of the reconstruction process is given in Fig. 1 and it can be expressed by

$$\hat{x} = f(\hat{x}; \Theta)$$

(11)

in which $\hat{x}$ and $x$ are the input and the output of the DNN and they are the data matrix of the degraded image and the recovered image, respectively. $f$ is the reconstruction function, $f : C^M \times N \rightarrow C^M \times N$. $\Theta$ is the parameter set.

The framework of the proposed method is given in Fig. 2. First we need construct a DNN architecture, i.e. determine the hyperparameters and architecture of a DNN to get $f$. To recover image from the degraded one using the DNN, we need two more stages—in the training stage, the parameter set is decided from the training dataset; in the restoration stage, we can use the trained DNN to recover the degraded image.

In the following, we give the details of the method.
C. NETWORK ARCHITECTURE

A U-shaped DNN (U-net) is constructed to obtain the reconstruction function model $f$. The U-net includes 22 layers. Its architecture and hyperparameters are shown in Fig. 3. Three different types of layers—the convolution layer, the max-pooling layer and the up-convolution layer—are included in the network. The nonlinear operations in these three layers can be represented by

$$y_{v, \theta, i} = \text{Relu}\left(\sum_{i,j=-1,0,1; \kappa=1,...,K} y_{v+i, \theta+j, \kappa} \alpha_{i,j,k,i} + \beta_i\right)$$

$$\tilde{y}_{v, \theta, i} = \max_{i,j=0,1} \{y_{2v+i, 2\theta+j, i}\}$$

$$\tilde{y}_{2v+i, 2\theta+j, i} = \text{Relu}\left(\sum_{\kappa=1,...,K} y_{v, \theta, \kappa} \alpha_{i,j,k,i} + \beta_i\right)$$

(12)
in which $y$ and $\tilde{y}$ are the input and the output of each layer. $\alpha$ is the weighting coefficient and $\beta$ is the bias. Three subscripts of $y$ and $\tilde{y}$ denote the 2-D index in the channel and the index of the channel, respectively. $\text{Relu}(z) = 0$ for $z < 0$ and $\text{Relu}(z) = z$ for $z > 0$.

The proposed U-net architecture in Fig. 3 is inspired by the DNN architecture in [22] and [23], which is shortened as a JHK U-net in this paper. To fit the SAR aliasing reconstruction application, there are three modifications on the JHK U-net.

1) MODIFICATIONS

First, the input and output data volume is increased from one channel to two channels. Second, the number of channels of the data volume in each layer is reduced to the half of that in JHK U-net, except the input layer and the output layer. In addition, we remove 4 convolution layers, 1 up-convolution layer and 1 maxpooling layer from the JHK net. Besides, the number of layers is reduced from 28 to 22. Third, the residual skip connection between the input and the output layer of the JHK U-net is abandoned in our proposed U-net.

2) REASONS OF MODIFICATIONS

The major reason for the above modifications is based on the different features between the spectrum-aliased SAR image and the bio-medical image, for which the JHK-Unet was designed. The first modification is due to the fact that the SAR image is complex-valued, while the bio-medical image is real-valued. The 2-channel input/output is necessary to process the SAR image because this enables that both the real part and the imaginary part of the image can be simultaneously processed.

The next two modifications are due to the fact that the image degradation in the SAR aliasing is caused by several (generally no more than 4) folded segments of the spectrum. Thus it has a much simpler structure and a smaller coverage than the streaking artifact in the biomedical imaging reconstruction. Therefore, the number of convolution layers can be reduced. Such modification can greatly reduce the computation complexity, the storage requirement and the risk of overfitting. Meanwhile, the network is kept deep enough for expressing complicated nonlinear reconstruction function in our SAR image reconstruction from the spectrum aliasing. Furthermore, according to [27], the residual skip connection need be removed to obtain a better reconstruction performance in the SAR application.

D. TRAINING

In the training stage, the parameters assembling $\Theta = \{\alpha_{i,j,k,i}, \beta_i\}$ of the proposed U-net are determined.

