Article:

Li, S., Xu, H., Gao, S. et al. (2020) An interferometric phase noise reduction method based on modified denoising convolutional neural network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13. pp. 4947-4959. ISSN: 1939-1404

https://doi.org/10.1109/jstars.2020.3017808

An Interferometric Phase Noise Reduction Method Based on Modified Denoising Convolutional Neural Network

Shuo Li, Huaping Xu, Shuai Gao, Wei Liu, Chunsheng Li and Aifang Liu

Abstract—**Traditional interferometric synthetic aperture radar (InSAR) denoising methods normally try to estimate the phase fringes directly from the noisy interferogram. Since the statistics of phase noise are more stable than the phase corresponding to complex terrain, it could be easier to estimate the phase noise. In this paper, phase noises rather than phase fringes are estimated first, and then they are subtracted from the noisy interferometric phase for denoising. The denoising convolutional neural network (DnCNN) is introduced to estimate phase noise and then a modified network called IPDnCNN is constructed for the problem. Based on the IPDnCNN, a novel interferometric phase noise reduction algorithm is proposed, which can reduce phase noise while protecting fringe edges and avoid the use of filter windows. Experimental results using simulated and real data are provided to demonstrate the effectiveness of the proposed method.**

Index Terms—**Interferometric synthetic aperture radar, Phase noise reduction; Denoising convolutional neural network.**

I. INTRODUCTION

SYNTHETIC Aperture Radar Interferometry (InSAR) is an all-time and all-weather remote-sensing technique and can be used for generating digital elevation models (DEMs) or detecting surface deformation [1], [2]. However, phase noise cannot be avoided due to the existence of thermal noise, temporal decorrelation, spatial decorrelation, and mis-coregistration, etc., which increases the difficulty of phase unwrapping and reduces the accuracy of DEM and deformation reconstruction [3]. Consequently, noise reduction is crucial for improving the quality of SAR interferograms before phase unwrapping [4].

Traditional phase noise reduction approaches are usually divided into two categories: spatial-domain filtering and transform-domain filtering. In spatial-domain filtering, local phase estimation methods are widely used. Classic boxcar filters estimate the parameters over a rectangular sliding window and require the samples to be homogeneous [5]. The algorithms proposed in [6], [7] are based on the noise subspace and the projection of the signal subspace. The subspace of noise is obtained from a local window after coarse co-registration, and the window size may influence its performance. Meanwhile, it is difficult to estimate the signal subspace dimension in regions with low coherence. The complex-valued Markov random field filter (MRF) is employed in [8],[9] to estimate the noise-free phase term by minimizing the energy function in a local window. The energy function is further developed in [10] based on a joint probability and the phase value is computed with a genetic algorithm. Nevertheless, some complicated areas could be over-smoothed due to the fixed local window for MRF methods. A common issue with these local phase estimation methods is that they have difficulty in adapting to different features with the fixed window size.

The Lee filter is designed to achieve a balance between residual noise and detail information loss [5], where a window with the adjustable size and direction is employed according to the local gradient of the interferogram. However, this method only calculates 16 discrete orientations, which brings distortion to curved fringes. Following the Lee filter, the Intensity-Driven Adaptive-Neighborhood (IDAN) method carries out a complex multi-look operation on an adaptive neighborhood [11], where the adaptive-window filters can achieve a tradeoff between noise reduction and detail preservation. However, noise reduction is not effective because the adjacent pixels are limited within the local window.

To overcome the limitation of estimating the phase in a local window, non-local phase estimation is proposed [12], [13], which suppresses noise while preserving textures utilizing weighted averaging of similar pixels, with phase similarity calculated by a matching window. In [14], a refined non-local filter is proposed, which measures the similarity between the central pixel and the remaining pixels in the matching window by a normalized probability density function. However, non-local methods cannot provide accurate similarity estimation in highly sloped terrains because a fixed-size matching window is used to capture the varied fringe curvature [15].

On the other hand, the transform domain filtering approach mainly includes the wavelet transform and the frequency transform. In [16], a complex wavelet interferometric phase filter (WInPF) is implemented utilizing the discrete wavelet packet transform decomposition to extract and amplify the useful signal in the interferogram. There are several adapted versions of the WInPF, such as those studied in [17], [18],

The authors, Shuo Li, Huaping, Xu, Shuai Gao and Chunsheng Li are with the School of Electronic and Information Engineering, Beihang University, Beijing 100083, China (e-mail: shuo201@buaa.edu.cn, xuhuaping@buaa.edu.cn; gaoshuai@buaa.edu.cn, lics@buaa.edu.cn).

The author, Wei Liu, is with the Electronic and Electrical Engineering Department, University of Sheffield, S1 3JD, U.K (e-mail: w.liu@sheffield.ac.uk).

The author, Aifang Liu is with Nanjing Research Institute of Electronics Technology, Nanjing 211113, China (e-mail: lafsx1997@163.com).

This work was supported in part by the National Natural Science Foundation of China under Grant 61471020 and Shanghai Aerospace Science and Technology Innovation Fund under Grant SAST2019-026.

where by employing the Wiener filter or simultaneous detection and estimation techniques, better performance is achieved in filtering complicated areas. The phase information and noise can be more easily separated in the wavelet domain, but the wavelet domain filters greatly depend on the scales of wavelet decomposition and the threshold of wavelet coefficients.

For the frequency domain methods, the Goldstein filter suppresses phase noise by enhancing the main frequency components, but its performance is affected by the window size and filter parameter [19]. A modification is proposed in [20] to construct a filtering parameter dependent on the coherence value to keep more texture details in the interferogram. However, a biased coherence estimation result usually leads to an inaccurate estimation of the filtering parameter. To solve this problem, the filtering parameter is modified using an optimal nonlinear model with homogenous regions and a bootstrapping technique [21], or using a combination of correlation and multi-look factors [22], [23]. These extended Goldstein filters preserve phase fringes well, but frequency domain filtering still suppresses high-frequency components of fringes, resulting in the loss of fringe details.

In order to further enhance the fringe edge-preserving ability, E. Trouve proposed a local frequency compensation filtering algorithm [24]. The local fringe frequency (LFF) is removed in each local window, and then the residual phase is smoothed. Finally, the removed fringe frequency is added to the filtered residual phase to generate the filtered interferogram [25]. In [26], an adaptive multiresolution technique was proposed to modify the LFF estimation by setting a threshold to eliminate the “bad LFF values” which have a large difference compared to its neighboring pixels. It provides better protection for phase fringes, but it is still hard to estimate the fringe frequency for highly sloped terrain. In [27], multi-frequency data is used to achieve an accurate LFF in abruptly changing terrain and the Goldstein filter is applied to the residual phase. In [28], the local fringe frequency is removed before Goldstein filtering and the filter parameters are then optimized, which improves edge preservation. Nevertheless, the performance of local frequency compensation filters relies on frequency estimation accuracy, which is heavily influenced by phase noise and window size.

