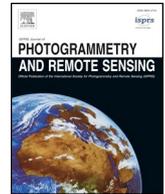




Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

A residual convolutional neural network for polarimetric SAR image super-resolution

Huanfeng Shen^{a,d}, Liupeng Lin^a, Jie Li^{b,*}, Qiangqiang Yuan^{b,d}, Lingli Zhao^c

^a School of Resource and Environmental Sciences, Wuhan University, Wuhan, China

^b School of Geodesy and Geomatics, Wuhan University, Wuhan, China

^c School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

^d Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan, China

ARTICLE INFO

Keywords:

Polarimetric SAR
Super-resolution
Residual convolutional neural network
Remote sensing
Deep learning

ABSTRACT

PolSAR images provide rich polarimetric information, however, due to the limitations of the imaging system, the spatial resolution decreases while the richer polarimetric information is obtained. The lower resolution limits the application, so it is necessary to use super-resolution technology to improve the spatial resolution. In this paper, in response to the low spatial resolution of PolSAR images, a PolSAR super-resolution framework is proposed to improve the spatial resolution by the use of a residual convolutional neural network. Within this framework, deconvolution is used to up-sample the PolSAR images, PReLU is added to maintain the numerical properties. A complex structure block is also designed to accommodate the PolSAR data structure. In addition, prior information on the low-resolution image itself is used to reduce the artifacts. The proposed method shows a superior performance when compared to the traditional methods in both the quantitative evaluation and visual assessment. The proposed method improved the spatial resolution significantly, especially in terms of detail information retention, and it improves the mean PSNR by more than 12% when compared to the traditional methods. By analyzing the phase statistics and polarimetric response, it is shown that the proposed method has a good polarimetric information retention ability, and can obtain a higher classification accuracy.

1. Introduction

Fully polarimetric synthetic aperture radar (PolSAR) images can provide different scattering characteristics from multiple polarimetric channels (HH, HV, VH, VV) of the same scene, which can provide richer information for ground interpretation than single-polarimetric SAR. PolSAR images can thus be applied in multiple fields, such as building damage detection after an earthquake, marine ship monitoring, agriculture, wetland mapping, and classification (Qi et al., 2015; Lê et al., 2015; Zhao et al., 2013; Shi et al., 2015; Liu and Gierull, 2007; Liu et al., 2019; Zhao et al., 2017; Shi et al., 2013; Liu et al., 2019; Hänsch and Hellwich, 2018; Dabboor and Shokr, 2013; White et al., 2015; Mahdianpari et al., 2017). However, in the process of PolSAR imaging, in order to obtain the information of the multiple polarimetric channels, it is inevitably to decrease the spatial resolution to obtain more abundant polarimetric features (Song et al., 2018). For some SAR sensors, the limitation of this system results in the spatial resolution of the PolSAR images often being lower than the spatial resolution of single-polarimetric SAR images. The reduction of the PolSAR image resolution

limits its application in a variety of fields. Therefore, it is necessary to enhance the resolution of PolSAR images by means of image processing. Recently, super-resolution reconstruction technology (Yue et al., 2016) has been widely used in optical and single-polarimetric SAR images to enhance the resolution. SAR super-resolution is an image enhancement technique used to improve the spatial resolution and reduce the side-lobes of SAR images (Zhang and Zhang, 2013), however, to date, there have been relatively few studies of the super-resolution of PolSAR images. For the purpose of improving the resolution of PolSAR images, Suwa et al. (Suwa and Iwamoto, 2007) proposed the 2-D polarimetric bandwidth extrapolation (2D-PBWE) method, which used 2-D polarimetric linear prediction model to extend the spatial frequency bandwidth of the azimuth and distance. This method extended bandwidth extrapolation to PolSAR images, and improved the resolution of the azimuth and range directions, but the polarimetric information of the PolSAR imagery is not fully utilized. Zou et al. (Zou et al., 2008) proposed a super-resolution method based on target decomposition and quadrant pixel (SRQP). In this method, target polarimetric decomposition was first performed on the PolSAR images, the central pixel

* Corresponding author.

E-mail address: jli89@sgg.whu.edu.cn (J. Li).

<https://doi.org/10.1016/j.isprsjprs.2020.01.006>

Received 30 August 2019; Received in revised form 28 December 2019; Accepted 6 January 2020

Available online 21 January 2020

0924-2716/ © 2020 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

was then divided into four sub-pixels, and finally obtained the coherency matrix values of the four sub-pixels by weighting the central pixel and the adjacent pixels. However, by only considering the relationship between the central pixel and the neighboring pixels, the reconstruction results have a distinct grid effect. The PolSAR super-resolution method based on polarimetric spatial correlation (SRPSC) (Zhang et al., 2011) was subsequently proposed to introduce the polarimetric spatial correlation between sub-pixels to reduce the grid effect and maintain the polarimetric characteristics of the results. However, the inappropriate polarimetric target decomposition for this super-resolution method limits the accuracy of the super-resolution. In addition, super-resolution algorithms for PolSAR images have been developed based on specific targets. The spectral-analysis-based techniques (Pastina et al., 2003) and the projection onto convex sets (POCS) algorithm (Jiong and Jian, 2007) were applied to the super-resolution of PolSAR images of ship targets. This type of method focused on a single type of target, and other types of features was not considered.

For the super-resolution of PolSAR images, robust reconstruction results can be achieved by taking the multiple polarimetric channel information into account (Pastina et al., 2001). However, due to the special imaging mode by a coherent superposition of the received echoes during imaging, the imaging process is complex, and the relationship between the different polarimetric channels is also relatively complicated. Therefore, it is difficult to fit the relationship between different polarimetric channels by the use of a simple linear model. In recent years, deep learning has been widely adopted to fit complex nonlinear relationships. For image processing, convolutional neural networks (CNNs) have exhibited better fitting capabilities than the traditional linear models. For natural image super-resolution, a number of deep-learning-based image super-resolution methods have been proposed. The super-resolution convolutional neural network (SRCNN) model (Dong et al., 2014) showed a superior super-resolution effect when compared to the traditional reconstruction model. Subsequently, in order to reduce the loss of precision and the computational burden caused by the image interpolation, the fast super-resolution convolutional neural network (FSRCNN) model (Dong et al., 2016) introduced deconvolution into the network. Some studies have also focused on the deepening of the convolutional layers. The very deep convolutional network (VDSR) model (Kim et al., 2016) used residual learning to solve the model degradation problem caused by the increase in the number of layers (He and Sun, 2015; He et al., 2016). Furthermore, the global residual learning of VDSR allowed it to converge quickly (Tai et al., 2017). However, as the number of network layers increase, the network parameters also increase. To solve this problem, the deeply-recursive convolutional network (DRCN) model (Kim et al., 2016) reduced the number of network parameters by using the same convolutional layer in multiple recursions, but this recursive approach increased the network training time. Thereafter, based on residual blocks, the enhanced deep super-resolution network (EDSR) model (Lim et al., 2017) removed the modules and simplified the network. In addition, the deep back-projection network (DBPN) model (Haris et al., 2018) introduced an up-projection unit and a down-projection unit and reconstructed the high-resolution imagery by mutually connecting the up-sampling and down-sampling stages. CNNs have also been successfully used in remote sensing (Zhang et al., 2016; Zhang et al., 2018; Yuan et al., 2019; Wang et al., 2017; Chierchia et al., 2017; Zhang et al., 2018; Zhou et al., 2018; Zhang et al., 2017). For example, the Very Deep Convolutional Neural Network (Sun and Wang, 2018) and fully convolutional neural networks have been used in remote sensing applications (Mahdianpari et al., 2018). However, the application of deep learning in the field of super-resolution of PolSAR images is still blank. Hence, in this paper, we aim to fit a relationship between the multiple polarimetric channels of PolSAR images, and then combine the multiple polarimetric channels to reconstruct the resolution of the PolSAR imagery by the use of CNN model.

With regard to the super-resolution of PolSAR imagery, the first

difficulty is the incompatibility between the multiple polarimetric channel SAR data and the single luminance channel networks. The relationship between the polarimetric channels of the PolSAR imagery is complicated. If the PolSAR imagery is reconstructed channel by channel, the relationship between the polarimetric channels is ignored, which will inevitably affect the reconstruction accuracy. In addition, the numerical characteristic of PolSAR images is diverse. With the existence of both negative and complex numbers, PolSAR data far exceed the range of natural image pixel values, which are generally neither negative nor complex values. Therefore, directly using a natural image super-resolution network will lose some of the numerical information of the PolSAR imagery. Hence, in this paper, a residual convolutional neural network based PolSAR image super-resolution method is proposed, considering the complex structure of PolSAR imagery from the following aspects. On the one hand, the number of channels in the network is changed to adapt to the properties of the multi-channel PolSAR imagery. On the other hand, according to the numerical characteristics of the PolSAR imagery, a complex structure block is designed to extract the mixed features of the complex domain of the SAR imagery, and the activation function is improved to adapt to the data characteristics.

The rest of this paper is organized as follows. Section 2 introduces the PolSAR data and PolSAR image degradation. The framework of the proposed PolSAR image super-resolution method and the details of its structure are also introduced in Section 2. The experimental results are provided in Section 3. A discussion on the residual compensation (RC) strategy is presented in Section 4, followed by our conclusions and future work in Section 5.

2. The proposed method

2.1. Data organization and model framework

Different echo emitting and receiving modes enable PolSAR data to have multiple channels with different polarimetric properties. However, this structural complexity makes the PolSAR data super-resolution process even more difficult. Therefore, it is necessary to reorganize the data structure. Under the condition of satisfying the theory of reciprocity (Lee and Pottier, 2009; Chitroub et al., 2002), and ignoring the system noise in PolSAR imaging systems, the backscattering matrix can be converted into a coherency matrix:

$$T_3 = k \cdot k^{*T} = \frac{1}{2} \begin{bmatrix} |S_{HH} + S_{VV}|^2 & (S_{HH} + S_{VV}) & 2(S_{HH} + S_{VV})S_{HV}^* \\ (S_{HH} - S_{VV})^* & |S_{HH} - S_{VV}|^2 & 2(S_{HH} - S_{VV})S_{HV}^* \\ (S_{HH} - S_{VV})(S_{HH} + S_{VV})^* & 2S_{HV}(S_{HH} + S_{VV})^* & 4|S_{HV}|^2 \end{bmatrix} \quad (1)$$

where S_{HH} and S_{VV} represent the co-polarized channels, and S_{HV} and S_{VH} represent the cross-polarized channels.

Considering the elements on the diagonal of the coherency matrices being real numbers, and the non-diagonal elements of the coherency matrices being complex numbers, the coherency matrix can be rewritten as:

$$T_3 = \begin{bmatrix} R_{11} & R_{12} + I_{12}j & R_{13} + I_{13}j \\ R_{21} + I_{21}j & R_{22} & R_{23} + I_{23}j \\ R_{31} + I_{31}j & R_{32} + I_{32}j & R_{33} \end{bmatrix} \quad (2)$$

where $j = \sqrt{-1}$ is the imaginary unit. For the coherency matrices, two elements of symmetric position on a non-diagonal line are conjugate complex numbers, for example, $R_{12} = R_{21}$, $I_{12} = -I_{21}$. In this paper, we extracted the value of the upper triangular matrix on the coherency matrix T_3 and converted it into a 1×9 real-valued vector T_{value} .

$$T_{value} = [R_{11}, R_{12}, I_{12}, R_{13}, I_{13}, R_{22}, R_{23}, I_{23}, R_{33}]^T \quad (3)$$

Thus, the super-resolution process for PolSAR imagery can be

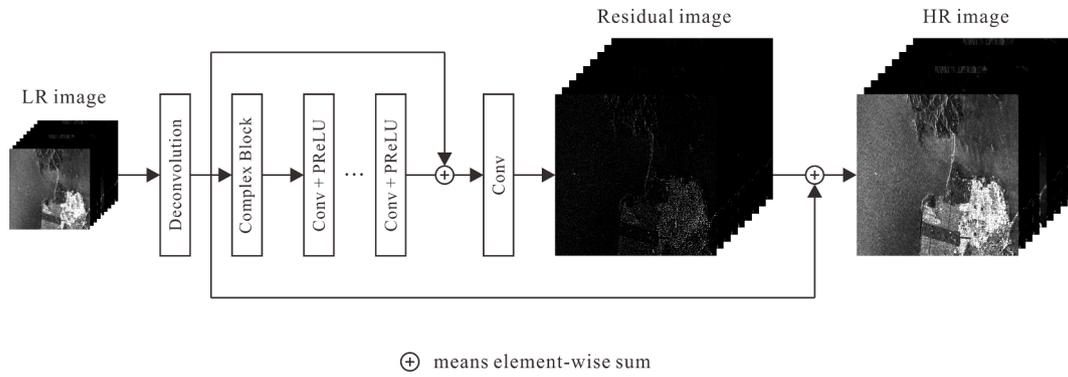


Fig. 1. The framework of the proposed method.

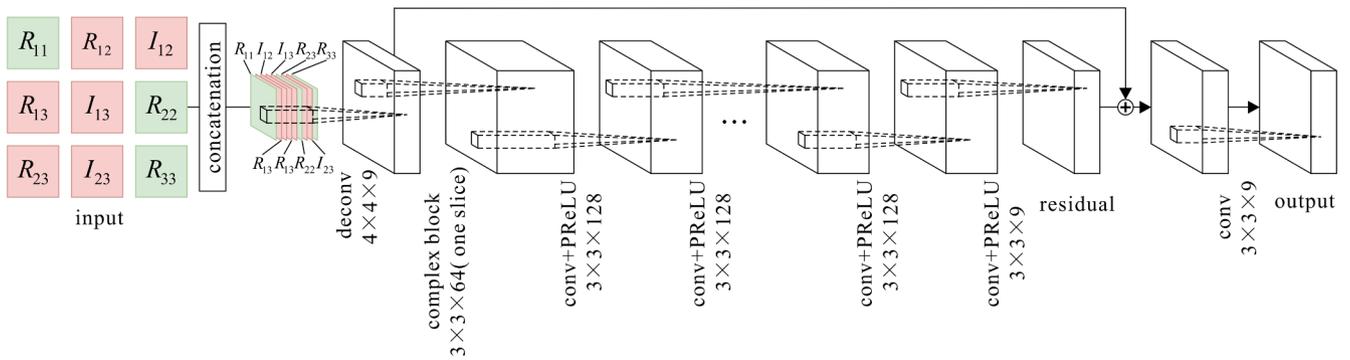
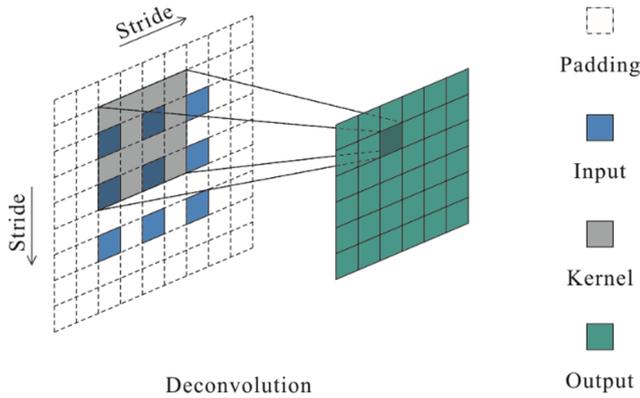


Fig. 2. The convolutional layers of the proposed network.