Due to the lack of training samples from the real SAR system, we resort to artificially construct the training dataset. Algorithm1 shows the detail steps. The strategy is to first simulate the echo data pair about the same point target, and then focus the echo data pair through imaging processing method to obtain the image pairs for training.
Denote $s(k)$ and $\hat{s}(k)$ as the k-th pair of aliasing-free and aliasing-present echo with respect to the target located at $(a_k, r_k)$, and they are generated from the following echo models

$$s(k)[m, n] := \sigma_k \hat{w}(m\Delta t) h_k(m\Delta t, n\Delta \tau) \text{rect}(\frac{n\Delta \tau}{T_r})$$

$$\hat{s}(k)[m, n] := \sigma_k \hat{w}(m\Delta t) h_k(m\Delta t, n\hat{\tau}) \text{rect}(\frac{n\Delta \hat{\tau}}{T_r})$$

in which $\sigma_k = \exp(-j2\pi \Phi_k)$ and $\Phi_k$ is a random phase within $[0, 2\pi)$. $h_k(t, \tau) = \exp(-j2\pi \frac{k(t)}{c}) \exp(j\pi \mu (\tau - \frac{R_k(t)}{c})^2)$ and $R_k(t) = \sqrt{r_k^2 + (\nu t - a_k)^2}$ is the range history.

The key procedure of generating training data is to determine the system parameters in (13) to obtain the aliasing-free echo data, we set $\Delta t = \frac{1}{\max[e_1 B_s F_y]}$, $\Delta \tau = \frac{1}{\max[e_2 B_s F_y]}$. $e_1, e_2$ are constant coefficient larger than 1. In addition, the AAP is ideal with $\hat{w}(t) = \text{rect}(\frac{t - t_{DC}}{T_0})$ where $t_{DC} = \frac{a_k - r_k \tan \theta}{c}$, $\theta$ is the squint angle and $T_0$ is the synthetic aperture lasting duration. On the other hand, to generate the aliasing-present echo $\hat{s}(k)$, the 2-D sampling rate and the AAP are set to be the same as the tested SAR system. In practice, we can approximate the realistic AAP with a rectangular function $\hat{w}(t) = \text{rect}(\frac{t - t_{DC}}{\epsilon_3 T_0})$, where $\epsilon_3 > 1$ is the constant coefficient.

Another important parameter is the target position $(a_k, r_k)$. In this method, it is randomly selected within the interested scenario. Due to the fact that the space-variant transformation function $\gamma$ is embedded in the alias-free and alias-present image pair, sufficient training image pairs from random and different targets within the scenario can ensure the training dataset comprehensively accommodate the space-variance. By training the deep U-net with such dataset, the SAR image restoration problem dependent on the space-variant transformation function can be well solved.

After generating the echo pair $s(k)$ and $\hat{s}(k)$, the training image pair $x(k)$ and $\hat{x}(k)$ are obtained by focusing them with the image focusing processing. Conventional SAR imaging techniques, e.g. the direct back-projection (DBP), the range-Doppler method (RD) and so on, can be used.

With the constructed training dataset, we take the L2-norm of the error as the loss function

$$J(\Theta) = \sum_{k=1, \ldots, K} \|f(\hat{x}(k); \Theta) - x(k)\|_2^2.$$  \hspace{1cm} (14)

We train the U-net to obtain the values of the weighting parameters and the bias parameters in the CNN by

$$\Theta = \arg \min_{\Theta} J(\Theta).$$  \hspace{1cm} (15)

The gradient descent and the back propagation method are used to iteratively solve the optimization problem in (15) by

$$\Theta(\ell + 1) = \Theta(\ell) - \rho(\ell) \cdot \nabla_{\Theta} J(\Theta(\ell)).$$  \hspace{1cm} (16)

where $\ell = 0, 1, \ldots, L$ denotes the sequence number of iteration during gradient descent, and $L$ is the total iteration number. The initial values of the parameter vector, $\Theta(0)$, are randomly selected. $\rho(\ell)$ is the learning rate at the $\ell$-th iteration and $\nabla_{\Theta}$ is the gradient operator operation with respect to $\Theta$.