Window size selection is an important issue for traditional noise reduction methods. A large window denoises better at the cost of losing details such as edges, vice versa [29]. However, it is difficult to select a suitable window for all pixels of the interferogram due to the diversity of terrain. Although the adaptive-window filter can be used according to the coherence or other criterions, it only reaches a tradeoff between noise reduction and edge preservation. So the performance improvement from adaptive window is limited for complicated terrains. Fortunately, the phase noise statistics are more stable than the phase fringes since noise is almost from the same types of error sources, and thus estimation of phase noise could be easier than estimating phase fringes in areas with complicated terrains [30]. Therefore, in this work, we intend to estimate the phase noise first and then subtract it from the noisy phase to obtain the denoised one.

In recent years, convolutional neural networks (CNNs) has been developing rapidly and widely applied to image noise reduction [31]. CNNs have a powerful mapping approximation capability and can extract the noise characteristics from massive training data [32], [33]. For noise reduction in optical images, a large-scale multi-layer perceptron model is adopted in [34] with superior performance to traditional methods such as block-matching and 3-D filtering (BM3D) in image detail retention [35]. The denoising convolutional neural network (DnCNN) proposed in [36] can quickly and steadily remove optical image noises. In addition, sparse encoding [37], Trainable Nonlinear Reaction Diffusion (TNRD) [38], and self-coder [39] have achieved good results in optical image denoising through phase training. Among these methods, DnCNN is more effective in removing Gaussian noise from optical images.

In this paper, a new approach to remove the interferometric phase noise via a modified DnCNN is presented. The original DnCNN is modified to adapt to interferometric phase noise estimation and the denoised phase is obtained by removing the estimated noise from the original noisy interferogram. In the proposed method, the number of samples used for noise training, 300000 here, is huge and all pixels of noisy interferogram are exploited in phase noise estimation with the well-trained network. Therefore, it can effectively suppress noise while preserving phase fringe edges.

The remainder of this paper is organized as follows. The interferometric phase denoising method based on the modified DnCNN is proposed in Section II. Experimental results based on both simulated and real SAR data are presented in Section III, where the results are compared with those of slope adaptive filtering and improved Goldstein filtering algorithms. Conclusions are drawn in Section IV.

II. PRINCIPLE OF MODIFIED INTERFEROMETRIC PHASE NOISE REDUCTION METHOD

Traditional denoising methods normally estimate the interferometric fringes directly from the noisy interferogram with the pixels in a window. However, it is difficult to extract all of the fringes accurately especially for a complicated terrain interferogram with low coherence or low signal-to-noise ratio (SNR). As mentioned earlier, a new strategy is adopted in this work, where noise is estimated first and then removed from the image. Given the strong mapping approximation ability of CNN, it is suitable for processing low-SNR interferograms with heavy phase noise. Therefore, estimation of noise is achieved by modifying a DnCNN in this paper.

A. DnCNN Denoising Network [36]

DnCNN is modified from the VGG network [40] for image denoising. VGG is a typical CNN architecture proposed by the Visual Geometry Group of Oxford at ILSVRC 2014 based on the Alexnet network. Compared with the Alexnet network, VGG uses several groups of small convolution filters with a size of 3×3 instead of larger convolution filters. Under the condition of the same receptive field, the network expression capability is improved by increasing the network depth.

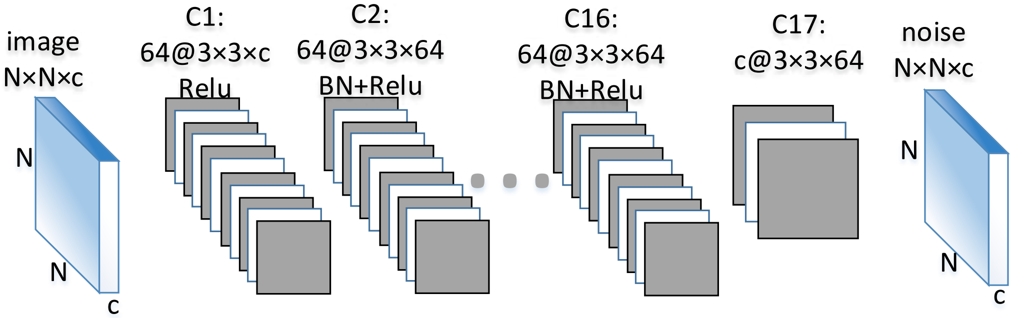


Fig. 1. The structure of DnCNN.

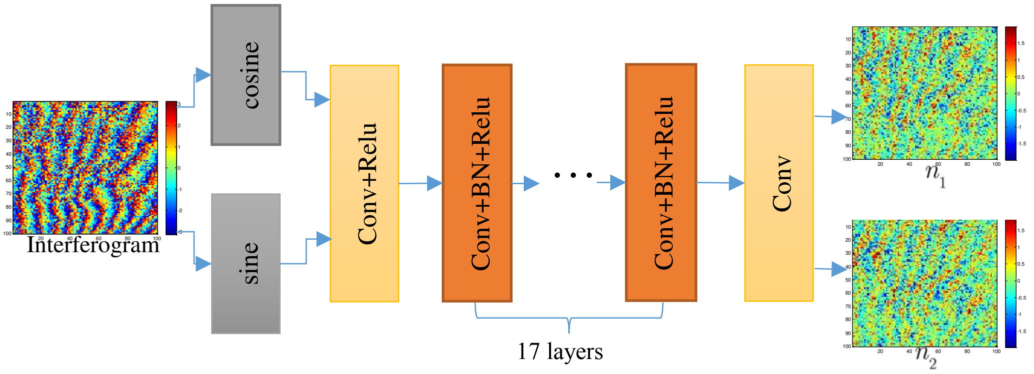


Fig. 2. The structure of IPDnCNN.

DnCNN removes all of the pooling layers in a VGG network, learns the noise distribution, and combines batch normalization (BN) for fast training and better denoising. It sets the depth of the network according to the patch size used in the most advanced denoising algorithms [36]. The network structure of DnCNN is shown in Fig. 1. Assuming that the original image size is $N \times N \times c$, the corresponding output is noise with the same size. c is the number of channels, i.e., $c=1$ in grey case and $c=3$ in color case. The size of the convolution filter is 3×3, the number of feature maps is 64, and the size of the receptive field is $(2d+1)\times(2d+1)$ for a depth of d.