Deconvolution

Fig. 3. The deconvolution scheme.

transformed into the super-resolution reconstruction of vector T_{value} . Based on the above description, the PolSAR image degradation process can be expressed as follows:

$$T_y = f_{ds}(T_x) \tag{4}$$

where T_y is the observed low-resolution PolSAR image, and T_x is the high-resolution PolSAR image. Both T_y and T_x are a 3-D matrix composed of multiple T_{value} , and the vector represents the value vector of the upper triangular matrix of one pixel of the PolSAR images T_y and T_x . f_{ds} represents the down-sampling function, describing the underlying relationship between the observed image and the high-resolution results. Owing to the complex nonlinear relationship of the PolSAR image degradation process, deep learning is introduced in this paper to fit its degradation process so as to better reconstruct low-resolution PolSAR images.

In order to improve the resolution of PolSAR data, an end-to-end network for PolSAR image super-resolution is proposed, as illustrated in Fig. 1. Fig. 2 is a specific description of the multi-channel residual structure and parameters. In the network, instead of interpolating the low-resolution PolSAR images as input directly, a deconvolutional layer is used in the proposed network. Subsequently, the complex structure block is used to extract the mixed features of complex values. Moreover, the network contains 20 convolutional layers with parametric rectified

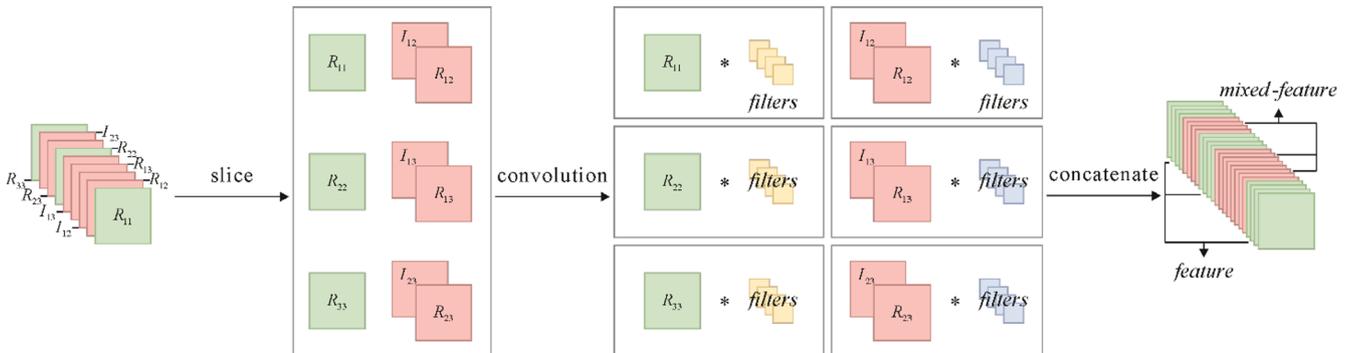


Fig. 4. The complex structure block.

Table 1
The Definitions and Configurations of the Parameters.

Parameter	Definition	Size
M	The size of the feature map for each kernel output	40×40
K_{de}	The size of the kernel in the deconvolutional layer	4×4
C_{inde}	Number of input channels in the deconvolutional layer	9
C_{outde}	Number of output channels in the deconvolutional layer	9
K_{cb}	The size of the kernel in the complex structure block	3×3
C_{incb}	Number of input channels in the complex structure block	9
C_{outcb}	Number of output channels in the complex structure block	384
K_l	The size of the kernel in the convolutional layer	3×3
$C_{inconvf}$	Number of input channels in the first convolutional layer	384
$C_{outconvf}$	Number of output channels in the first convolutional layer	64
C_{l-1}	Number of input channels in the l th convolutional layer	64
C_l	Number of output channels in the l th convolutional layer	64
$C_{inconvl}$	Number of input channels in the last convolutional layer	64
$C_{outconvl}$	Number of output channels in the last convolutional layer	9
D	The depth of the convolutional layer	19

linear units (PRELU) and one convolutional layer without an activation function. The details of the super-resolution framework are given in the following.

2.2. Deconvolution

In consideration of the nonlinear relationship between the high-resolution imagery and the degraded imagery, it is not suitable to use a simple up-sampling operator to up-sample low-resolution PolSAR images. The predefined up-sampling, which is fixed and cannot be learned, is inflexible and may generate new noise (Haris et al., 2018). In addition, due to the negative values in the PolSAR imagery, if the up-sampling is performed directly using the predefined up-sampling operator, the sign of the up-sampling result may be abnormal. Hence, we tend to up-sample low-resolution images in a non-fixable and learnable way. Deconvolution (Zeiler et al., 2010), which is also known as transposed convolution, can be regarded as a learnable up-sampling technique. Compared to bilinear up-sampling, the filters in the deconvolutional layers are flexibly learned, along with the different polarimetric data.

For the deconvolutional layer, the up-sampling process can be defined as follows:

$$F_{deconv} = W \circ F_{input} + b \quad (5)$$

where F_{deconv} is a feature map up-sampled by the deconvolutional layer. W represents the convolutional kernel. The size of W is $C_{in} \times S_{filter} \times S_{filter} \times C_{out}$, where S_{filter} is the size of a filter, C_{in} and C_{out} represent the number of channels of the deconvolutional layer input and output, respectively. b represents the bias term of the convolution which is a C_{out} -dimensional vector. F_{input} represents the input feature map. The size of the deconvolutional feature map F_{deconv} is a_{scale} times the size of the input feature map F_{input} , where a_{scale} is the scale factor of the super-resolution. \circ is the convolution operation. The deconvolution scheme is depicted in Fig. 3.

Since the deconvolution is done inside the network, the utilization of the deconvolutional layer can not only automatically learn the

Table 2
The number and proportion of pixels used by each sensor for training and testing.

Data	Training (num of pixels)	Testing (num of pixels)	Training (proportion)	Testing (proportion)
RADARSAT-2 (San Francisco)	37,273,600	2,867,200	59.52%	4.58%
RADARSAT-2 (Vancouver)	0	18,874,368	0.00%	30.14%
ESAR (Munich)	0	2,168,320	0.00%	3.46%
PiSAR (Niigata)	0	1,440,000	0.00%	2.30%
Total	37,273,600	25,349,888	59.52%	40.48%

sampling relationship between the low-resolution and high-resolution images, but it can also reduce the influence of the interpolation on the super-resolution results, without interpolation preprocessing.

2.3. Complex structure block

In Section 2.1, we mentioned that the non-diagonal elements of the coherency matrices are complex. For the complex structure in PolSAR imagery, the simplest strategy is to extract the values from the real and imaginary fields separately, then convert the complex numbers into two separate real values. Super-resolution can then be finally performed on the separate values. However, the strategy of separating the real and imaginary values from the complex values will destroy the complex features, to a certain extent.

Therefore, the complex structure block is proposed to solve the complex feature extraction problem. As illustrated in Fig. 4, according to the numerical attribute, T_{value} is sliced into three real parts (R_{11}, R_{22}, R_{33}) and three pairs of complex parts (R_{12}, I_{12}), (R_{13}, I_{13}), (R_{23}, I_{23}). The six portions respectively correspond to the values of the six positions of (2). Instead of directly extracting the features of the nine values of T_{value} , the features of the above six parts are learned in the proposed model. The proposed complex structure block contains two types of convolution. For the diagonal elements, common convolution is adopted for the feature extraction. For the real and imaginary values at the same position on the non-diagonal elements, mixed convolution is employed to extract the mixed features.

For the complex structure block layer, the convolutional process can be defined as follows:

$$F_R^i = W_{CBR}^i \circ F_{RS}^i + b_{CBR}^i, \quad i = 11, 22, 33 \quad (6)$$

$$F_C^i = W_{CBC}^i \circ F_{CS}^i + b_{CBC}^i, \quad i = 12, 13, 23 \quad (7)$$

where F_R^i is a real feature map; F_C^i is a complex feature map; and W_{CBR}^i , W_{CBC}^i are convolutional kernels of the real-valued feature map and complex-valued feature map, respectively. b_{CBR}^i and b_{CBC}^i are the bias terms of the real-valued feature map and complex-valued feature map, respectively. F_{RS}^i is the slice of the real-valued feature map, and F_{CS}^i is the slice of the complex-valued feature map. i corresponds to the subscripts of the upper triangular matrix in (2).

After the convolution, all the features are concatenated in the order before the slicing. The concatenation can be defined as follows:

$$F_{CB} = Concat(F_R^{11}, F_C^{12}, F_C^{13}, F_R^{22}, F_C^{23}, F_R^{33}) \quad (8)$$

where F_R^{11} , F_C^{12} , F_C^{13} , F_R^{22} , F_C^{23} , F_R^{33} represent the feature maps of the real-valued slice and the complex-valued slice after convolution.

The complex structure block, on the one hand, can extract the complex mixed features of the non-diagonal elements. On the other hand, the subsequent convolution operation is performed on the complex mixed-feature map, which can reduce the precision loss caused by separating the complex numbers.

2.4. The parametric rectified linear unit for complex values

The polarimetric coherency matrices have both positive and negative values. As an activation function, the unilateral suppression characteristics of the rectified linear unit (ReLU) (He et al., 2015) can cause

Table 3
Test datasets information.

Sensor	Region	Looks	Resolution	Size	Source
RADARSAT-2	San Francisco	1	6 m	1024 × 2800	Canadian Space Agency
RADARSAT-2	Vancouver	1	6 m	6144 × 3072	Canadian Space Agency
ESAR	Munich	4	3 m	1408 × 1540	European Space Agency
PiSAR	Niigata	4	3 m	1200 × 1200	European Space Agency

Table 4
The Proposed Network Configuration.

Layer	Configuration
Deconvolutional	Deconv + PReLU: $9 \times 4 \times 4$, stride = 2, pad = 1, type='bilinear'
Complex structure block	Conv + PReLU: $64 \times 3 \times 3$, stride = 1, pad = 1, type='msra'
	Conv + PReLU: $64 \times 3 \times 3$, stride = 1, pad = 1, type='msra'
Convolutional	Conv + PReLU: $128 \times 3 \times 3$, stride = 1, pad = 1, type='msra'
	Conv: $9 \times 3 \times 3$, stride = 1, pad = 1, type='msra'
	layer1-layer20
	layer21

Table 5
Quantitative Evaluation Results (San Francisco).

Method	Bicubic	SRPSC	MSSR	PSSR
MAE ($ P_1 ^2$)	2.06	1.71	<u>0.27</u>	0.26
MAE ($ P_2 ^2$)	2.12	1.82	<u>0.57</u>	0.37
MAE ($ P_3 ^2$)	2.23	1.85	<u>0.35</u>	0.34
MAE (mean)	2.14	1.79	<u>0.39</u>	0.32
PSNR ($ P_1 ^2$)	38.63	39.69	<u>51.58</u>	52.24
PSNR ($ P_2 ^2$)	38.14	38.87	<u>50.42</u>	52.03
PSNR ($ P_3 ^2$)	39.62	40.78	<u>50.50</u>	50.87
PSNR (mean)	38.80	39.78	<u>50.83</u>	51.71

zero values to abnormally appear in the negative number field and the gradient in the backward propagation. This phenomenon means that the weight of the filters is no longer updated, and the negative field may remain inactive during the subsequent training process. The negative information will then be lost. Hence, in this paper, we introduce the parametric rectified linear unit (PReLU) (He et al., 2015) as the activation function. Unlike the original ReLU, the PReLU is adapted to consider the numerical characteristics of SAR images. Compared with the ReLU, the PReLU has non-saturation in the negative domain. By adaptively learning parameters a_i , the negative domain is preserved. The PReLU can be defined as:

$$f_{prelu}(y_i) = \begin{cases} y_i, & y_i > 0 \\ a_i y_i, & y_i \leq 0 \end{cases} \quad (9)$$

where y_i represents the input of the PReLU f_{prelu} on the i th component, and a_i is a negative domain slope control coefficient. After the PReLU

processing, the data integrity of the coherency matrices of the PolSAR images is maintained, which can promote the accuracy of the super-resolution reconstruction.

Compared to natural images, due to the negative elements in both the real and imaginary parts of PolSAR image coherency matrices, the utilization of the PReLU in the PolSAR image super-resolution can not only correct the data distribution but, more importantly, it can maintain the details of the negative domain. Hence, the PReLU shows better adaptability in the super-resolution of PolSAR images.

2.5. Polarimetric loss function

In the proposed method, since the input is a low-resolution image, without up-sampling, the size of the input and output is inconsistent. Therefore, instead of using the input and output, the results of the deconvolution and output are used as the identity shortcut. In the proposed network, T_D denotes the deconvolution result of the low-resolution image T_y , and the residual can be defined as follows:

$$\rho = T_x - T_D \quad (10)$$

where T_x is the high-resolution PolSAR image.

Given a training set with N training image pairs $\{T_x^i, T_y^i\}_{i=1}^N$, T_x^i is the high-resolution PolSAR image patch, and T_y^i is the low-resolution PolSAR image patch. In the proposed network, we employ the mean-squared error (MSE) between the patch residual ρ^i and the output of the network $\xi(T_y^i)$ as the loss function. The loss function is defined as follows:

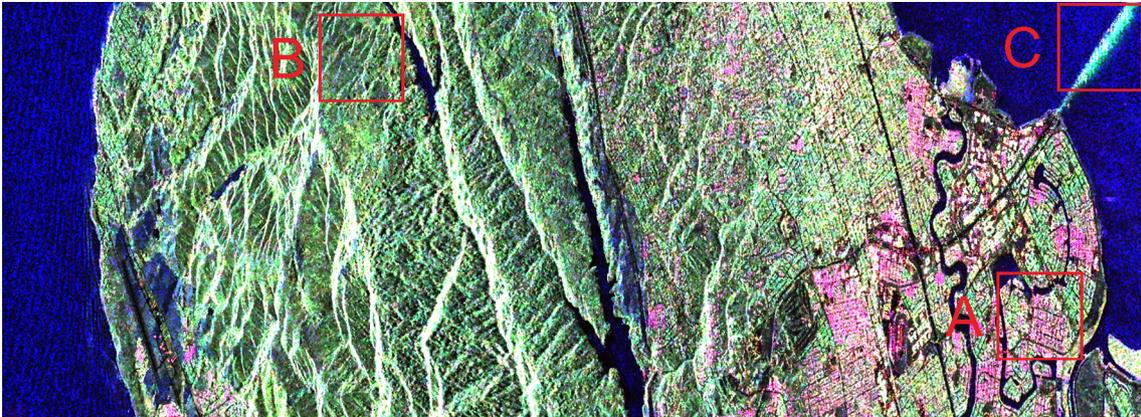


Fig. 5. The PSSR super-resolution result for the RS-2 imagery (San Francisco).