### E. Restoration

After training, the U-net can be used to restore the degraded SAR image. As shown in Fig.2, the SAR echo data with spectrum aliasing were preprocessed through the same focusing method with the training data to get the degraded image. Then the degraded image is sent into the well-trained network. The input image goes through each layer in Fig. 3 and the output is the reconstructed image.

### F. Computation and Memory Complexity

In this method, the computation burden mainly comes from the training dataset construction and the U-net training procedure. The complexity of constructing the training dataset relies on the employed focusing algorithm and the number of the training pairs in the training dataset. When the DBP method is used, the computation cost is in the order of $O(K(MN)^{2/3})$. The complexity of training the U-net is the total numbers of the multiplications operations in all of the convolution and up-convolution layers [28]. Thus, according to Fig. 3, the computation complexity of training the U-net is in the order of $O(10^3 MNL)$, where $L$ is the iteration number during training. In the restoration stage, the computation complexity of the proposed method is in the order of $O(10^5 MN)$.
The memory requirement of the proposed method results from storing the weighting and bias parameters as well as the feature map of each layer. According to Fig. 3, the number of the weighting and bias parameters in the proposed U-net is about $6 \times 10^6$; while the total size of the saved feature map is $10^2 \text{ MN}$. The required memory of the proposed method is determined by the summation of the above two values.

IV. RESULTS

In this section, the synthetic echo as well as the experimental Radarsat-1 data [29] are utilized to verify the proposed method. The system parameters are shown in Table 1.

| Parameters                  | Values         |
|-----------------------------|----------------|
| Transmitting central frequency (GHz) | 5.3            |
| Transmitting bandwidth (MHz)  | 30.116         |
| Transmitting pulse duration (µs) | 41.75          |
| Doppler bandwidth (Hz)       | 1103.9         |
| Platform velocity (m/s)      | 7062           |
| Synthetic aperture time (s)  | 0.637          |
| Equivalent squint angle (degree) | 1.58        |

To verify the generality of the proposed method, we test the image degraded by the spectrum aliasing of 3 different causes. Here, we denote the azimuth subNyquist as case-1, the range & azimuth subNyquist as case-2 and the non-ideal AAP as case-3, respectively. The three cases are discriminated by different sampling rates, which are listed in Table 2.

| Azimuth (Hz) | Range (Hz) | Source of spectrum aliasing                  |
|--------------|------------|-----------------------------------------------|
| Case-1       | 628.49     | Azimuth subNyquist                            |
| Case-2       | 628.49     | Azimuth & range subNyquist                    |
| Case-3       | 1256.98    | non-ideal AAP                                 |

The ideal-residual-ratio (IRR), which is used as one criterion to evaluate the overall recovery performance, is defined as

$$ IRR = \max_{a,b} 20 \log_{10} \left( \frac{\| x \|_2}{\| x - a\tilde{x} + b \|_2} \right) $$  \hspace{1cm} (17)

in which $x$ and $\tilde{x}$ are the desired image and the reconstructed image, respectively. The definition of IRR is the same with the criteria named as ‘SNR’ in [22]. Here, we change the notation as IRR to avoid the confusion with the signal-noise-ratio in the following.

Two more criterions, target-artifact-ratio (TAR) and resolution loss percentage (RLP), which are used to evaluate the ability of the artifact removal and the resolution improvement of point target, respectively, are given by

$$ TAR = 10 \log_{10} \left( \frac{i_T}{i_a} \right), \text{RLP} = \frac{\delta_D}{\delta_0} \times 100\% $$  \hspace{1cm} (18)

in which $i_T$, $i_a$ are the maximum energy within the region of the true target and the artifact replica in the same image, respectively. $\delta_D$, $\delta_0$ are the area of the 3dB resolution cell of the reconstructed image and the desired image at the true target position, respectively. Higher values of IRR, TAR and RLP correspond to better reconstruction performances.