Denote the noisy image by y, the clean image by x, and the noise by n. Then, the input of the DnCNN model is $y = x + n$. Unlike most denoising networks, such as multi-layer perceptron (MLP) [34], which trains the mapping function $F(y) = x$ to estimate the clean image directly, the DnCNN uses the residual learning method to estimate noise by training the mapping function $G(y) = n$, and then obtain a clean image by applying $x = y - G(y)$. The loss function in the DnCNN is the mean squared error of the noise

$$J(\theta) = \frac{1}{2N}\sum_{i=1}^{N}(G_\theta(y^{(i)}) - (y^{(i)} - x^{(i)}))^2 \tag{1}$$

where θ is the trainable parameters to be learned, $x^{(i)}$ and $y^{(i)}$ are the *i-th* clean image and noisy image, respectively, and $G_\theta(y^{(i)})$ is the noise trained by the *i-th* noisy image.

B. Interferometric Phase Denoising Network Based on DnCNN

Interferometric noise is considered as additive complex Gaussian in complex interferogram [5], which makes it suitable for denoising with the residual learning strategy [36]. However, significant errors occur when DnCNN, the classic optical image denoising network, is directly used to process the interferometric phase. By modifying DnCNN, an interferometric phase denoising network (IPDnCNN) is designed to suppress phase noise in a more robust way. The network structure of IPDnCNN is shown in Fig. 2, where sine and cosine values of phase are used as the input of the network and two more layers are added. Due to the wrapping characteristic of the interferometric phase, if we use the interferometric phase value as the input of the network directly, the fringe edge tends to be judged as noise, which leads to unstable network training and poor denoising result. To avoid the instability of the fringe edge on network training, IPDnCNN uses sine and cosine values of the interferometric phase as input to the network. Fig. 3 displays a cross-section of several fringes together with its sine and cosine values. As can be seen, phase jumps appear in the fringe edges. These jumps are similar to the characteristics of phase noise. Meanwhile, sine and cosine values are continuous even at fringe edges; thus, they won't be confused with the noise in network training.

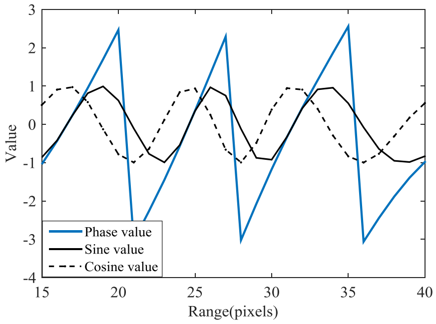


Fig. 3. Inputs of DnCNN and IPDnCNN.

Therefore, with the introduction of sine and cosine values of the interferometric phase, the number of channels becomes two in the proposed IPDnCNN, i.e., $c = 2$. Assuming noise n_1 and n_2 are outputs of the IPDnCNN, the interferometric phase φ is calculated by

$$\varphi = angle(\cos(x') + j(\sin(x') - n_1)) \tag{2}$$

where $j = \sqrt{-1}$ and

$$\begin{cases} \sin(x') = \sin(y) - n_1 \\ \cos(x') = \cos(y) - n_2 \end{cases} \tag{3}$$

The loss function of IPDnCNN is adopted to learn the residual mapping for prediction, which is changed as

$$\begin{aligned} J'(\theta) = & \frac{1}{2N}\sum_{i=1}^{N}(n_1^{(i)} - (\sin(y^{(i)}) - \sin(x^{(i)})))^2 \\ & + \frac{1}{2N}\sum_{i=1}^{N}(n_2^{(i)} - (\cos(y^{(i)}) - \cos(x^{(i)})))^2 \end{aligned} \tag{4}$$

where $n_1^{(i)}$ and $n_2^{(i)}$ are obtained through the mapping $G'_\theta(\sin(y^{(i)}), \cos(y^{(i)}))$.

For general optical image denoising tasks, DnCNN typically sets depth as 17 with a reception field of 35×35. Since the interferometric phase usually has a low SNR, a larger receptive field is needed to capture enough spatial information for denoising. In order to balance efficiency and performance, the depth of IPDnCNN in this work is increased to 19 with a receptive field size of 39×39. Simulation experiments show further increase in network depth will increase computational cost, but without clear improvement in denoising performance [36]. The first layer is a convolution layer with 64 filters of size $3 \times 3 \times 2$. 64 feature maps are obtained and the rectified linear units (Relu) activation is applied for nonlinearity. 64 filters of size $3 \times 3 \times 64$ are utilized for convolution from the second layer to the 18th layer, and the batch normalization technique (BN) is used for 64 feature maps to accelerate convergence at these 17 layers where Relu activation works. The last layer uses two filters of size $3 \times 3 \times 64$ to reconstruct two noisy images. Then, denoised sine and cosine images are obtained by removing the estimated noise from the noisy images, and finally, the denoised interferometric phase φ is calculated according to (2).

C. *Interferometric Phase Denoising Based on IPDnCNN*

Based on IPDnCNN, an interferometric phase denoising method is proposed. The phase noise is predicted by the network and then removed from the noisy phase to obtain the latent clean phase. As shown in Fig. 4, IPDnCNN is mainly composed of three steps: firstly, a large amount of training data with different noise intensity is prepared; then, the network is trained many epochs, including adjustment of parameters and other experiments; finally, the denoising network is tested with both simulated and real data.

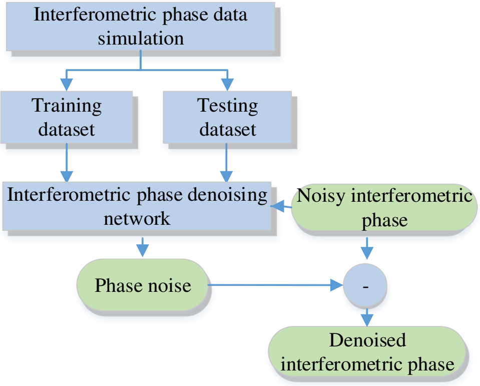


Fig.4. The flowchart of IPDnCNN.

1) Data preparation

Datasets are particularly critical for deep learning. Reasonable training data of IPDnCNN is produced through simulation. The training datasets are generated according to the observation geometry of InSAR using the real DEM data in Lanzhou, China. The simulation parameters are listed in Table 1.

TABLE I
Simulation Parameters

Parameters	Value
Baseline	600m
Wavelength	0.05666m
Near range	630km
Baseline obliquity angle	10°
Slant range resolution	1m
view angle	30°

The process of data preparation is shown in Fig. 5. Firstly, the slant distance Δr is calculated with DEMs and the satellite position. The clean wrapped phase can be expressed as:

$$\varphi = \mathrm{mod}\left(\frac{2\pi \cdot \Delta r}{\lambda},\ 2\pi\right) - \pi \tag{5}$$

where λ is wavelength, $\mathrm{mod}(\cdot)$ operator retains the principal value, and the actual phase is wrapped within the period $(-\pi, \pi]$.