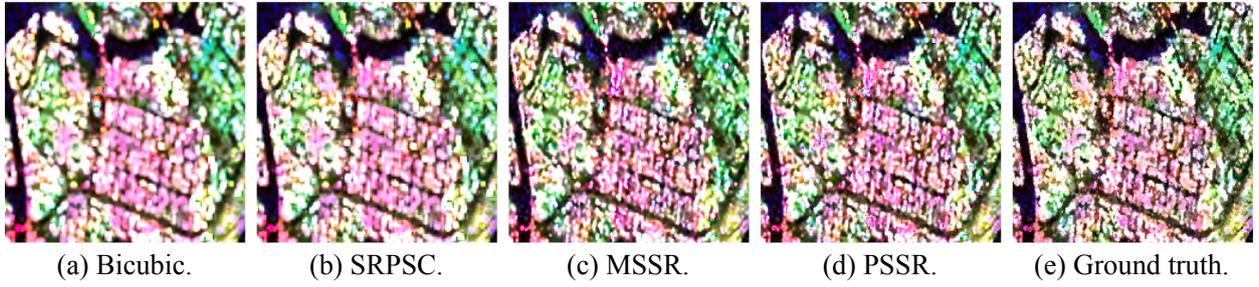


Fig. 6. PolSAR image super-resolution performance comparison for the urban in San Francisco.

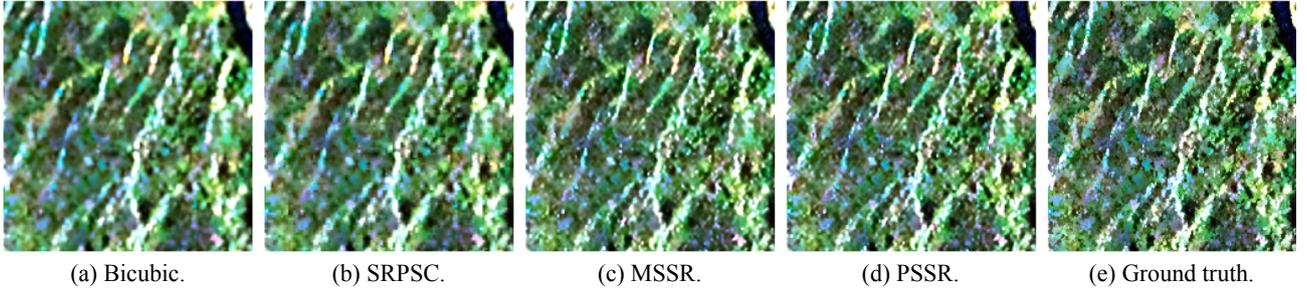


Fig. 7. PolSAR image super-resolution performance comparison for the forest in San Francisco.

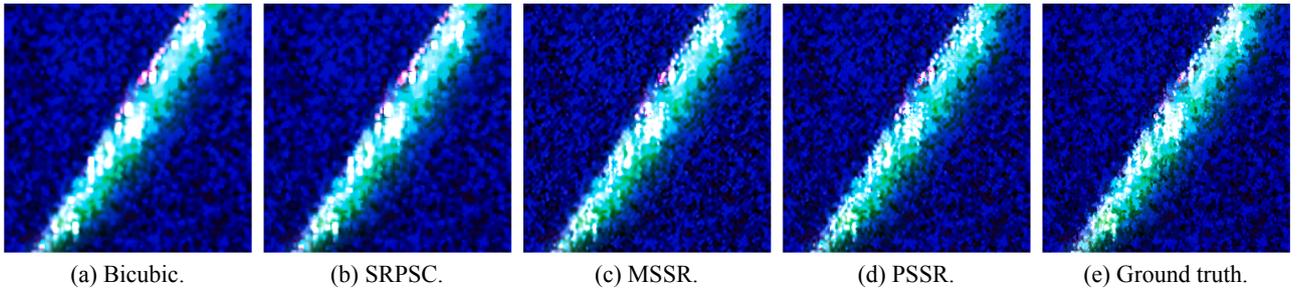


Fig. 8. PolSAR image super-resolution performance comparison for the sea in San Francisco.

$$L(\Theta) = \frac{1}{2N} \sum_{i=1}^N \|\rho^i - \xi(T_y^i)\|_2^2 \quad (11)$$

where Θ is the parameters of the proposed network. The proposed network uses the training dataset to simulate the relationship between the high-resolution and low-resolution PolSAR image patches, then trains a learned network with a converged loss.

2.6. Complexity analysis

In the following, the complexity of the algorithm is analyzed from the aspects of time complexity and spatial complexity.

1) Time complexity analysis

The time complexity of the proposed network framework is as follows:

$$Time \sim O \left(\begin{array}{l} M \cdot K_{de} \cdot C_{de_in} \cdot C_{de_out} \\ + M \cdot K_{cb} \cdot C_{cb_in} \cdot C_{cb_out} \\ + M \cdot K_l \cdot C_{convf_in} \cdot C_{convf_out} + M \cdot K_l \cdot C_{convl_in} \cdot C_{convl_out} \\ + \sum_{l=1}^D M \cdot K_l \cdot C_{l-1} \cdot C_l \end{array} \right) \quad (12)$$

The definitions and configurations of the parameters are given in Table 1.

2) Spatial complexity analysis

The spatial complexity of the network includes the total amount of parameters and the output feature map. The spatial complexity of the proposed network framework is as follows:

$$Space \sim O \left(\begin{array}{l} K_{de} \cdot C_{inde} \cdot C_{outde} + M \cdot C_{outde} \\ + K_{cb} \cdot C_{incb} \cdot C_{outcb} + M \cdot C_{outcb} \\ + K_l \cdot C_{inconvf} \cdot C_{outconvf} + M \cdot C_{outconvf} + K_l \cdot C_{inconvl} \cdot C_{outconvl} \\ + M \cdot C_{outconvl} \\ + \sum_{l=1}^D K_l \cdot C_{l-1} \cdot C_l + \sum_{l=1}^D M \cdot C_l \end{array} \right) \quad (13)$$

The definition and configuration of the parameters are given in Table 1.

3. Experimental results

3.1. Experimental data

All the data used in this study were PolSAR data with four polarimetric channels (HH, HV, VH, VV). The data from the San Francisco region (RADARSAT-2) were C-band data, for which the size was 14416 × 2823 and the resolution was about 6 m. The data from the San Francisco region were divided into two parts, nine-tenths for network

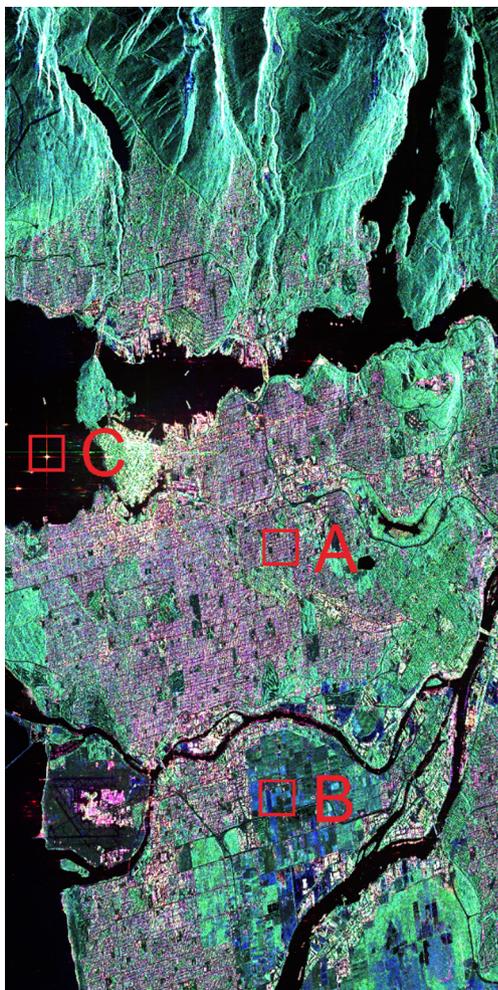


Fig. 9. The PSSR super-resolution result for the RS-2 imagery (Vancouver).

training and one-tenth for network testing. The data from the Vancouver region (RADARSAT-2) were C-band data, for which the size was 6144×3072 and the resolution was about 6 m. The data from the Munich region (ESAR) were L-band data, for which the size was 1408×1540 and the resolution was about 3 m. The data from the Niigata region of Japan (PiSAR) were L-band data, for which the size was 1200×1200 and the resolution was about 3 m. For all the data used in this paper, 60% were used for network training and 40% for testing. The proportion of training and test data for each sensor is shown in Table 2. All the data contained a variety of land-cover types, such as urban areas, mountains, vegetation, and water. For the speckle noise, the original PolSAR image was despeckled using the non-local means polarimetric speckle filter in PolSARpro version 5.0, before the experiments. The data were then converted into the form of a coherency matrix.

In the network training process, only RADARSAT-2 (RS-2) data were used as training data. The size of the high-resolution patches was set to 40×40 , and the size of the corresponding low-resolution patches was set to 20×20 . In total, 26,048 pairs of training images were used. The elements on the diagonal of the matrix data used in the experiments were positive real numbers, and the non-diagonal elements were complex and contained both positive and negative numbers.

We used four scenes of PolSAR data as the test data, for which the size of each scene was 1024×2800 , 6144×3072 , 1408×1540 , and 1200×1200 , respectively. The test data were divided into three groups and used in three sets of experiments. The effectiveness of the proposed method for the super-resolution reconstruction of PolSAR images was verified from the following three aspects. Details of the test datasets are provided list in Table 3.

1) *The same sensor and the same imaging conditions*

This set of experiments involved one scene of PolSAR images from RS-2, which were acquired under the same imaging condition as the training data.

2) *The same sensor but different imaging conditions*

This set of experiments involved one scene of PolSAR data from the Vancouver region, which were acquired by RS-2 under different imaging conditions, including different imaging times and different imaging angles.

3) *Different sensors and different imaging conditions*

This set of experiments involved two scenes of PolSAR images, one of which was acquired by ESAR and the other scene was acquired by PiSAR.

Table 6
Quantitative Evaluation Results (Vancouver).

Method	Bicubic	SRPSC	MSSR	PSSR
MAE ($ P_1 ^2$)	1.17	<u>1.00</u>	0.21	0.21
MAE ($ P_2 ^2$)	0.87	0.76	<u>0.17</u>	0.16
MAE ($ P_3 ^2$)	1.67	1.45	<u>0.28</u>	0.25
MAE (mean)	1.24	1.07	<u>0.22</u>	0.21
PSNR ($ P_1 ^2$)	40.60	41.36	<u>51.84</u>	52.14
PSNR ($ P_2 ^2$)	40.99	41.56	<u>51.78</u>	52.30
PSNR ($ P_3 ^2$)	40.55	41.31	<u>51.88</u>	52.52
PSNR (mean)	40.71	41.41	<u>51.83</u>	52.32

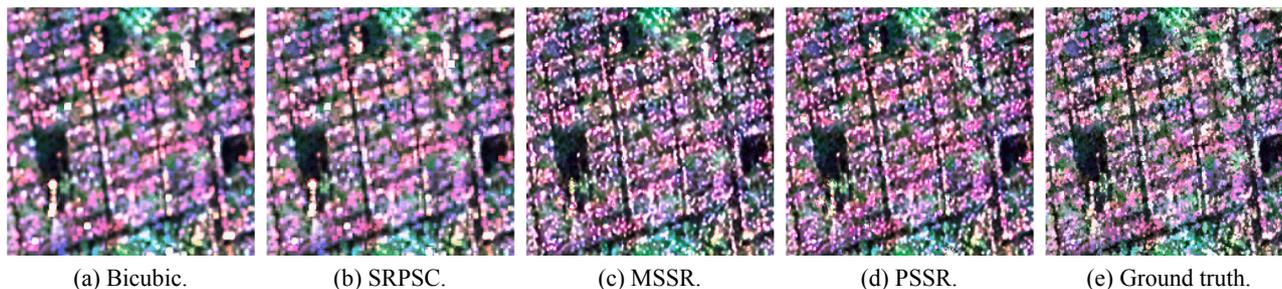


Fig. 10. PolSAR image super-resolution performance comparison for the urban in Vancouver.

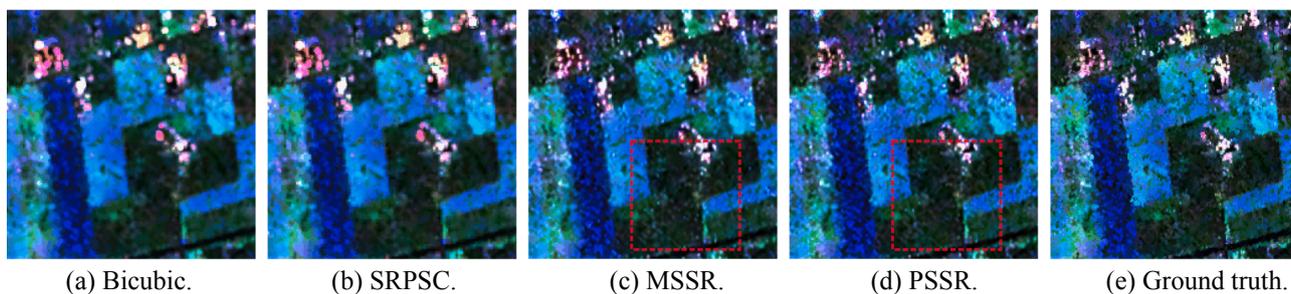


Fig. 11. PolSAR image super-resolution performance comparison for the farmland in Vancouver.