For comparison, traditional methods, e.g. the SF method [6] and the change detection (CD) method [10], are employed. Meanwhile, the JHK U-net architecture in [22] and a recently proposed residual dense U-net model in [21] (abbreviated as RU-net in the following) are also used for comparison. The single-channel input of the two U-net network is the absolute value of the degraded SAR image.

The proposed U-net and the RU-net are implemented with Tensorflow 1.13.0. The JHK U-net is realized through the MatConvNet toolbox [30]. All networks run on a GPU graphic processor of GTX2080Ti (NVIDIA cooperation).

A. TRAINING PROCEDURE

1) TRAINING DATASETS FOR CASE-1 AND CASE-2

Each training dataset with 100 training pairs is constructed by the Algorithm 1 with parameters $\epsilon_1 = 1.07$, $\epsilon_2 = 1.13$, $\epsilon_3 = 1.01$, $r_{\min} = 1011.6km$, $r_{\max} = 1016.4km$, $a_{\min} = -4.31km$, $a_{\max} = 1.44km$. The validation dataset with 10 image pairs are constructed with the same parameters for each case.

The targets’ positions of the training pairs are randomly chosen from the imaging scene. These positions are different from those of test images in the test dataset. DBP is used for echo focusing.

2) TRAINING DATASET FOR CASE-3

We construct the training dataset with 350 image pairs for case-3 according to steps in the Algorithm 1 with parameters $\epsilon_1 = 1.07$, $\epsilon_2 = 1.13$, $\epsilon_3 = 2.01$. $r_{\max} = 1021.09km$, $r_{\min} = 1011.61km$, $x_{\min} = -4.31km$, $x_{\max} = 7.18km$. The validation dataset of 35 image pairs are also constructed with the same parameters. RD is used for imaging.

3) TRAINING U-net

The training datasets for case-1, case-2 and case-3 are sent into the U-net shown in Fig.3. The U-net is trained and the parameters are learned by using the mini-batch gradient descent with the learning rate of 0.001.

To demonstrate the procedure of the parameter adjustment and the hyper-parameter determination of the network,
the average mean square error (MSE) values of the reconstructed images in the validation dataset of the three cases are calculated for each epoch iteration. We change the number of layers while keeping the other hyper-parameters fixed. Due to the symmetry of the U-net, we increase or decrease the layers with the same output size. The number of layers is changed with an increment of 6. The MSE curves as well as the flop counting of the computation cost for the U-nets with different layer numbers are plotted in Fig. 4.

From Fig. 4(a), it can be found that, as the number of layers is increased from 10 to 22, the reconstruction performance is improved. While when the layer number is increased from 22 to 28, the performance is barely changed. As can be seen in Fig. 4(b), the computational complexity will be significantly increased. Therefore, the layer number of 22 is the best balance between the reconstruction performance and the computation cost. In addition, the training termination epoch round is selected to be 20 as the MSE curves vary little around 20 epoch according to Fig. 4(a). The training duration time for 20 epoch training of the three cases is about 12 minutes, 12 minutes and 42 minutes, respectively.

B. TEST DATA PREPARATION AND RESTORATION
1) TEST DATA FOR CASE-1 AND CASE-2
Both test datasets of case-1 and case-2 are synthetic data. For each case, the test dataset is composed with 51 image pairs. Each pair includes one aliasing-present and one aliasing-free desired image of one point target with the same position and they are generated according to Algorithm 1. Among the 51 image pairs, positions of 50 point targets are randomly chosen within the imaging scene which extends about 5.1 km × 6.6 km in the range and azimuth direction. This is to test the performance of the proposed method in processing the space-variance of the reconstruction. One test pair with target position located at the center of the imaging scene P₀ is included to demonstrate the reconstruction performance. DBP is used in the image focusing procedure.
To test the robustness of the proposed method in presence of noise, 5 more test datasets with white Gaussian noise of different powers are added to the previously generated test dataset for case-1 and case-2, respectively. The signal-to-noise-ratio (SNR) of the 5 datasets varies from 10dB to 50dB with the interval of 10dB.