To train the network for denoising with different noise levels, random complex Gaussian noise is added during SAR image simulation [42]. Then, the noisy interferometric phase is obtained through a complex conjugate cross-product of two SAR images.

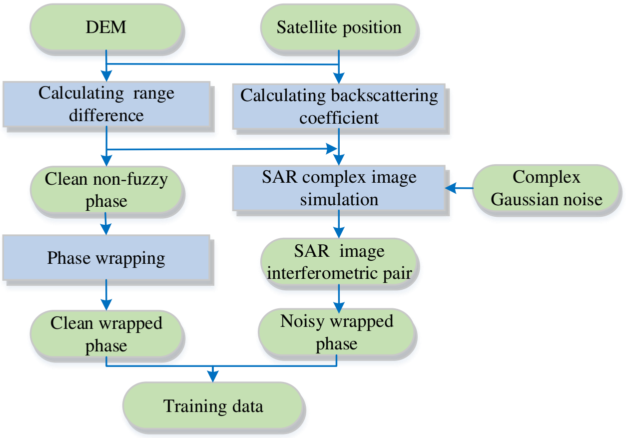


Fig. 5. Process of training data generation.

Following the steps above, 6000 groups of clean and noisy wrapped phases of size 591×591 are generated. To reduce overfitting issues during the training process, sufficient training data is needed. Augmentation techniques [43] including horizontal flip, vertical flip, rotation, and so on, are used to expand the training set. The patch size of the DnCNN is 40×40. It is increased to 80×80 in our method to capture more context information since the interferometric phase usually has a low SNR. After these steps, more image patches are produced. In this work, 300000 groups of phase patches are used as training data and 30000 groups are used as testing data. Four typical training samples are shown in Fig. 6.

2) *Network training*

The sine and cosine values of the noisy interferometric phase are input to the network. The output is obtained by subtracting sine and cosine values of the clean interferometric phase from those of the noisy interferometric phase.

Some network parameters are set according to the DnCNN network to learn the residual map for predicting phase noise. 40 epochs are trained using the stochastic gradient descent (SGD) method. The learning rate is manually adjusted based on empirical value according to the DnCNN network. The learning rate of the first 30 epochs is set as 0.001 to speed up the convergence. The learning rate in the last 10 epochs is 0.0001 to reduce the final error. Instead of setting a dropout rate to prevent overfitting, the BN and residual learning strategy are employed to stabilize and enhance the training performance [36]. The initial value of the network weight matrix in SGD also has a significant impact on the training process. For multi-layer networks, the initial values should be random while ensuring that the input and output of each hidden layer have the same statistical characteristics [41]. In order to speed up the convergence, the mini-batch size is set as 32, which means that 32 interferometric phases are randomly fed into the network each time.

During the training process, the value of the loss function given in (4) is observed. The network is said to have converged if the value of loss function gradually becomes smaller and finally stabilizes. The DnCNN method requires a GPU that is able to accommodate the computational load. Based on the Hardware Environment of DnCNN, all the experiments are implemented on a PC with Intel(R) Core(TM) i5-5200U@2.2GHz CPU and a Quadro P4000 GPU. The training in this experiment took about 3 days.

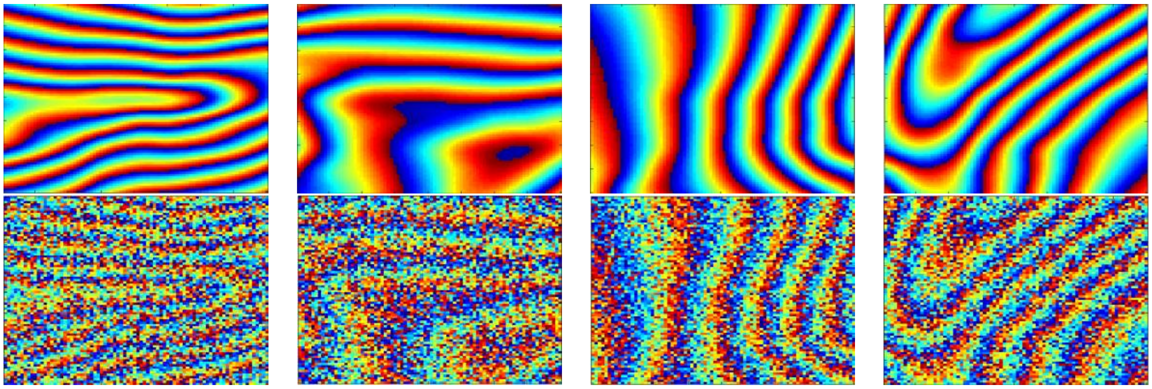
3) *Network testing*

Using the trained network, the simulated phase data, not involved in the training, are used to test the generality of IPDnCNN. The sine and cosine values of the noisy phase are fed into the trained IPDnCNN through two channels, and then the noise is obtained at the output. The interferometric phase after denoising is reconstructed according to (2) and (3). The performance is evaluated by phase MSE and residual points.

To test the generalization ability of the trained network, extra 100 groups of data simulated for different occasions are used to evaluate its ability to handle unknown phase noise. The coherence value of simulated data is randomly set from 0.03 to 0.97. Four groups of noisy phases and denoised results are provided in Fig. 7. The evaluation results are shown in TABLE II. After phase denoising using IPDnCNN, the average number of phase residues of 100 denoised images is improved from 1528 to 13, EPI from 4.119 to 0.957 and MSE from 1.338 to 0.134. It shows that the trained network performs well in handling more general cases.

TABLE II
EVALUATION RESULT

Iterferogram	Residues	EPI	MSE
Noisy phase	1528	4.119	1.338
Denoised phase	13	0.957	0.134



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

Fig. 6. Different training samples of clean/noisy interferometric fringe.

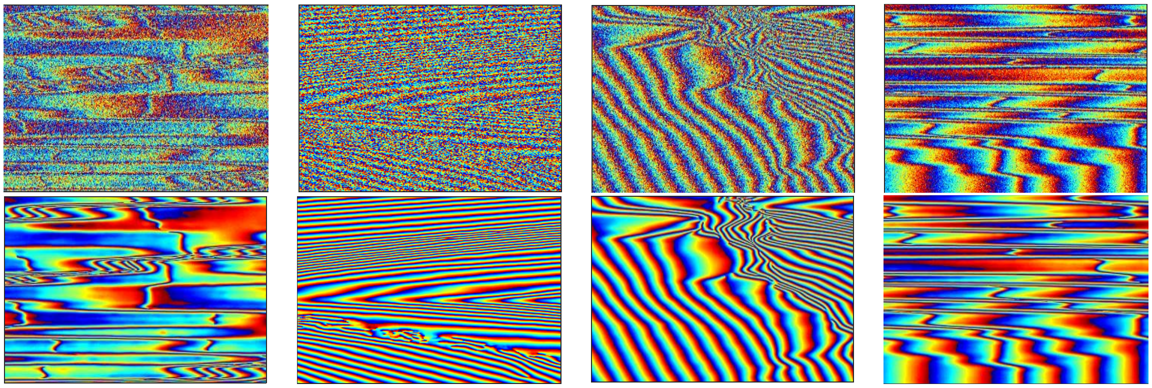

(a) (b) (c) (d)

Fig. 7. Four groups of noisy interferometric fringe and denoised result.