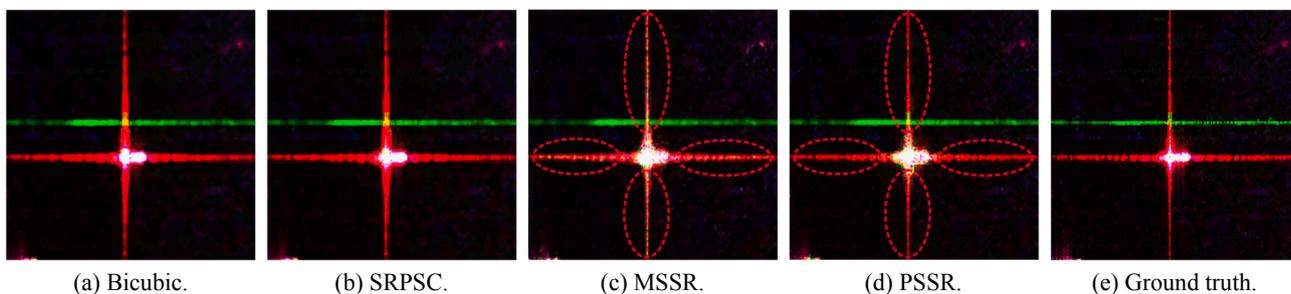


Fig. 12. PolSAR image super-resolution performance comparison for the strong scattering points in Vancouver.

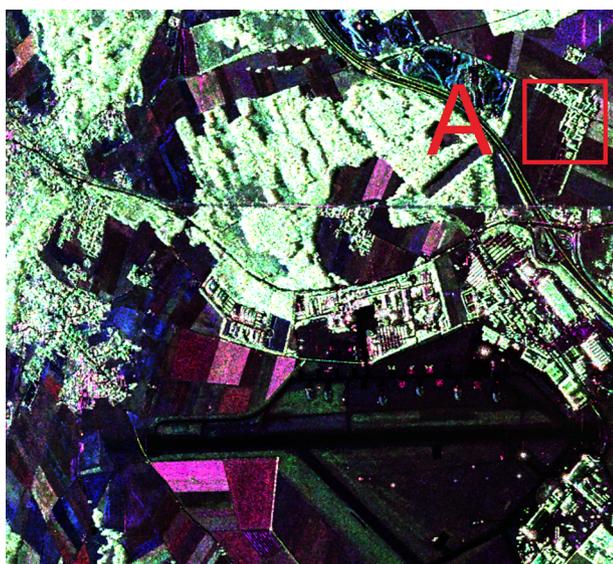


Fig. 13. The PSSR super resolution result for the ESAR data.

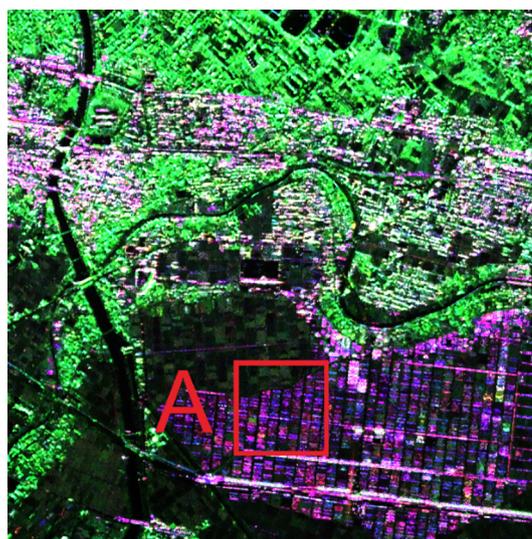


Fig. 14. The PSSR super resolution result for the PiSAR data.

3.2. Parameter settings

In the proposed network, the weights of the deconvolutional filters were initialized by bilinear interpolation. The proposed network uses a 21-layer convolutional structure, in which the first 20 layers are the convolutional layers with the PReLU, and the last layer is a convolutional layer. In addition, we initialized all the convolutional filters using “MSRA” (He et al., 2015). The detailed configuration of the proposed network is provided in Table 4. The PReLU was used with an initial negative slope a_i of 0.25. The learning rate α was initialized to 0.0001 for the whole network, which was fixed. In the network training process, the maximum number of iterations was set to 1000000. The Adam algorithm (Kingma and Ba, 2015) was used for the optimization in the network training. We used the Caffe framework to train the proposed method in the Windows 10 environment, 64 GB RAM, and an Nvidia Quadro P4000 GPU. During network training, 64 patches were trained in parallel.

Table 7
Quantitative Evaluation Results (ESAR).

Method	Bicubic	SRPSC	MSSR	PSSR
MAE ($ P_1 ^2$)	2.93	2.46	<u>1.25</u>	1.20
MAE ($ P_2 ^2$)	3.12	2.62	0.80	<u>0.90</u>
MAE ($ P_3 ^2$)	4.20	3.46	1.64	<u>1.81</u>
MAE (mean)	3.41	2.85	1.23	<u>1.30</u>
PSNR ($ P_1 ^2$)	38.23	39.18	<u>48.33</u>	48.67
PSNR ($ P_2 ^2$)	38.44	39.33	<u>49.06</u>	49.69
PSNR ($ P_3 ^2$)	37.18	38.01	46.11	<u>45.61</u>
PSNR (mean)	37.95	38.84	<u>47.83</u>	47.99

3.3. Quantitative Evaluation indicators

In order to verify the effectiveness of the proposed method, we compared it with bicubic interpolation (Keys, 1981), SRPSC (Zhang et al., 2011), and Multi-channel SAR Super Resolution (MSSR). Among

Table 8
Quantitative Evaluation Results (PiSAR).

Method	Bicubic	SRPSC	MSSR	PSSR
MAE ($ P_1 ^2$)	0.73	0.65	<u>0.23</u>	0.20
MAE ($ P_2 ^2$)	0.56	0.50	<u>0.14</u>	0.15
MAE ($ P_3 ^2$)	1.94	1.73	<u>1.55</u>	1.43
MAE (mean)	1.08	0.96	<u>0.64</u>	0.59
PSNR ($ P_1 ^2$)	41.09	41.68	<u>48.48</u>	49.24
PSNR ($ P_2 ^2$)	42.93	43.39	<u>50.04</u>	49.94
PSNR ($ P_3 ^2$)	42.60	43.30	<u>44.72</u>	45.84
PSNR (mean)	42.21	42.79	<u>47.75</u>	48.34

the comparison algorithms, the MSSR method used is a multi-channel VDSR (Kim et al., 2016) network for polarimetric SAR image super-resolution, which were adapted for use with PolSAR images.

In the experiments, Pauli decomposition for the coherency matrix was performed after the super-reconstruction. The obtained Pauli decomposition results were then used to calculate the quantitative indices. The following is the equation for Pauli decomposition.

$$P_1 = \frac{S_{HH} + S_{VV}}{\sqrt{2}}, P_2 = \frac{S_{HH} - S_{VV}}{\sqrt{2}}, P_3 = \frac{2S_{HV}}{\sqrt{2}} \quad (14)$$

The three components after Pauli decomposition correspond to the three scattering mechanisms, where P_1 represents single-bounce scattering from a rough surface, P_2 represents double-bounce scattering, which usually comes from buildings, and P_3 represents volume scattering from the canopy. $|P_1|^2$, $|P_2|^2$, $|P_3|^2$ stand for the scattered power by single-bounce scattering, double-bounce scattering, and volume scattering, respectively.

The following indicators are used to quantitatively evaluate the super-resolution accuracy of the PolSAR images: mean absolute error (MAE) and peak signal-to-noise ratio (PSNR). The MAE is used to reflect the absolute error between the predicted and actual images. The PSNR is the most widely used indicator for evaluating the degree of image distortion. The lower the value of MAE, the closer the reconstruction result is to the ground truth, and the higher the value of PSNR, the better the super-resolution result is. The phase of the non-diagonal elements of the coherency matrix was also calculated to measure the polarimetric information retention. In addition, the equivalent number of looks (ENL) (Anfinson et al., 2009; Ren et al., 2019) was calculated and used to check whether the super-resolution reconstruction process introduces artifacts. The ENL can be defined as follows:

$$ENL = \frac{Tr(\langle T_3 \rangle)^2}{\langle Tr(T_3 T_3) \rangle - Tr(\langle T_3 \rangle \langle T_3 \rangle)} \quad (15)$$

where T_3 is the coherency matrices of the PolSAR imagery, and $Tr(\cdot)$ is the trace operator.

3.4. Experimental analysis

For the proposed method, the scale factor was set to 2. Three sets of experiments were undertaken to quantitatively and qualitatively verify

the proposed method.

1) The same sensor and the same imaging conditions

The first set of experiments involved the San Francisco area data acquired by RS-2. As shown in Table 5, the proposed framework shows a superior performance in the quantitative evaluation. Compared with bicubic interpolation, the PSNR of the proposed method is significantly increased. At the same time, the MAE of the proposed method is reduced to around 0.5 or even lower. Compared with the traditional SRPSC method, the proposed framework shows a state-of-the-art performance in the super-resolution of PolSAR images.

We evaluated three typical targets (see Fig. 5): urban area, forest, and sea. In the urban part, as shown in Fig. 6, the bicubic interpolation and SRPSC methods show an over-smoothing effect, as a whole, and the edge of the urban structure is blurred. The structural information of the two deep-learning-based methods is richer. Compared with MSSR, the result of the proposed method is sharper at the edges, and the details are more obvious. In the forest part, as shown in Fig. 7, the proposed method performs better than the other algorithms in reconstructing vegetation texture information, and the ridgeline is clearer. In the sea part, as shown in Fig. 8, the sharpness of the edge of the bridge in the sea is higher than that of other comparison algorithms, and the structure is closer to the ground truth, indicating that the proposed method shows a superior performance in detail reconstruction.

2) The same sensor but different imaging conditions.

The second set of experiments involved the RS-2C-band data from the Vancouver area as shown in Fig. 9. The imaging conditions in the Vancouver area were different from the conditions in the San Francisco area. As shown in Table 6 and Figs. 10–12, the two deep-learning-based methods show a significant improvement over the traditional SRPSC super-resolution method, in both quantitative indicators and visual results. As shown in Table 6, the quantitative results of the proposed method are superior to those of the comparison algorithms, indicating that the proposed method results in fewer errors and can better preserve texture information.

In this set of experiments, three typical targets: urban, farmland, and strong scatter points were evaluated. In the urban part, as shown in Fig. 10, the results allow a similar conclusion to the first set of experiments, and the proposed method reconstructs more detailed information. In the farmland part, as shown in Fig. 11, the results of the bicubic interpolation and SRPSC methods are generally smoother. However, the boundaries between the different crops are blurred. The details in the results of the MSSR and PSSR methods are better, and the boundaries between the different crops are clear. Compared to MSSR, PSSR results in fewer artifacts, and MSSR has obvious point-like artifacts as shown by the dotted rectangle in Fig. 11. We also selected strong scattering points and evaluated them. Strong scattering points are point targets with strong scattering intensity, and such point targets are susceptible to defocusing and blurring, due to the degradation processes.

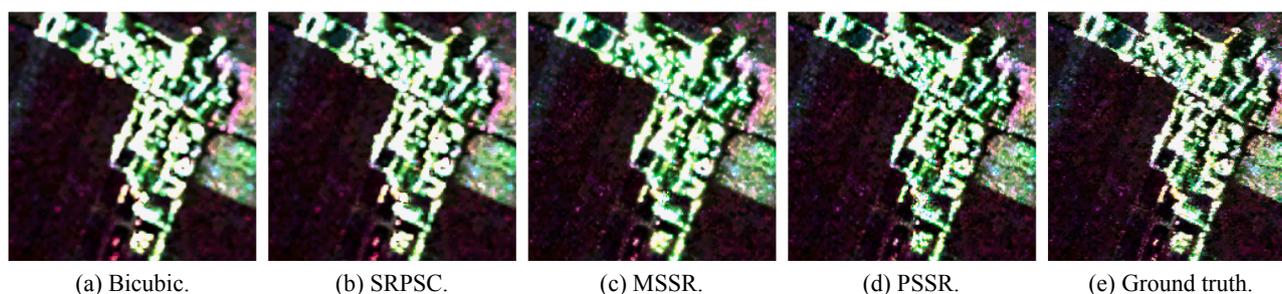


Fig. 15. PolSAR image super-resolution performance comparison for the building in Munich.

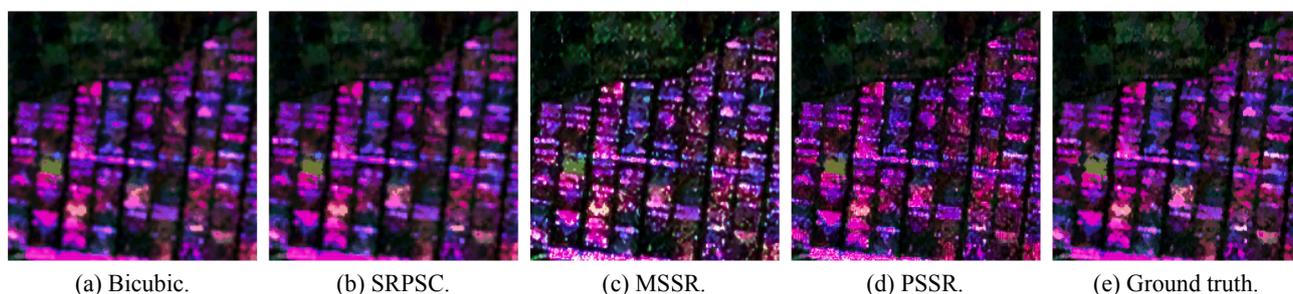


Fig. 16. PolSAR image super-resolution performance comparison for the building in Niigata.



Fig. 17. The PSSR super-resolution result for the RS-2 imagery (Vancouver).

As shown in Fig. 12, for strong scattering points in the sea area, in the results of the bicubic interpolation and SRPSC methods, defocusing occurs at strong scattering points, and the strong scattering points of SRPSC show a grid effect in the reconstruction of the strong scattering

points. The reconstruction of the two deep-learning-based methods has a higher sharpness, but does contain certain artifacts at the same time, which may be related to the small number of target samples of such strong scattering points in the training data. The results of the MSSR method also show a certain color distortion in the point target extension, as shown by the dotted oval in Fig. 12, whereas the PSSR method results in less color distortion.

3) Different sensors and different imaging conditions

In the third set of experiments, one scene of 1408×1540 ESAR data and one scene of 1200×1200 PiSAR data were used as test data to verify the effectiveness of the proposed method for non-RADARSAT-2 data as shown in Fig. 13 and 14.

Although the training data did not include data from the above two sensors, the proposed method still has achieves a good reconstruction effect in the super-resolution reconstruction of non-RADARSAT-2 data as shown in Table 7 and 8. As shown in Fig. 15, for the building area, the bicubic interpolation, SRPSC, and MSSR methods show varying degrees of smoothing, while the PSSR method shows a better generalization performance, and the details of the building area are well preserved. Similar conclusions can also be made from Fig. 16. Also, as shown in Fig. 16, the results of the MSSR contains speckle-like artifacts. The results reconstructed by the proposed method have fewer artifacts while reconstructing more details, and the overall results are closer to the original image.