2) TEST DATA FOR CASE-3
We use the raw echo data from the well-known Radarsat-1 dataset to generate test data for case-3. The system parameters of the Radarsat-1 are listed in Table 1 and the 2-D sampling rates are given in Table 2. More details about the data can be found in [31].

The experimental test image is obtained by focusing the Radarsat-1 echo with the RD imaging method, and we select two marine zones from the imaging results. The scene extension of the zone-1 is 17.4km $\times$ 26.9km and that of the zone-2 is 13.4km $\times$ 18.6km along the azimuth and range direction.

3) TEST DATA RESTORATION
The test images in the dataset of each case are sent into the corresponding well-trained U-net sequentially to evaluate the performance of different methods.

FIGURE 6. Processing results of different methods for case-2 with SNR = Inf.

FIGURE 7. IRR vs different SNR. From left to right are the curves of case-1 and case-2.
FIGURE 8. Processing results of different methods for zone-1 of the experimental RADARSAT-1 data.

restoration performance. The processing time to restore one test image is about 3 second for the three cases.

C. RESTORATION QUALITY ASSESSMENT

1) RESULTS OF CASE-1 AND CASE-2
Fig.5 and Fig.6 show the 2-D contour of the reconstruction results of one point target located at $P_0$ under the case-1 and case-2 configurations without noise. In each figure, the true target and the artifacts are marked with black solid and black dashed rectangular, respectively. The overview and the enlarged view around the position of the true target and the two artifact replicas are shown in each figure.

From the results, we can observe that, the spectrum aliasing caused by the subsampling arises obvious artifacts in the azimuth direction and along the 2-D directions. Moreover, there is resolution loss at the true target location according to the RLP values. After the processing of the proposed U-net, the energy of the artifacts are significantly suppressed. A very high TAP values can be obtained and the artifacts can hardly be found in the restored image. In addition, the resolution is improved to be close to the ideal one. Comparing with the other three methods, the proposed method has the best reconstruction performance with the highest IRR, TAR and RLP. In the following, we demonstrate the restoration quality confronting space-variance and noise by evaluating the average IRR of the reconstruction results of all images in each test dataset. The curves of the average IRR for the proposed method, the DBP method, the SF method and the JHK U-net method are shown in Fig. 7. The curves of the average IRR of the CD processing results are not present due to its average IRR retaining nearly zero for all test samples.

From Fig. 7, it can be found that, for the two cases, the IRR values of the proposed method stay above 20dB with SNR no less than 30dB. Meanwhile, they gradually decline at higher noise levels. These results show that, the proposed method can deal with the space-variance and it is robust to tolerate a broader interference range of noise, even though it is trained under the noise-free configuration. On the other hand, the curves of the SF method and the degraded image obtained from the DBP imaging processing coincide. This means that the SF method has no impact on the restoration performance under the subNyquist conditions. In addition, the IRR values of the JHK U-net method degrade more dramatically with the increase of the noise levels compared to the proposed method. Such results illustrate that the proposed method is more robust than the JHK U-net method and it can work in a wider range of noise background.

2) RESULTS OF CASE-3
The imaging results of the zone-1 and zone-2 are shown in Fig. 8(a) and Fig. 9(a), respectively. The interested scenario within the two zones are marked with yellow solid rectangulars and they are enlarged in the leftmost of Fig. 8(b) and Fig. 9(b).
FIGURE 9. Processing results of different methods for zone-2 of the experimental RADARSAT-1 data.

As can be found, artifacts replicas appear in the two scenarios, they are circled by green dashed rectangulars. The true targets are marked with red solid ovals. Only 2 true targets in the upper right of the interested scenario of zone-1 and their replicas are present together, while the artifact replicas of the other four targets are not included in the image. Similarly, for the interested scenario in zone-2, no corresponding replica of the true target within the scene is present, but two replicas of targets outside the scene are shown. Thus, these two figures illustrate the variance of the replica sections caused by targets’ locations in the same image.