III. RESULTS AND ANALYSIS

In this section, both simulated and real interferograms are used to demonstrate the performance of the proposed phase denoising method. The training data sets are the same as those in Section II, which is simulated according to the observation geometry of InSAR and the real DEM data. The simulation parameters are shown in Table 1. The slope adaptive filter [27] and improved Goldstein filter [28] are used for comparison.

A. Basic experiments

SAR complex images of mountains, in Lanzhou, China, are simulated with the method in [42] and noisy interferometric phase with 591 × 591 pixels is produced. The noise-free phase is created with the same steps as in Section II. Clean and noisy phases are respectively shown in Fig. 8(a) and (b). As can be seen in Fig. 8(b), the phase fringes are submerged by noise because of low coherence.

Slope adaptive filter, improved Goldstein filter, DnCNN and the proposed IPDnCNN are applied to this simulated dataset. The window size for the former two filters is set as 11 × 11 and 32 × 32, whereas the DnCNN and IPDnCNN methods do not need a filter window. The results are shown in Fig. 9. In each group, the left image is the denoised phase, and the right one is the phase difference between denoised and clean phases, as well as the distribution of residual points. Among them, the purple dots represent positive residual points, while the blue points represent negative ones.

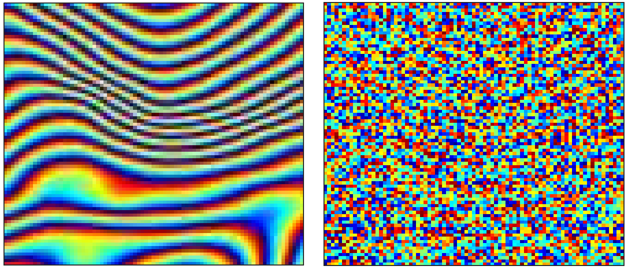
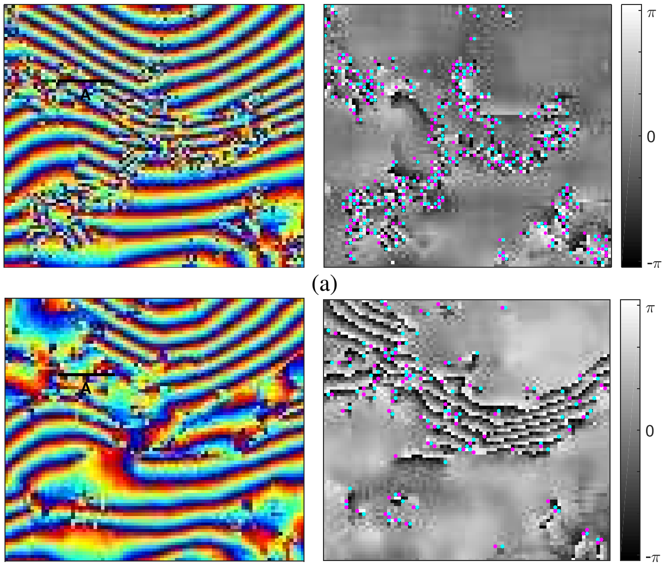

(a) (b)

Fig. 8. Simulated data: (a) clean phase, (b) noisy phase.

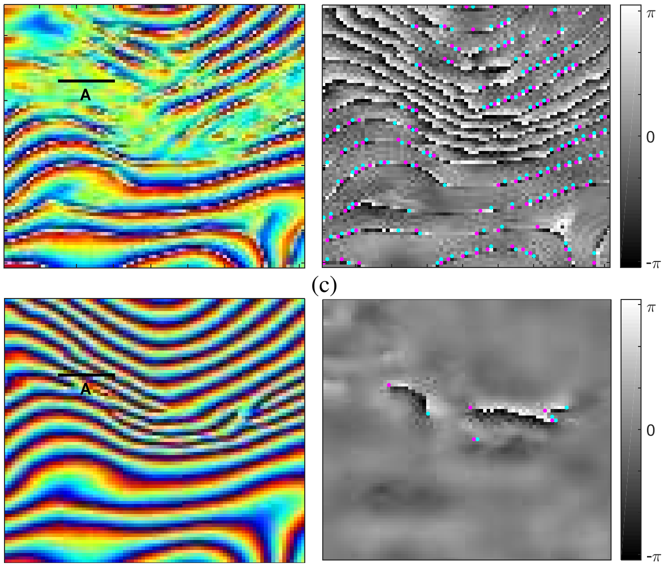
Clearly, the fringes in Fig. 9 (d) contain less noise than those in Fig. 9 (a)–(c), especially in the region with dense interferometric fringes. In Fig. 9(a), the slope adaptive filter can protect fringe edges better, but leaving more phase residues. In Fig. 9(b), the improved Goldstein filter has excessive filtering strength in dense fringe areas, leading to broken fringes. In Fig. 9(c), for the phase-as-input DnCNN method, significant errors have resulted in fringe edges because of misjudging the phase jumps as noise. It is obvious that there are too many errors when DnCNN is applied directly to the interferometric phase. So the modifications in IPDnCNN are necessary.

Comparing the result of the IPDnCNN method in Fig. 9(d) with the existing filters, the noise reduction effect is significant, and the fringes are much better preserved. From the phase error diagrams, it is clear that the IPDnCNN method performs better than the other filters.