In the three sets of experiments, the resolution of the RADARSAT-2 ground truth was 6 m, and the resolution of the ESAR and PiSAR ground truth was 3 m. The degraded image resolutions of Radarsat-2, ESAR, and PiSAR were 12 m, 6 m, and 6 m, respectively. With super-resolution, we attempt to enhance the details of low-resolution images so that their resolution is as close as possible to the original high-resolution image. When compared to the ground truth, the experiments proved that the details of the texture structure in the super-resolution results of the proposed method are very close to the ground truth. Hence, for the RADARSAT-2 test data set, the resolution of the super-resolution results is close to 6 m, and on the ESAR and PiSAR test data sets, the resolution of the super-resolution results is close to 3 m.

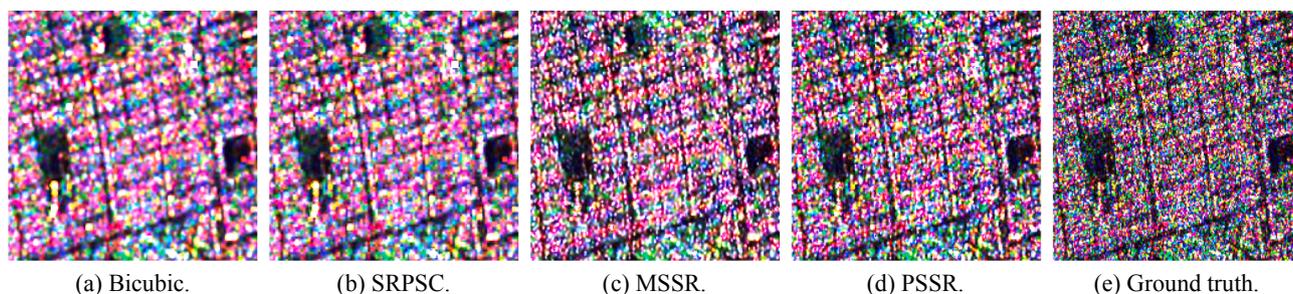


Fig. 18. PolSAR image super-resolution performance comparison for the urban in Vancouver.

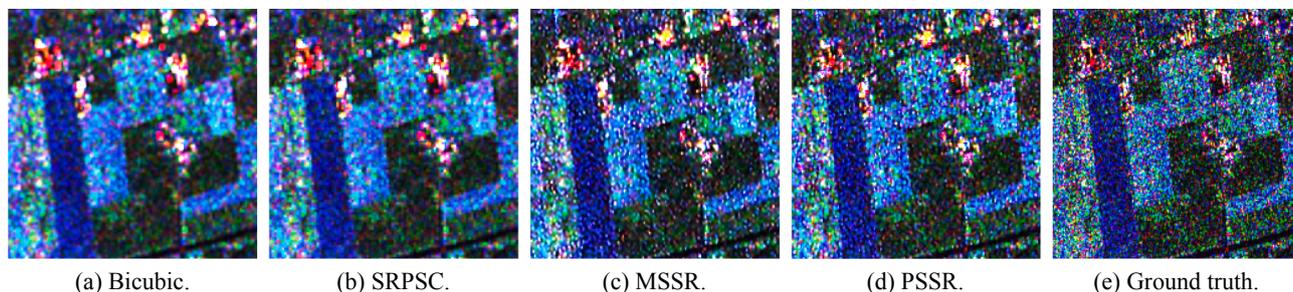


Fig. 19. PolSAR image super-resolution performance comparison for the farmland in Vancouver.

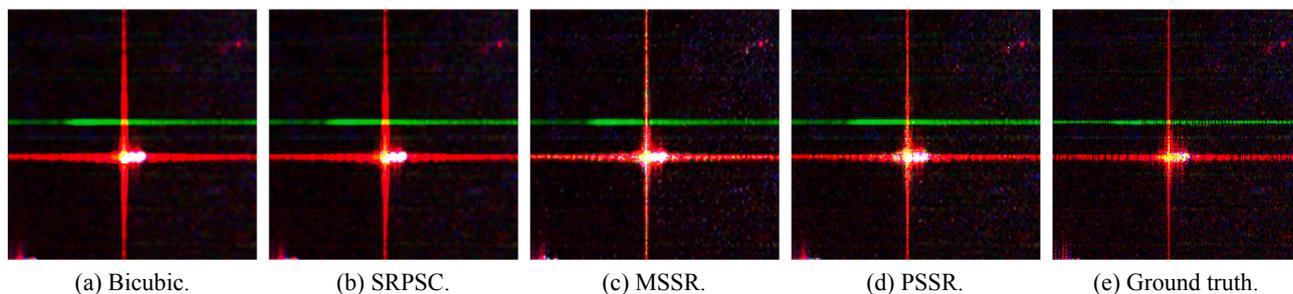


Fig. 20. PolSAR image super-resolution performance comparison for the strong scattering points in Vancouver.

Table 9

Quantitative Evaluation Results for the PolSAR Images containing Speckle (Vancouver).

Method	Bicubic	SRPSC	VDSR	PSSR
MAE ($ P_1 ^2$)	3.16	3.01	<u>2.80</u>	2.70
MAE ($ P_2 ^2$)	2.17	2.07	<u>1.99</u>	1.92
MAE ($ P_3 ^2$)	0.90	0.87	<u>0.87</u>	0.86
MAE (mean)	2.08	1.98	<u>1.89</u>	1.83
PSNR ($ P_1 ^2$)	31.43	31.63	<u>31.25</u>	31.71
PSNR ($ P_2 ^2$)	31.78	31.99	<u>31.21</u>	31.53
PSNR ($ P_3 ^2$)	38.53	38.76	38.07	<u>38.05</u>
PSNR (mean)	33.91	34.13	<u>33.51</u>	33.76

4) Experiments with PolSAR image containing speckle noise

To some extent, the speckle noise in the PolSAR imagery contains some of the polarimetric information. Therefore, in these experiments, PolSAR imagery with speckle noise was reconstructed. From the quantitative evaluation and visual assessment, it can be seen that the proposed method is still effective for PolSAR images containing speckle noise. The quantitative indicators and visual assessment of the results of the proposed method for PolSAR images containing speckle are superior to the results of the traditional methods. In the following, we provide the super-resolution results for the experiments with PolSAR images containing speckle noise as shown in Fig. 17–20 and Table 9.

In these experiments, the trained model was trained with despeckled data. For the super-resolution of PolSAR images containing speckle, the trained model was used directly. The results show that the speckle noise of the PolSAR images is suppressed while the spatial resolution of the PolSAR images is also improved.

4. Further discussion

4.1. Evaluation of the polarimetric information retention

To further evaluate the proposed method in the preservation of polarimetric information, phase calculation of the T12 channel was performed, which is more commonly used in applications in the coherency matrix. For the RADARSAT-2 dataset from the San Francisco

region, the frequency of the phase appearance was then counted.

In Fig. 21(a) and (b), compared with the phase of the original image, the phase of the results reconstructed by the bicubic interpolation method and the SRPSC method is significantly changed, indicating that the polarimetric information changes during the reconstruction process. The phase changes of two deep-learning-based methods are smaller than those of the traditional methods. The phase change result of the proposed method is the smallest among all the methods, and the result is closer to the original image. Overall, the proposed method shows a better effect on the preservation of polarimetric information.

It is important to evaluate the polarimetric information retention of different targets in the super-resolution reconstruction. Hence, the polarimetric response of different land covers was analyzed. The polarimetric response was analyzed for the four common land cover types: sea, vegetation, bare soil, and urban.

On the RADARSAT-2 data from the San Francisco region, four types of land cover types were selected, and we compared the polarimetric response of the super-resolution result with the polarimetric response of the ground truth as shown in Fig. 22–25. In all four types of land cover, the polarimetric response of the super-resolution result is well maintained, and the polarimetric response of the ground truth is basically consistent.

4.2. Evaluation of the polarimetric decomposition

In order to verify the validity of the PolSAR super-resolution results reconstructed by the proposed method, polarimetric decomposition was used to decompose the PolSAR image after super-resolution. In the experiment, the Four-Component Scattering Power Decomposition With Rotation of Coherency Matrix (Y4R) (Yamaguchi et al., 2011) was used to decompose the PolSAR image into the odd bounce scattering, the double bounce scattering, the volume scattering and helix scattering.

In the experiment, three areas were selected for analysis, including forest (area A), dense built-up area (area B) with complex scattering mechanisms, and airport (area C). As shown in Fig. 27, for the forest area, volume scattering shows stronger energy, while odd bounce scattering and double bounce scattering have weaker energy. Since helix scattering is a form of scattering unique to man-made targets, and

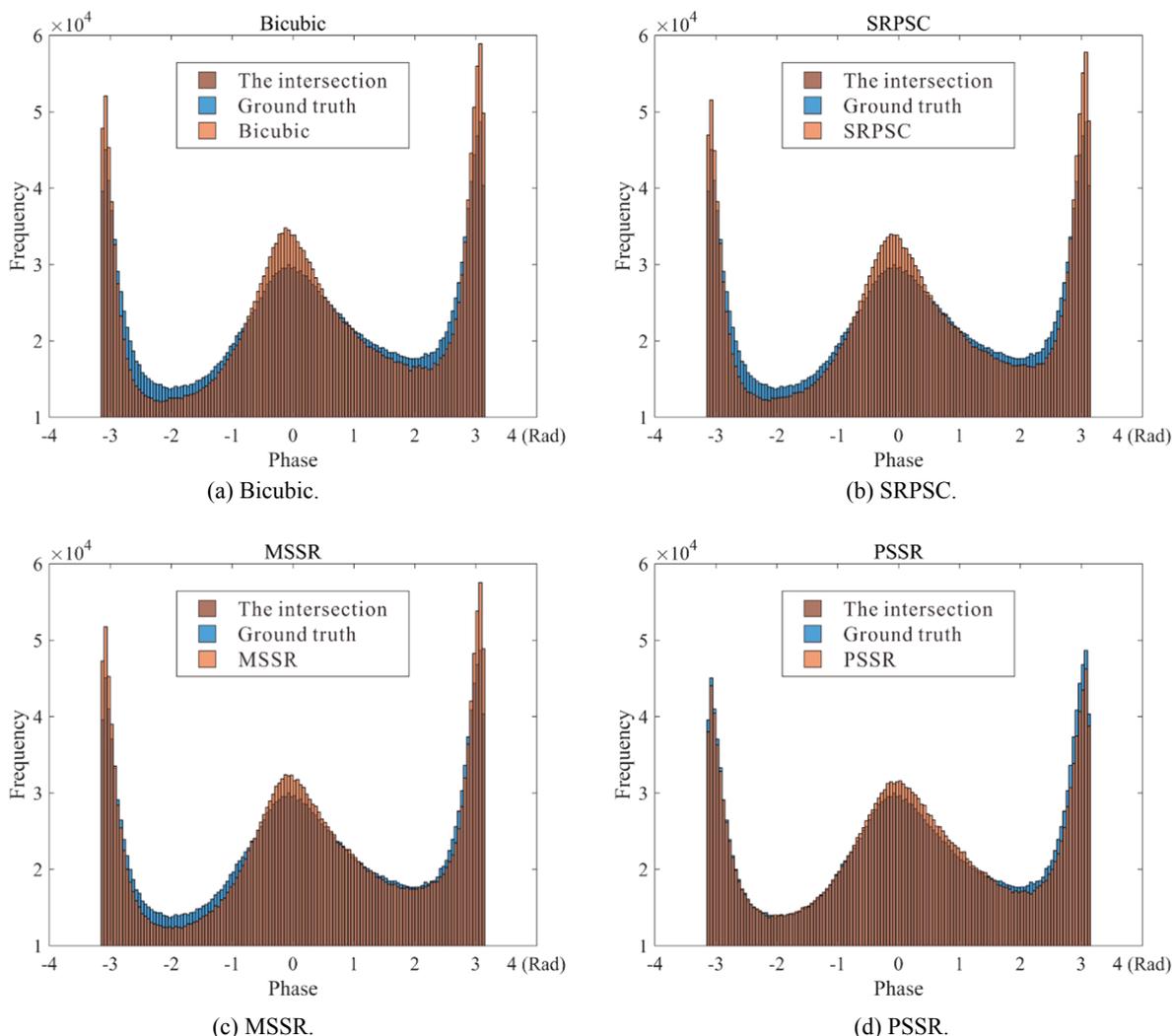


Fig. 21. Phase statistics histograms.

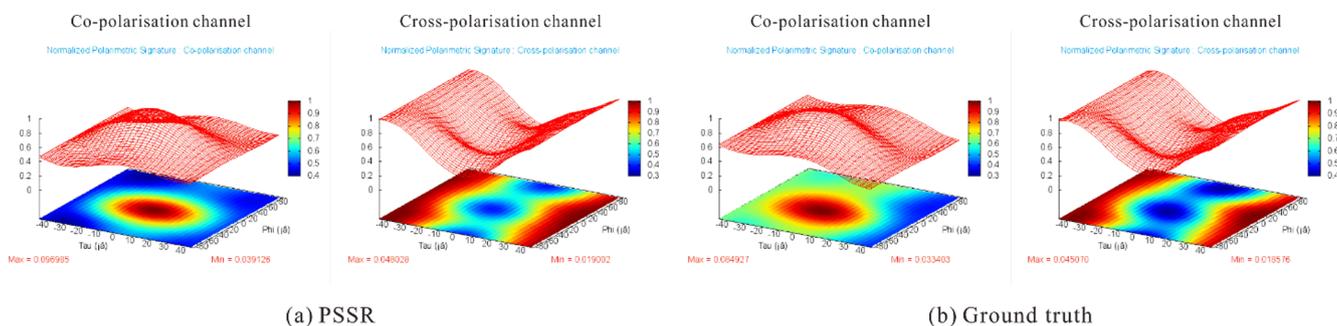


Fig. 22. Polarimetric response in the sea.

its energy is close to zero in this area. In Fig. 28, for building areas, double bounce scattering is the main scattering mechanism. In addition, compared to non-building areas, helix scattering has relatively strong energy in dense building areas. In Fig. 29, due to the large flat soil in the airport area, single-bounce scattering is strong. The existence of large buildings also results in strong energy of double bounce scattering in the area. At the same time, there is helix scattering in some areas. In addition, the volume scattering energy is low.