The enlarged reconstruction results of the proposed method as well as those of the other three methods are given in Fig. 8(b) and Fig. 9(b), respectively. It can be seen that, despite the space-variance, the artifacts are removed while the true targets are reserved through our proposed method. On the contrary, there are residuals of artifact replicas in the results of the CD method and the JHK U-net. Although the artifact is mitigated in the result of the SF method, the reconstruction performance of the proposed method is much better due to the higher value of the TAP. As for the RU-net method, the true targets as well as the artifacts are both removed in the restored image.

D. SUMMARY
In the above, the validity of the proposed method has been illustrated in different SNR scenarios by different spectrum aliasing causes. The test results demonstrate that, the proposed method can work well for the reconstruction of the image degradation caused by the non-ideal AAP, azimuth subsampling and 2-D subsampling. Meanwhile, it can deal with the space-variance problem and can achieve a good recovery performance in the noise-present background with a broad range of noise levels. Compared to the traditional SF method, the CD method, the JHK U-net and the RU-net, the proposed method has the best recovery performance.

V. CONCLUSION
In this paper, the restoration of degraded SAR image from the spectrum aliasing by different causes is studied. We first demonstrate a general transformation model between the desired alias-free SAR image and the degraded alias-present image. It shows that the challenge of recovering SAR image in a unified way comes from the diversity and the space-variance of the transformation. In order to conquer such problem, we propose a deep learning method based on a 22-layer U-net, which can handle the diversity and
the space-variance problem after well training. Meanwhile, through a general SAR echo model, the training data which accommodate different spectrum aliasing features can be artificially generated. The validity, the robustness and the generality of the proposed method have been demonstrated by the Radarsat-I data as well as the synthetic data.

ACKNOWLEDGMENT
The authors would like to thank the anonymous reviewers for their valuable suggestions.

REFERENCES
[1] S. Bruschi, S. Lehner, T. Fritz, M. Soccorsi, A. Soloviev, and B. van Schie, “Ship surveillance with TerraSAR-X,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 1092–1103, Mar. 2011.

[2] D. Velotto, M. Soccorsi, and S. Lehner, “Azimuth ambiguities removal for ship detection using full polarimetric X-band SAR data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 76–88, Jan. 2014.

[3] K. Aberman and Y. C. Eldar, “Sub-nygquist SAR via Fourier domain range-Doppler processing,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6228–6244, Nov. 2017.

[4] A. Moreira, “Suppressing the azimuth ambiguities in synthetic aperture radar images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 4, pp. 885–895, Jul. 1993.

[5] J. Chen, K. Wang, W. Yang, and W. Liu, “Accurate reconstruction and suppression for azimuth ambiguities in spaceborne stripmap SAR images,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 102–106, Jan. 2017.

[6] A. M. Guarnieri, “Adaptive removal of azimuth ambiguities in SAR images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 625–633, Mar. 2005.

[7] G. Di Martino, A. Iodice, D. Riccio, and G. Ruello, “Filtering of azimuth ambiguity in stripmap synthetic aperture radar images,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 9, pp. 3967–3978, Sep. 2014.

[8] Y. Wu, Z. Yu, P. Xiao, and C. Li, “Suppression of azimuth ambiguities in spaceborne SAR images using spectral selection and extrapolation,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6134–6147, Oct. 2018.

[9] J. Chen, M. Iqbal, W. Yang, P.-B. Wang, and B. Sun, “Mitigation of azimuth ambiguities in spaceborne stripmap SAR images using selective restoration,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4038–4045, Jul. 2014.

[10] Y. Wang, Z. Zhang, N. Li, F. Hong, H. Fan, and X. Wang, “Maritime surveillance with undersampled SAR,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1423–1427, Aug. 2017.

[11] J. Fang, Z. Xu, B. Zhang, W. Hong, and Y. Wu, “Fast compressed sensing SAR imaging based on approximated observation,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 1, pp. 352–363, Jan. 2014.