(a)

(b)

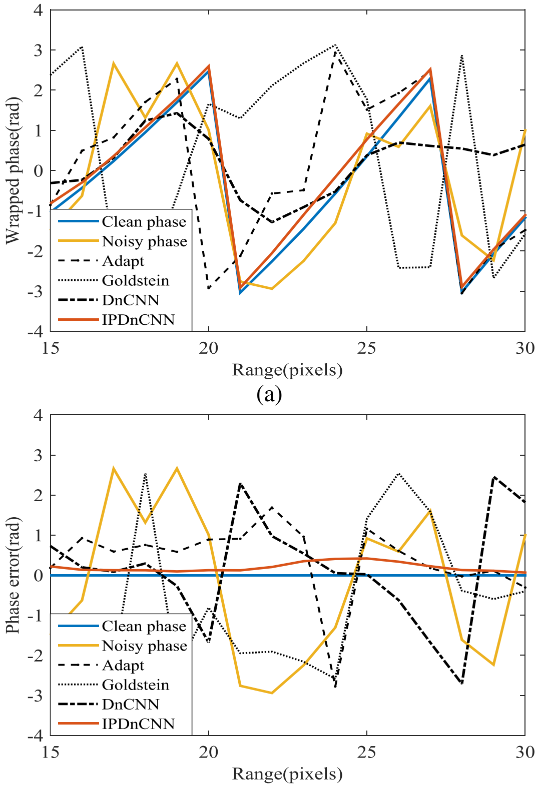


(c)

(d)

Fig. 9. Denoised phase and phase error: (a) slope adaptive filter with window size 11 × 11, (b) improved Goldstein filter with window size 11 × 11, (c) the DnCNN method, (d) the IPDnCNN method.

Fig. 10 displays a cross-section through the denoised phase and phase error map in region A. As clearly shown, in this low-coherence area, the result of the IPDnCNN method is most consistent with the original clean phase data.



(a)

(b)

Fig. 10. Cross-sections: (a) denoised phase, (b) phase error.

In order to quantitatively evaluate the results, mean square error (MSE), edge preservation index (EPI) and residues are used as the criteria [28]. MSE is to measure the deviation of the denoised phase from the clean one, given by

$$\mathrm{MSE} = \frac{\sum \left| \arg\left(\exp\left(j\varphi(i,j) - j\varphi_{clean}(i,j) \right) \right) \right|^2}{M} \tag{6}$$

where $\varphi(i,j)$ represents the denoised phase, $\varphi_{clean}(i,j)$ is the clean phase, and M is the number of pixels.

EPI is calculated by

$$\mathrm{EPI} = \frac{\sum \left(\left| \varphi(i,j) - \varphi(i+1,j) \right| + \left| \varphi(i,j) - \varphi(i,j+1) \right| \right)}{\sum \left(\left| \varphi_{clean}(i,j) - \varphi_{clean}(i+1,j) \right| + \left| \varphi_{clean}(i,j) - \varphi_{clean}(i,j+1) \right| \right)} \tag{7}$$

which is an indicator for performance in fringe and edge preservation and a value closer to 1 means a better edge preservation result.

Residues are the pixels where the gradient integral of adjacent pixels in a certain direction is not zero. More residues bring more difficulties in phase unwrapping, and thus reducing residues is one of the main purposes for phase denoising.

The evaluation results are presented in TABLE III. In this dense area, traditional filters using a small filtering window (11 ×11) perform better than using a large window (32×32). In terms of residues in the interferogram, the slope adaptive filter (11×11), improved Goldstein filter (11×11) and the DnCNN method have produced reductions of 90.12%, 97.15% and 84.61%, respectively, while by the IPDnCNN method it is 99.95%. The phase EPI for the IPDnCNN method is closer to 1 compared with the other methods, which means that it has a better performance in fringe preservation. Moreover, the MSE of the IPDnCNN method is the smallest due to an excellent phase smoothing performance.

TABLE III
EVALUATION RESULTS OF SIMULATED DATA

Iterferogram	Residues	EPI	MSE
Clean phase	0	1	0
Noisy phase	103767	2.3598	2.3416
Slope adaptive filter (11×11)	10253	1.1548	0.4103
Slope adaptive filter (32×32)	4910	1.2024	0.4636
Improved Goldstein filter (11×11)	2954	1.0402	0.9302
Improved Goldstein filter (32×32)	2261	1.0614	0.9600
DnCNN method	15970	0.7744	0.6453
IPDnCNN method	**48**	**0.9998**	**0.0793**

B. *Adaptability experiments*

In order to compare the adaptability of different methods under different noise levels, 30 additional interferograms with different coherence values from 0.38 to 0.83 are tested and the MSE of different methods are shown in Fig. 11. It can be seen that the proposed method always has the lowest MSE. Considering the better MSE of a smaller window for traditional filters, the 11 × 11 window size is used in them.

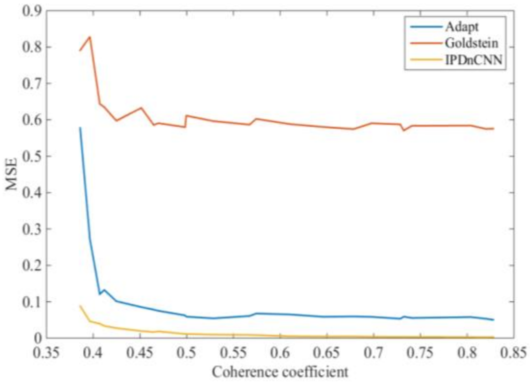
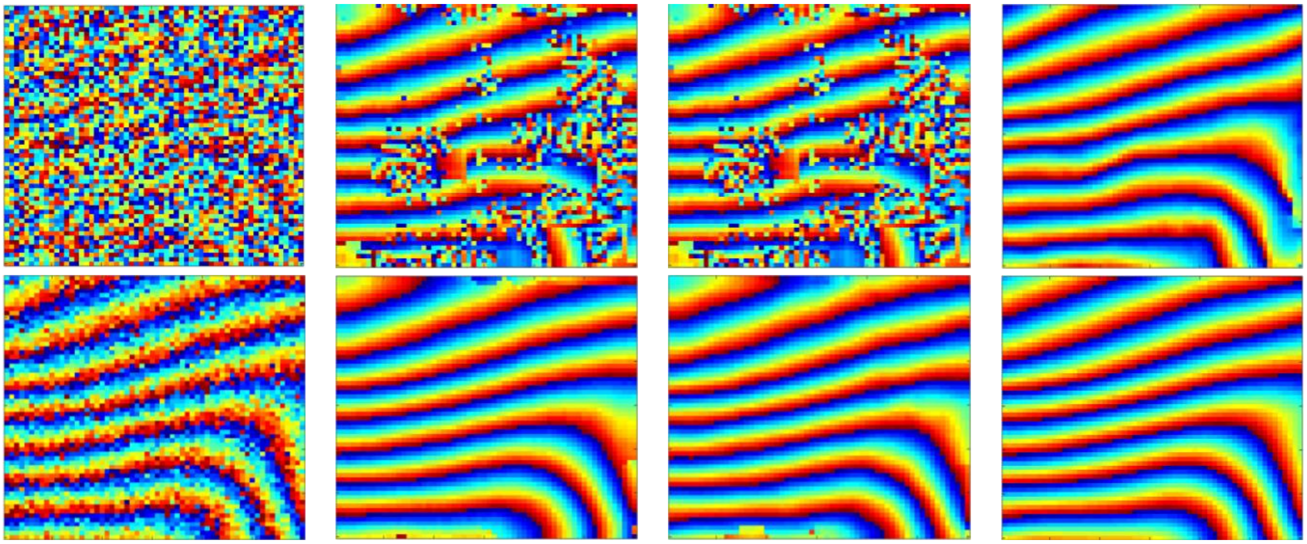


Fig. 11. MSE under different coherence values.