As shown in the Figs. 26–29, the scattering mechanism of the proposed method is consistent with the scattering mechanism of ground-

truth. Also, the polarimetric decomposition scattering intensity of the proposed method is close to ground-truth's.

4.3. Evaluation of the trace moment ENL

In order to verify whether artifacts are introduced into the homogenous regions during the super-resolution process, the artifacts were checked using the ENL. The ENL is often used to quantitatively evaluate the effect of despeckling in PolSAR images. The higher the ENL, the smoother the homogenous regions of the image. In this study, the ENL

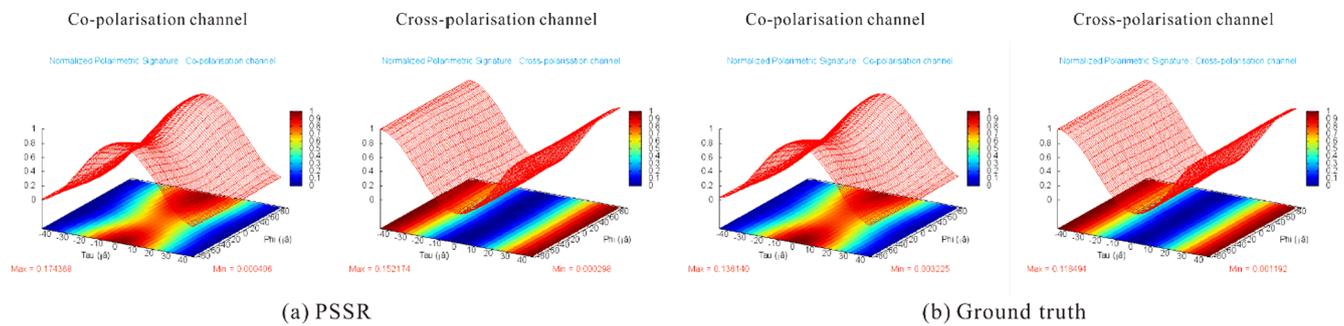


Fig. 23. Polarimetric response in the vegetation.

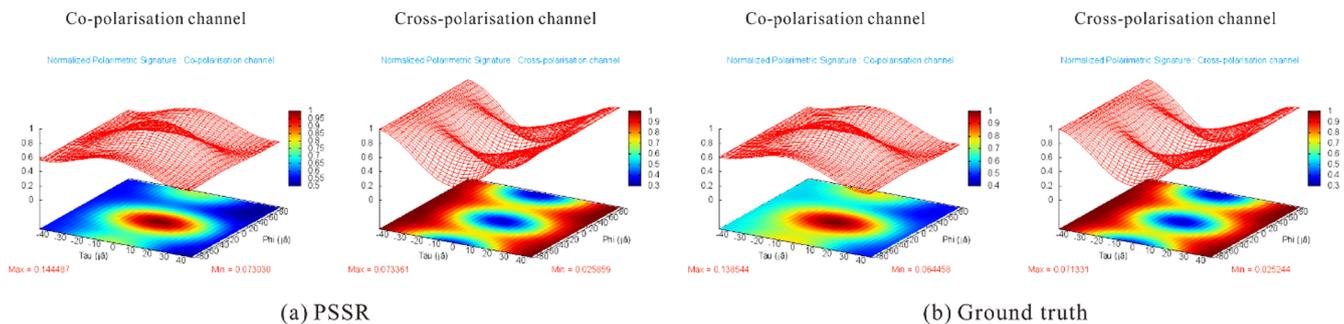


Fig. 24. Polarimetric response in the bare soil.

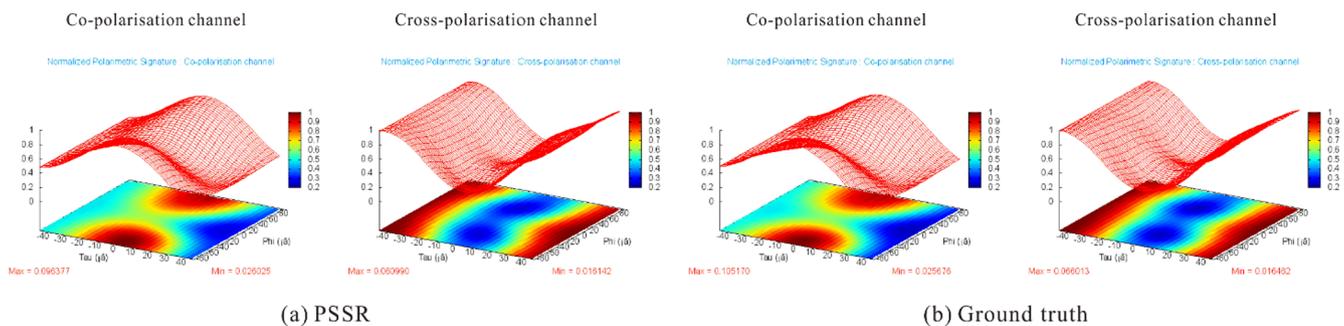


Fig. 25. Polarimetric response in the urban.

after super-resolution was calculated and compared with the ENL of the ground truth. If the ENL value after super-resolution reconstruction is lower than the ENL value of ground truth in the same position, it means that unnecessary artifacts have been produced.

ESAR images with a high spatial resolution were selected, and ENL was calculated. The ENL results for PSSR and the ground truth are shown in Fig. 30. Areas A and B covering an urban and farmland, respectively, were selected for further analysis of the ENL. As shown in Fig. 31(a) and (b), in area ‘A’, the ENL of the PSSR method is higher than that of the ground truth in the homogeneous region, indicating that unnecessary artifacts were not introduced in the super-resolution reconstruction process. In the heterogeneous region, the ENL of the PSSR method is equivalent to the ENL of the ground truth, indicating that the details of the heterogeneous region are well preserved. In area ‘B’, both PSSR and ground truth have low ENL values in the urban region, which means that PSSR reconstructs texture information that is

close to the ground truth. In the farmland area, the ENL of PSSR is higher than that of the ground truth, which means that the super-resolution reconstruction process did not introduce obvious artifacts.

4.4. The experiment in residual compensation

Due to the complexity of PolSAR imaging systems, there may be some differences in the data acquired in different flight angles at different times, so it is difficult for training samples to contain all the types of features. However, fewer types of training samples may cause network misprediction problems, in that the information reconstruction is not enough or artifacts are produced. Inspired by the downscaling of imagery in the geosciences (Wang et al., 2015; Wang and Atkinson, 2018), a simple and effective complex structure based residual compensation (RC) strategy is proposed for the post-processing of super-resolution results, as depicted in Fig. 32. In the RC strategy, the network

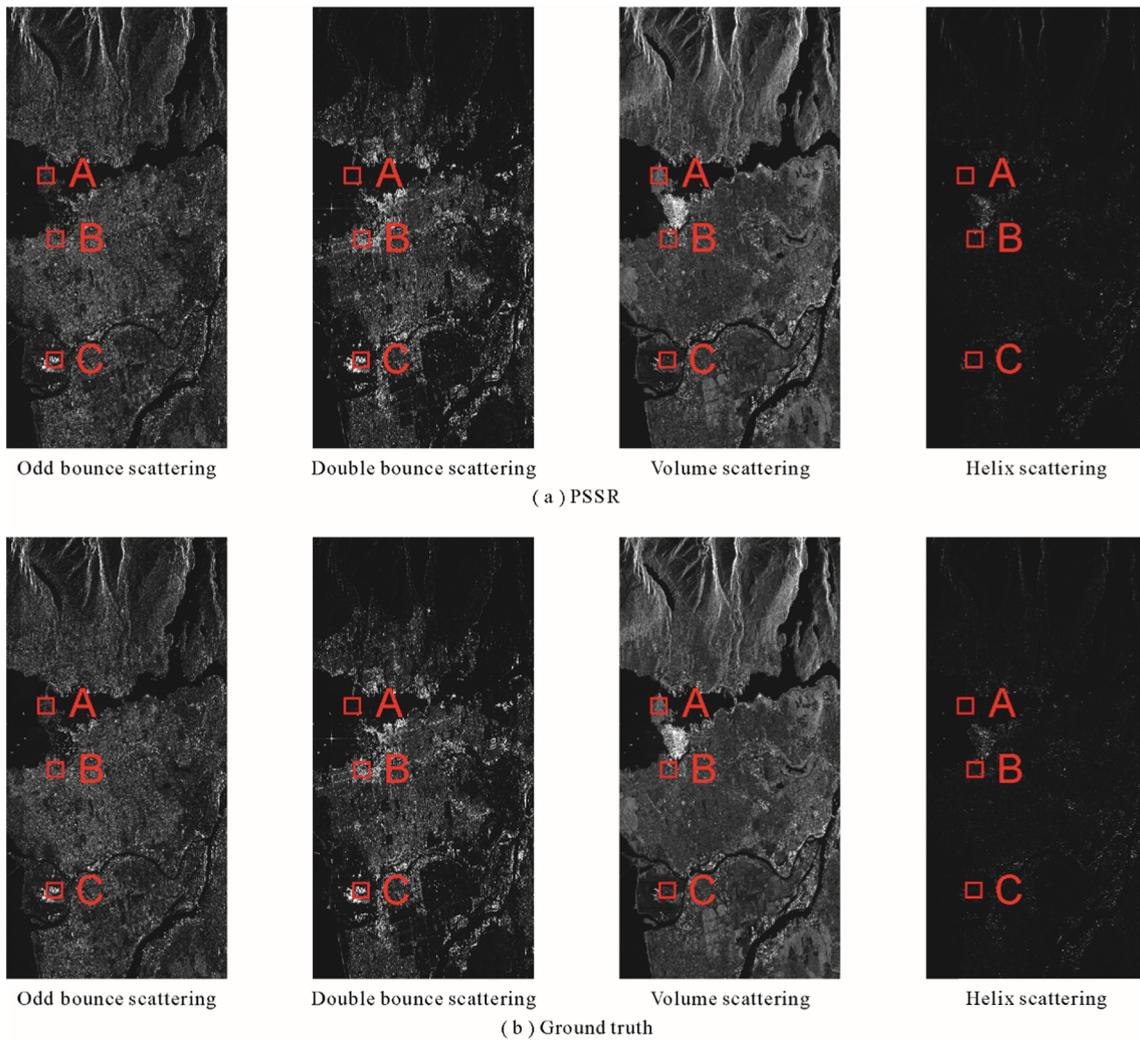


Fig. 26. PolSAR Image Polarimetric Decomposition Results.

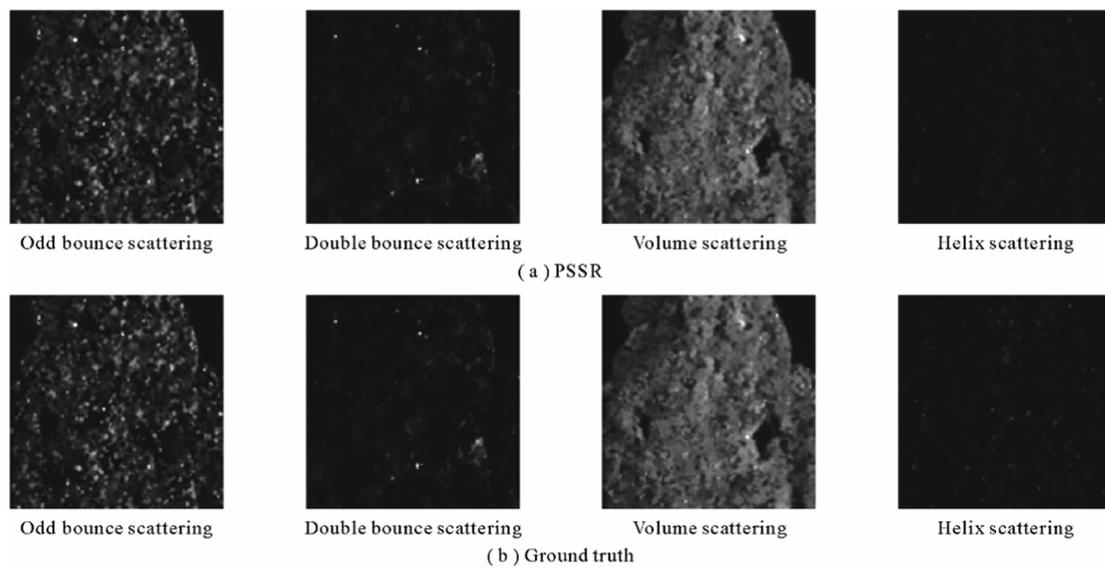


Fig. 27. PolSAR Image Polarimetric Decomposition Results in the Forest.

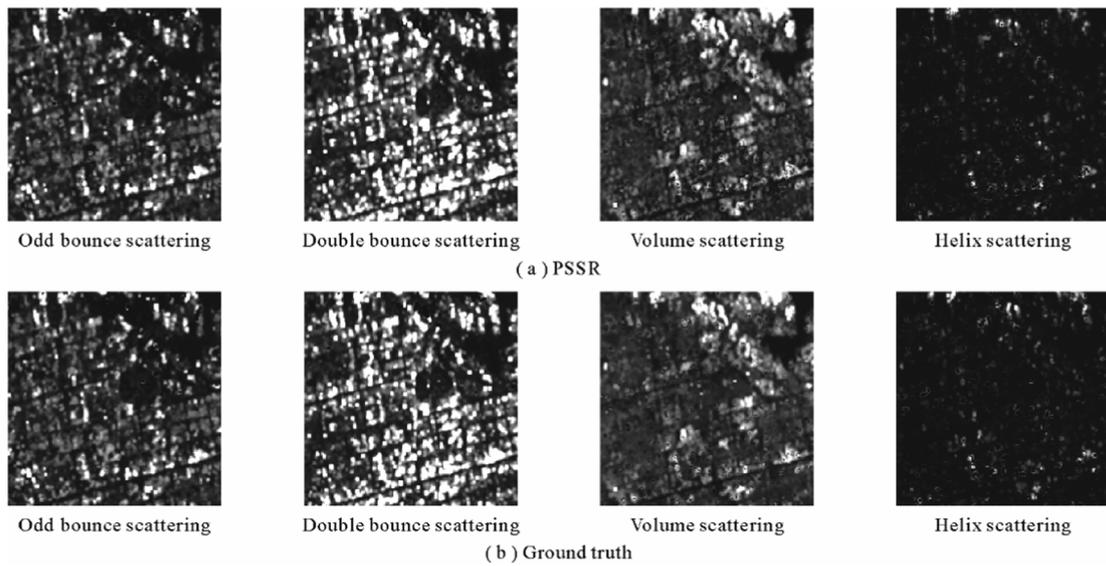


Fig. 28. PolSAR Image Polarimetric Decomposition Results in the Urban.