[12] X. Dong and Y. Zhang, “A novel compressive sensing algorithm for SAR imaging,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 2, pp. 708–720, Feb. 2014.

[13] G. Xu, X.-G. Xia, and W. Hong, “Nonambiguous SAR image formation of maritime targets using weighted sparse approach,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1454–1465, Mar. 2018.

[14] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.

[15] B. Yonel, E. Mason, and B. Yazici, “Deep learning for passive synthetic aperture radar,” *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 90–103, Feb. 2018.

[16] J. Pei, Y. Huang, W. Huo, Y. Zhang, J. Yang, and T.-S. Yeo, “SAR automatic target recognition based on multiview deep learning framework,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2196–2210, Apr. 2018.

[17] M. Wilkanski, C. Kreucher, and J. Lauer, “Modern approaches in deep learning for SAR ATR,” *Proc. SPIE*, vol. 9843, May 2016, Art. no. 98430N.

[18] Q. Yu, H. Hu, X. Geng, Y. Jiang, and J. An, “High-performance SAR automatic target recognition under limited data condition based on a deep feature fusion network,” *IEEE Access*, vol. 7, pp. 165646–165658, 2019.

[19] C. Henry, S. M. Azimi, and N. Merkle, “Road segmentation in SAR satellite images with deep fully convolutional neural networks,” *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 12, pp. 1867–1871, Dec. 2018.

[20] N. Lv, C. Chen, T. Qiu, and A. K. Sangaiah, “Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in SAR images,” *IEEE Trans Ind. Informat.*, vol. 14, no. 12, pp. 5530–5538, Dec. 2018.

[21] P. Shamsholmoli, M. Zareapoor, R. Wang, H. Zhou, and J. Yang, “A novel deep structure U-net for sea-land segmentation in remote sensing images,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3219–3232, Sep. 2019.

[22] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep convolutional neural network for inverse problems in imaging,” *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.

[23] Y. Seob Han, J. Yoo, and J. C. Ye, “Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis,” 2016, *arXiv:1611.06391*. [Online]. Available: http://arxiv.org/abs/1611.06391

[24] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, “Image reconstruction by domain-transform manifold learning,” *Nature*, vol. 555, no. 7697, pp. 487–492, Mar. 2018.

[25] B. Chalmonds, *Modeling and Inverse Problems in Image Analysis Plus*. New York, NY, USA: Springer, 2003.

[26] J. K. Seo and E. J. Woo, *Nonlinear Inverse Problems in Imaging*. New York, NY, USA: Wiley, 2012.

[27] J. C. Ye, Y. Han, and E. Cha, “Deep convolutional framelets: A general deep leaning framework for inverse problems,” *SIAM J. Imag. Sci.*, vol. 11, no. 2, pp. 991–1048, Jan. 2018.

[28] K. He and J. Sun, “Convolutional neural networks at constrained time cost,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5353–5360.

[29] G. W. Titi and D. F. Marshall, “The ARPA/NAVY mountaintop program: Adaptive signal processing for airborne early warning radar,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. Conf. Proc.*, vol. 2, May 1996, pp. 1165–1168 2.

[30] A. Vedaldi and K. Lenc, “MatConvNet: Convolutional neural networks for MATLAB,” in *Proc. 23rd ACM Int. Conf. Multimedia (MM)*, New York, NY, USA: ACM, 2015, pp. 689–692.

[31] I. G. Cumming and F. H. Wong, *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*. Norwood, MA, USA: Artech House, 2004.

ZHE LIU received the M.Sc. and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2002 and 2009, respectively. She is currently an Associate Professor with UESTC. Her research interests include radar signal processing and machine learning.

NING WU received the B.S. degree from the School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2018, where he is currently pursuing the M.Sc. degree. His research interest centers on deep learning and advanced signal processing methods for radar.

XINGXING LIAO received the B.S. degree from Yanbian University, China. She is currently pursuing the M.Sc. degree in information and communication engineering with the University of Electronic Science and Technology of China. Her research interest includes compressed sensing and advanced signal processing methods for radar.