For a detailed comparison, we present the denoised results of a low-coherence interferogram (coherence=0.39) and a high-coherence interferogram (coherence=0.71), as shown in Fig. 12. The proposed method suppresses noise effectively even for the low coherence case while the conditional filters are worse.

The evaluation results are shown in TABLE IV. According to the results above, the proposed method has the best performance on noise reduction (smallest MSE and least Residues) as well as fringe preservation (EPI closest to 1).

.



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

Fig. 12. Noisy and denoised phase (top: low coherence, bottom: high coherence): (a) noisy phase, (b) slope adaptive filter, (c) improved Goldstein filter, (d) the IPDnCNN method.

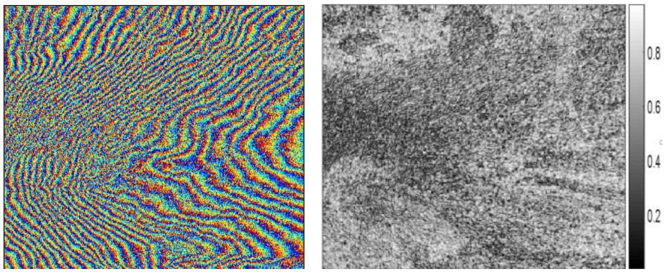
TABLE.Re. IV
Evaluation Results of Simulated Data

Interferogram	Interferogram with low coherence			Interferogram with high coherence		
	Residues	EPI	MSE	Residues	EPI	MSE
Clean phase	0	1	0	0	1	0
Noisy phase	1068	2.484	2.048	128	1.591	0.476
Slope adaptive filter	488	1.103	0.307	0	0.976	0.082
Improved Goldstein filter	92	1.033	0.765	0	0.990	0.601
IPDnCNN method	0	1.005	0.069	0	1.003	0.014

C. Experiments with real data

1) ERS SAR data

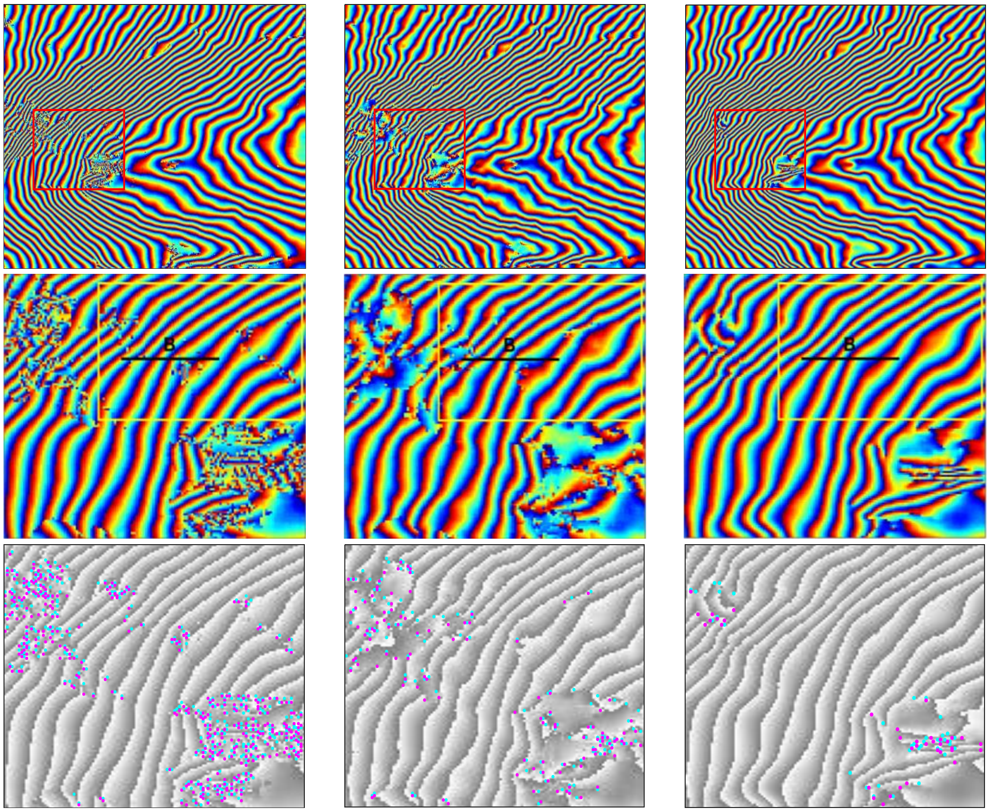
ERS SAR images over the ENTA Volcano in September and October 2000 are used as test data. The interferometric phase image of size 400 × 400 has dense fringes, and the mean coherence value is only 0.537. The interferometric phase and the coherence value are shown in Fig. 13.



(a) (b)

Fig. 13. ERS Interferogram: (a) interferometric phase, (b) coherence coefficient

The denoised results by the three methods are shown in Fig. 14. Each group contains the denoised phase, the enlarged area in the red rectangle, and the residue distribution. The interferometric fringes in the enlarged area are dense with a low coherence value of 0.419. It can be seen that the IPDnCNN method has reduced noise significantly while preserving the edge, whereas the slope adaptive filter and the improved Goldstein filter are less capable of denoising the interferometric phase.



(a) (b) (d)

Fig. 14. Denoised results of different methods: (a) slope adaptive filter, (b) improved Goldstein filter, (c) the IPDnCNN method.

To further verify the improvement produced by the IPDnCNN method, a cross-section is extracted in region B. As shown in Fig. 15, the phase obtained from the IPDnCNN method is relatively continuous, while those obtained from the other three still show some abnormality caused by residues and edge blur.

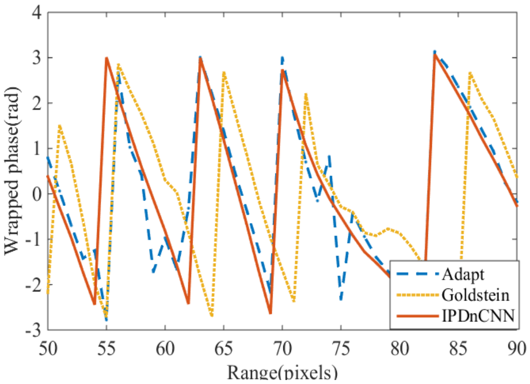


Fig. 15. Cross-sections through the denoised phase of real data.