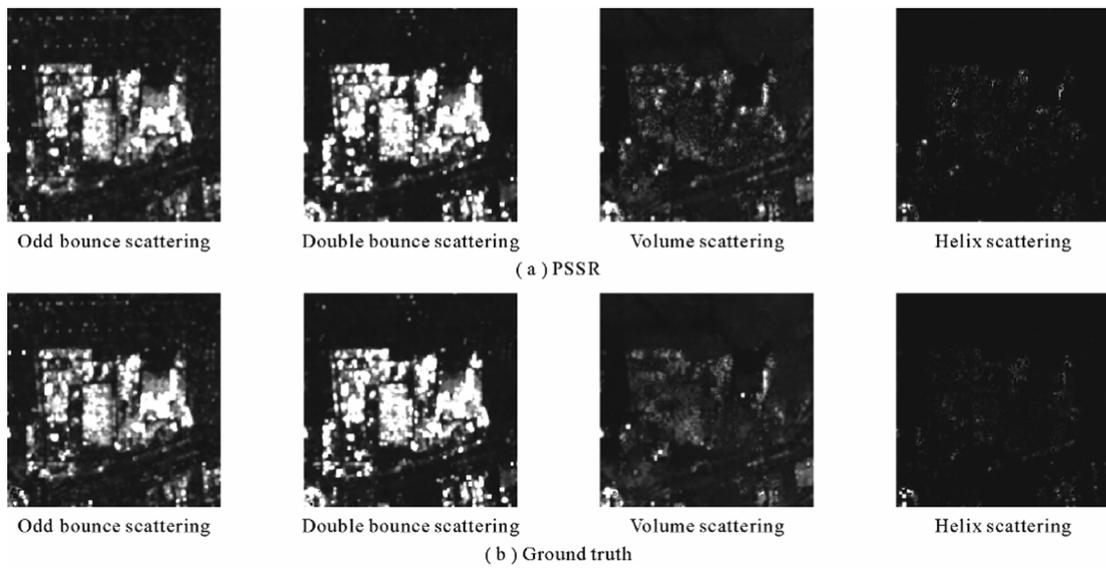


Fig. 29. PolSAR Image Polarimetric Decomposition Results in the Airport.

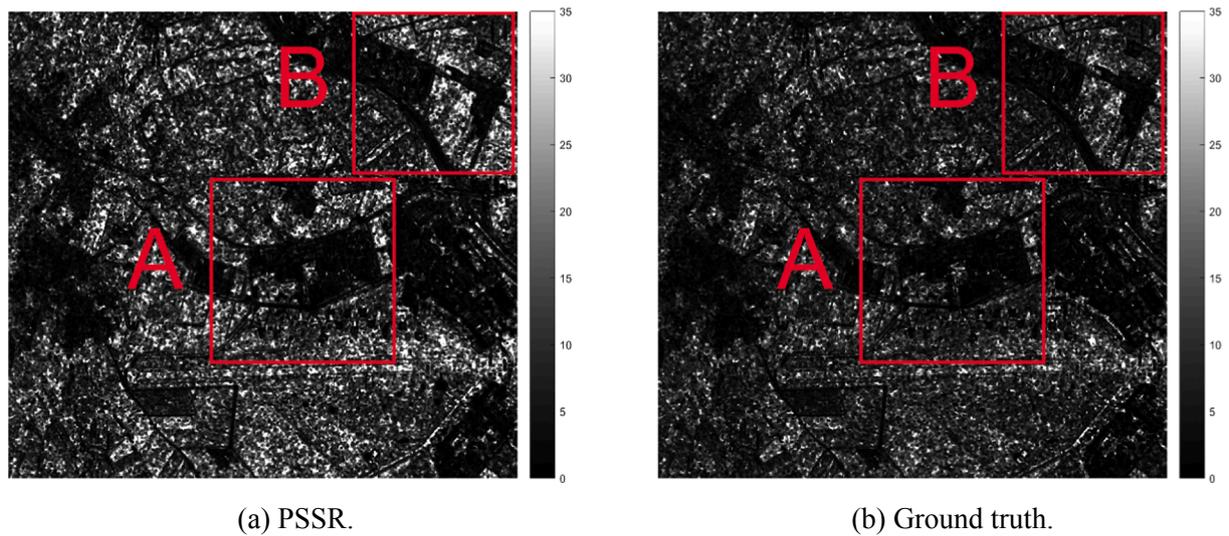


Fig. 30. Comparison of the ENL results.

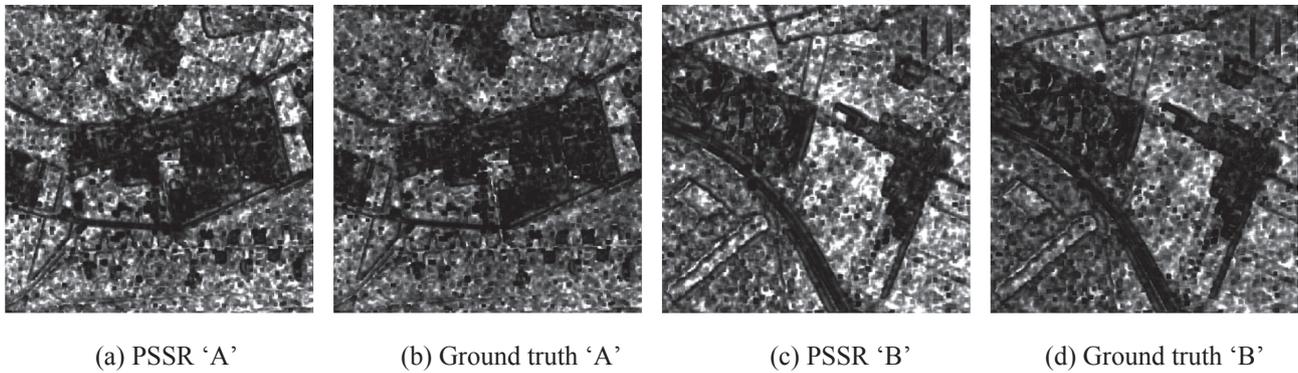


Fig. 31. Comparison of the ENL results in area ‘A’ and area ‘B’.

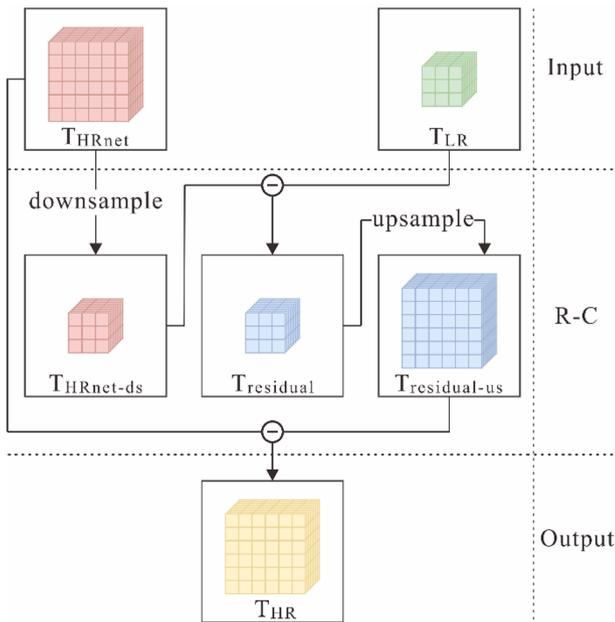


Fig. 32. The residual compensation strategy.

reconstruction results T_{HRnet} are deemed as an intermediate product, and prior information is used to perform RC on the network reconstruction results. The SRPSC method was used as the up-sampling function for the real and imaginary values, respectively, to maintain the complex structural characteristics of the data.

Three sets of data were used to verify the validity of the RC strategy. In Table 10, compared with the proposed method, after application of RC strategy, most of the quantitative indicators are improved. In Fig. 33, the result of the RC strategy, the phase change is smaller and is more in line with the phase distribution of the original high-resolution PolSAR image. The experimental results show that the RC strategy has a positive effect on correcting the network super-resolution results.

4.5. Experiments in PolSAR image classification

In order to further evaluate the effect of the proposed method, we used the support vector machine (SVM) method to classify the reconstruction results, to verify the effect of the super-resolution reconstruction results in practical applications. We classified the PolSAR images using SVM supervised classification in PolSARpro v5.0. The AIRSAR Flevoland data (Zhou et al., 2018) was used for the classification. The PolSAR image contains 15 types of features, as shown in Fig. 34.

We provide the overall accuracy and kappa coefficient value for the classification results, to verify the classification accuracy. It can be seen from Table 11 that the overall accuracy and kappa coefficient of the proposed method in the classification experiment are the highest, indicating that the classification results have high consistency. It is further shown that the proposed method can obtain good results in a practical application.

It can be seen from Table 11 that the correctness rate of the majority class is 90% or higher in the classification experiment. For the forest, beet, potatoes, rapeseed, and building classes, the classification accuracies of the proposed method are significantly better than the classification accuracies of the other comparison algorithms. In particular, for the building class, fewer training samples lead to poor classification results for the bicubic interpolation and traditional super-resolution

Table 10
Quantitative Evaluation Results.

Sensor	RS-2 (S)		RS-2 (V)		ESAR		PiSAR	
	PSSR	PSSRRC	PSSR	PSSRRC	PSSR	PSSRRC	PSSR	PSSRRC
MAE ($ P_1 ^2$)	0.258	0.256	0.215	0.213	1.198	1.242	0.197	0.194
MAE ($ P_2 ^2$)	0.373	0.380	0.164	0.159	0.900	0.929	0.153	0.150
MAE ($ P_3 ^2$)	0.335	0.327	0.249	0.245	1.813	1.568	1.428	1.314
MAE (mean)	0.322	0.321	0.209	0.205	1.304	1.246	0.593	0.552
PSNR ($ P_1 ^2$)	52.24	52.30	52.14	52.19	48.67	48.65	49.24	49.31
PSNR ($ P_2 ^2$)	52.03	52.09	52.30	52.39	49.69	49.68	49.94	50.08
PSNR ($ P_3 ^2$)	50.87	51.16	52.52	52.74	45.61	46.06	45.84	46.10
PSNR (mean)	51.71	51.85	52.32	52.44	47.99	48.13	48.34	48.50

Note: RS-2 (S) and RS-2 (V) mean the RADARSAT-2 data from the San Francisco and the Vancouver area respectively. PSSRRC means the proposed method with a residual compensation strategy.

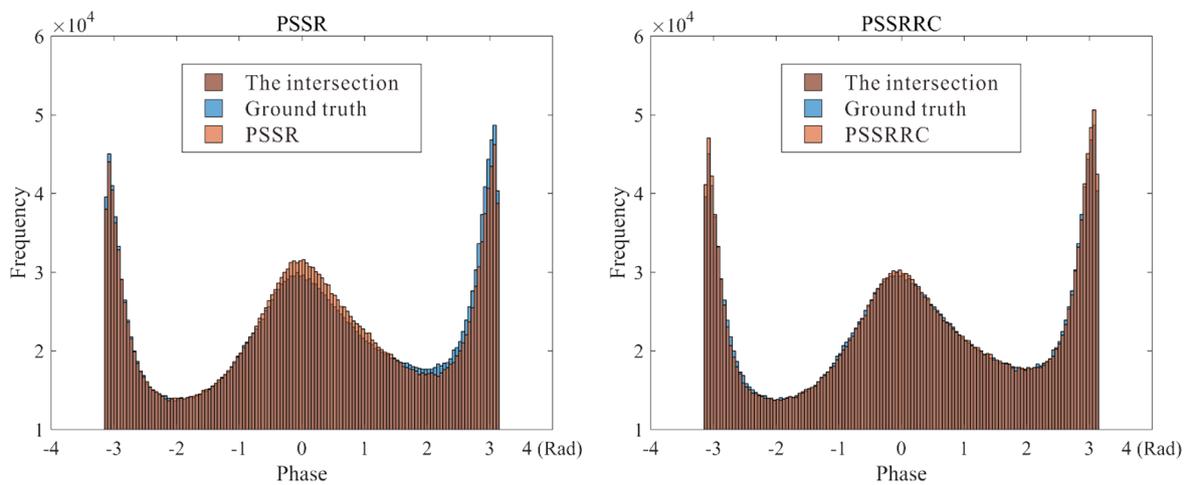


Fig. 33. Phase statistics histograms.

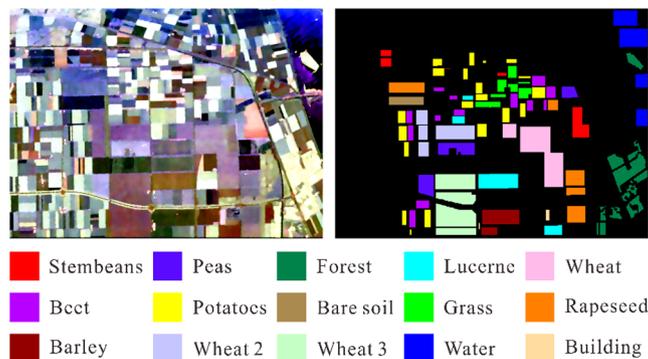


Fig. 34. Flevoland dataset and the labels of the dataset.

methods, and the proposed method obtains a superior classification accuracy. For the other categories, the proposed method is comparable to the results of the comparison algorithms. Compared with PSSR, the classification result of PSSR with RC strategy (PSSRRC) has been further improved as shown in Fig. 35, indicating that the RC strategy can promote the super-resolution reconstruction of PolSAR imagery.

Table 11
Classification accuracies of the comparison algorithms and the proposed method.

	Bicubic	SRPSC	MSSR	PSSR	PSSRRC
Stem beans	98.93	98.89	99.25	95.94	95.46
Peas	99.00	98.88	98.68	98.91	98.87
Forest	95.25	95.40	92.86	99.04	97.04
Lucerne	99.37	99.28	93.22	98.01	97.50
Wheat	95.18	95.24	93.99	94.01	96.33
Beet	87.23	88.10	82.03	95.07	95.26
Potatoes	75.67	76.18	81.17	89.16	85.50
Bare soil	99.38	99.35	99.84	97.60	99.51
Grass	91.34	90.94	88.93	85.91	89.87
Rapeseed	90.38	90.18	82.03	91.16	90.41
Barley	95.78	95.99	95.08	91.08	95.07
Wheat2	91.94	91.60	80.70	90.95	91.83
Wheat3	95.01	94.84	95.10	95.32	94.13
Water	90.47	89.92	87.72	79.41	87.80
Buildings	15.13	16.18	96.01	97.06	97.27
OA	92.08	92.08	90.02	<u>92.80</u>	93.40
Kappa	0.91	0.91	0.89	<u>0.92</u>	0.93

5. Conclusion

In this paper, we have proposed a PolSAR image super-resolution framework to improve the spatial resolution of low-resolution PolSAR images. Differing from the traditional methods, the proposed method is based on residual convolutional neural network, which used deconvolution to reduce the loss of precision caused by interpolation, added PReLU to preserve the negative information and improve the accuracy, and designed the complex block structure for extracting the mixed features of complex numbers. Taking into account the numerical characteristics and data structure of PolSAR imagery, the spatial resolution of the PolSAR imagery can be effectively improved by the use of the proposed method. Moreover, the use of the RC strategy reduced the artifacts that exist in the result of the super-resolution reconstruction. Compared with the traditional methods, the proposed method shows a state-of-the-art performance in both accuracy and visual results, especially in terms of detail information retention. In the phase statistics and polarization response experiments, the proposed method demonstrated good retention of polarimetric information. In addition, in practical application, the proposed method obtained a higher classification accuracy in the PolSAR image classification.

In the proposed method, amplitude information and phase information are not used to constrain the back-propagation. Therefore, considering amplitude and phase constraints in the network will be key points in our future work.

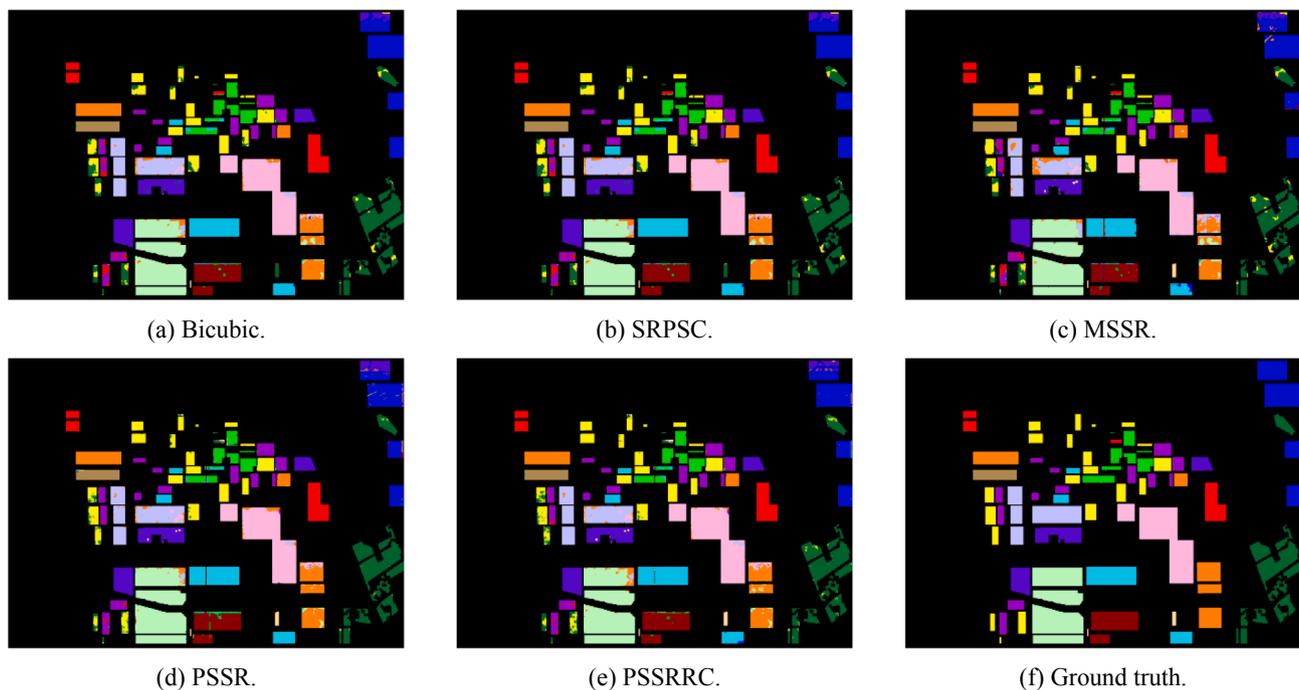


Fig. 35. Classification results of the comparison algorithms and the proposed method.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the National Natural Science Foundation of China under grant 61671334, 41701400, 61971319, and 41922008, and the Key Laboratory of Surveying and Mapping Science and Geospatial Information Technology of Ministry of Natural Resources under grant 201905. In addition, we would like to thanks to ESA and CSA for providing the PolSAR sample data.

References

- Qi, Z., Yeh, A.G.-O., Li, X., Zhang, X., 2015. A three-component method for timely detection of land cover changes using polarimetric SAR images. *ISPRS J. Photogramm. Remote Sens.* 107, 3–21.
- Kingma, D.P., Ba, L.J., 2015. Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations. (ICLR)*.
- Lê, T.T., Atto, A.M., Trouvé, E., Solikhin, A., Pinel, V., 2015. Change detection matrix for multitemporal filtering and change analysis of SAR and PolSAR image time series. *ISPRS J. Photogramm. Remote Sens.* 107, 64–76.
- Zhao, L., Yang, J., Li, P., Zhang, L., Shi, L., Lang, F., 2013. Damage assessment in urban areas using post-earthquake airborne PolSAR imagery. *Int. J. Remote Sens.* 34, 8952–8966.
- Shi, L., Sun, W., Yang, J., Li, P., Lu, L., 2015. Building Collapse Assessment by the Use of Postearthquake Chinese VHR Airborne SAR. *IEEE Geosci. Remote Sens. Lett.* 12, 2021–2025.
- Liu, C., Gierull, C.H., 2007. A New Application for PolSAR Imagery in the Field of Moving Target Indication Ship Detection. *IEEE Trans. Geosci. Remote Sens.* 45, 3426–3436.
- Liu, F., Jiao, L., Tang, X., Yang, S., Ma, W., Hou, B., 2019. Local Restricted Convolutional Neural Network for Change Detection in Polarimetric SAR Images. *IEEE Trans. Neural Networks Learn. Syst.* 30, 818–833.
- Zhao, L., Yang, J., Li, P., Shi, L., Zhang, L., 2017. Characterizing Lodging Damage in Wheat and Canola Using Radarsat-2 Polarimetric SAR Data. *Remote Sensing Letters* 8, 667–675.
- Shi, L., Zhang, L., Zhao, L., Yang, J., Li, P., Zhang, L., 2013. The potential of linear discriminative Laplacian eigenmaps dimensionality reduction in polarimetric SAR classification for agricultural areas. *ISPRS J. Photogramm. Remote Sens.* 86, 124–135.
- Liu, X., Jiao, L., Tang, X., Sun, Q., Zhang, D., 2019. Polarimetric Convolutional Network for PolSAR Image Classification. *IEEE Trans. Geosci. Remote Sens.* 57, 3040–3054.
- Hänsch, R., Hellwich, O., 2018. Skipping the real world: Classification of PolSAR images without explicit feature extraction. *ISPRS J. Photogramm. Remote Sens.* 140, 122–132.
- Dabboor, M., Shokr, M., 2013. A new Likelihood Ratio for supervised classification of fully polarimetric SAR data: An application for sea ice type mapping. *ISPRS J. Photogramm. Remote Sens.* 84, 1–11.
- White, L., Brisco, B., Dabboor, M., Schmitt, A., Pratt, A., 2015. A Collection of SAR Methodologies for Monitoring Wetlands. *Remote Sensing* 7, 7615–7645.
- Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Motagh, M., 2017. Random forest wetland classification using ALOS-2 L-band, RADARSAT-2 C-band, and TerraSAR-X imagery. *ISPRS J. Photogramm. Remote Sens.* 130, 13–31.
- Song, Q., Xu, F., Jin, Y.-Q., 2018. Radar Image Colorization: Converting Single-Polarization to Fully Polarimetric Using Deep Neural Networks. *IEEE Access* 6, 1647–1661.
- Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., Zhang, L., 2016. Image super-resolution: The techniques, applications, and future. *Signal Process.* 128, 389–408.
- Zhang, P., Zhang, S., 2013. Impact on polarimetric SAR calibration of SAR super-resolution imaging algorithm. 2013 IEEE International Geoscience and Remote Sensing Symposium.
- Suwa, K., Iwamoto, M., 2007. A Two-Dimensional Bandwidth Extrapolation Technique for Polarimetric Synthetic Aperture Radar Images. *IEEE Trans. Geosci. Remote Sens.* 45, 45–54.
- Zou, B., Hao, H., Guo, X., 2008. Super-Resolution of Polarimetric SAR Images Based on Target Decomposition and Polarimetric Spatial Correlation. In: *IEEE International Geoscience and Remote Sensing Symposium*, pp. 911–914.
- Zhang, L., Zou, B., Hao, H., Zhang, Y., 2011. A novel super-resolution method of PolSAR images based on target decomposition and polarimetric spatial correlation. *Int. J. Remote Sens.* 32, 4893–4913.
- Pastina, D., Lombardo, P., Farina, A., Daddi, P., 2003. Super-resolution of polarimetric SAR images of ship targets. *Signal Process.* 83, 1737–1748.
- Jiong, C., Jian, Y., 2007. Super-Resolution of Polarimetric SAR Images for Ship Detection. In: *IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, pp. 1499–1502.
- Pastina, D., Lombardo, P., Farina, A., Daddi, P., 2001. Super-resolution of polarimetric SAR images of a ship. In: *IEEE International Geoscience and Remote Sensing Symposium*, pp. 2343–2345.
- Dong, C., Loy, C.C., He, K., Tang, X., 2014. Learning a Deep Convolutional Network for Image Super-Resolution. In: *Computer Vision – ECCV 2014*. Springer International Publishing, pp. 184–199.
- Dong, C., Loy, C.C., He, K., Tang, X., 2016. Accelerating the Super-Resolution Convolutional Neural Network. In: *Computer Vision – ECCV 2016*. Springer International Publishing, pp. 391–407.
- Kim, J., Lee, J.K., Lee, K.M., 2016. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1646–1654.
- He, K., Sun, J., 2015. Convolutional neural networks at constrained time cost. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5353–5360.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image Recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778.
- Tai, Y., Yang, J., Liu, X., 2017. Image Super-Resolution via Deep Recursive Residual

- Network. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2790–2798.
- Kim, J., Lee, J.K., Lee, K.M., 2016. Deeply-Recursive Convolutional Network for Image Super-Resolution. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1637–1645.
- Lim, B., Son, S., Kim, H., Nah, S., Lee, K.M., 2017. Enhanced Deep Residual Networks for Single Image Super-Resolution. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1132–1140.
- Haris, M., Shakhnarovich, G., Ukita, N., 2018. Deep Back-Projection Networks for Super-Resolution. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1664–1673.
- Zhang, L., Zhang, L., Du, B., 2016. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geosci. Remote Sens. Mag.* 4, 22–40.
- Zhang, Q., Yuan, Q., Zeng, C., Li, X., Wei, Y., 2018. Missing Data Reconstruction in Remote Sensing Image With a Unified Spatial–Temporal–Spectral Deep Convolutional Neural Network. *IEEE Trans. Geosci. Remote Sens.* 56, 4274–4288.
- Yuan, Q., Zhang, Q., Li, J., Shen, H., Zhang, L., 2019. Hyperspectral Image Denoising Employing a Spatial-Spectral Deep Residual Convolutional Neural Network. *IEEE Trans. Geosci. Remote Sens.* 57, 1205–1218.
- Wang, P., Zhang, H., Patel, V.M., 2017. SAR Image Despeckling Using a Convolutional Neural Network. *IEEE Signal Process Lett.* 24, 1763–1767.
- Chierchia, G., Cozzolino, D., Poggi, G., Verdoliva, L., 2017. SAR, image despeckling through convolutional neural networks. In: IEEE International Geoscience and Remote Sensing Symposium, pp. 5438–5441.
- Zhang, Q., Yuan, Q., Li, J., Yang, Z., Ma, X., 2018. Learning a Dilated Residual Network for SAR Image Despeckling. *Remote Sensing*. 10, 196.
- Zhou, F., Fan, W., Sheng, Q., Tao, M., 2018. Ship Detection Based on Deep Convolutional Neural Networks for PolSAR Images. In: IEEE International Geoscience and Remote Sensing Symposium, pp. 681–684.
- Zhang, Z., Wang, H., Xu, F., Jin, Y.-Q., 2017. Complex-Valued Convolutional Neural Network and Its Application in Polarimetric SAR Image Classification. *IEEE Trans. Geosci. Remote Sens.* 55, 7177–7188.
- Sun, W., Wang, R., 2018. Fully Convolutional Networks for Semantic Segmentation of Very High Resolution Remotely Sensed Images Combined With DSM. *IEEE Geosci. Remote Sens. Lett.* 15, 474–478.
- Mahdianpari, M., Salehi, B., Rezaee, M., Mohammadimanesh, F., Zhang, Y., 2018. Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery. *Remote Sensing*. 10, 1119.
- Lee, J.S., Pottier, E., 2009. Polarimetric Radar Imaging From Basics to Applications. CRC Press, Boca Raton, FL.
- Chitroub, S., Houacine, A., Sansal, B., 2002. Statistical characterisation and modelling of SAR images. *Signal Process.* 82, 69–92.
- Zeiler, M.D., Krishnan, D., Taylor, G.W., Fergus, R., 2010. Deconvolutional networks. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2528–2535.
- He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In: IEEE International Conference on Computer Vision, pp. 1026–1034.
- Keys, R., 1981. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust. Speech Signal Process.* 29, 1153–1160.
- Anfinsen, S.N., Doulgeris, A.P., Eltoft, T., 2009. Estimation of the Equivalent Number of Looks in Polarimetric Synthetic Aperture Radar Imagery. *IEEE Trans. Geosci. Remote Sens.* 47, 3795–3809.
- Ren, Y., Yang, J., Zhao, L., Li, P., Shi, L., 2019. SIRV-Based High-Resolution PolSAR Image Speckle Suppression via Dual-Domain Filtering. *IEEE Trans. Geosci. Remote Sens.* 1–16.
- Yamaguchi, Y., Sato, A., Boerner, W.-M., Sato, R., Yamada, H., 2011. Four-Component Scattering Power Decomposition With Rotation of Coherency Matrix. *IEEE Trans. Geosci. Remote Sens.* 49, 2251–2258.
- Wang, Q., Shi, W., Atkinson, P.M., Zhao, Y., 2015. Downscaling MODIS images with area-to-point regression kriging. *Remote Sens. Environ.* 166, 191–204.
- Wang, Q., Atkinson, P.M., 2018. Spatio-temporal fusion for daily Sentinel-2 images. *Remote Sens. Environ.* 204, 31–42.