A quantitative evaluation is also performed to compare the denoised results. Due to the lack of clean phase, only the number of residuals and the residual phase standard deviation (RPSD) are calculated. The RPSD is carried out after removal of the local fringe frequency from the initial interferometric phase, and it reflects the smoothness of the residual phase. A smaller RPSD means a smoother phase with less noise. It is calculated using the following equation

$$\mathrm{RPSD}=\sqrt{\frac{\sum_{N}\left(\varphi_r(i,j)-\bar{\varphi}_r(i,j)\right)^2}{N-1}} \tag{8}$$

where $\varphi_r(i,j)$ is the residual phase obtained by removing the LFF from denoised phase, $\bar{\varphi}_r(i,j)$ is the linear phase ramp in a moving window of size 3×3, and N is the number of pixels in the whole image.

To reduce the possible effect of artifacts, we only evaluated the denoising performance in the yellow rectangle.

TABLE V
EVALUATION RESULTS OF REAL DATA

Iterferogram	Residues	RPSD
Noisy phase	1461	1.6586
Slope adaptive filter	78	1.1230
Improved Goldstein filter	19	1.0936
IPDnCNN method	**0**	**1.0694**

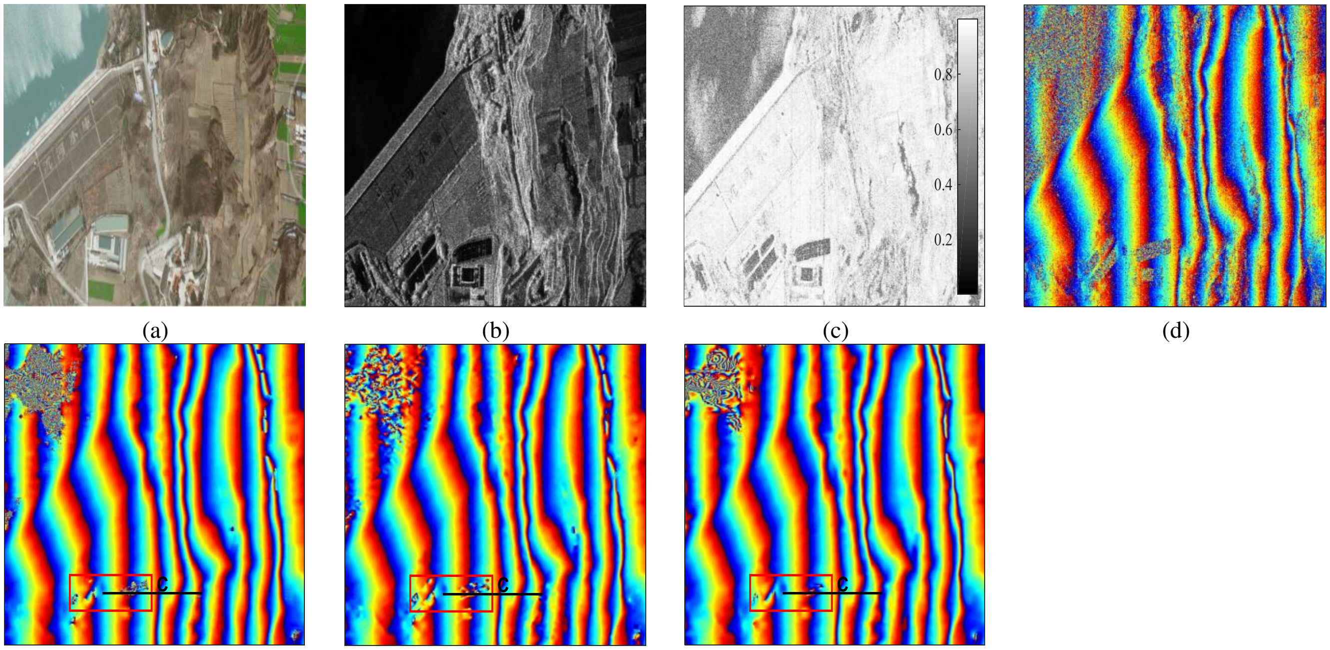
As shown in Table V, all methods can significantly reduce the number of residual points. Residues of the slope adaptive filter, improved Goldstein filter and the IPDnCNN method have been reduced by 94.6%, 98.7% and 100% respectively. Again, the IPDnCNN method gives the best result. For the RPSD results, we have a similar observation.

2) *NSAR data*

The interferograms obtained from a reservoir region in Shanxi, China, recorded by the NSAR system developed by the Nanjing Research Institute of Electronics Technology in March 2017, are chosen to conduct another experiment. The size of the observation area is of 775 × 775 pixels, and the terrain features are significantly different from that in the mountain area. The left side of this area is the reservoir. The slope of the reservoir dam is large, resulting in phase overlap. The amplitude image, coherence coefficient, interferometric fringes and denoised results are shown in Fig. 16.

Since the airborne data has high SNR and sparse stripes, all three methods have achieved a good noise reduction effect. In the low coherence region marked with the red box, the result of the IPDnCNN method, as shown in Fig. 16(g), contains less noise than that of the other two shown in Fig. 16(e) ~ (f). Fig. 17 shows a cross-section of the tangent C. It is obvious that the denoised phase from the IPDnCNN method is the cleanest and most continuous one, while those from the slope adaptive and improved Goldstein filters still show many unwanted phase jumps.

Since the signal in the water region is too weak to form coherent fringes, this region is excluded during quantitative evaluation, and the evaluation results are given in Table VI. Similar to the results of ERS data, the proposed method has produced the lest residues and smallest RPSD. It not only reduces the phase noise more effectively, but also preserves local fringe better.



(a) (b) (c) (d)
(e) (f) (g)

Fig. 16. NSAR data and the results: (a) optical image; (b) SAR amplitude image, (c) coherence coefficient, (d) interferometric phase, and denoised phase with the (e) slope adaptive filter, (f) improved Goldstein filter, (g) IPDnCNN methods.

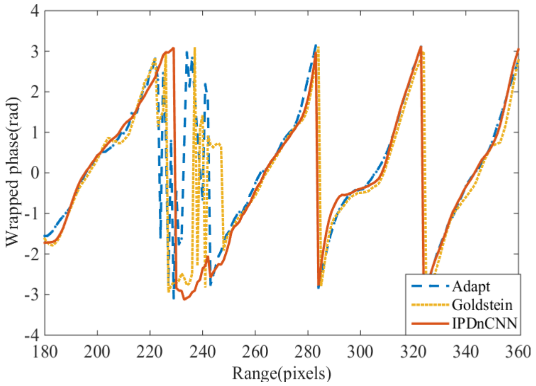


Fig. 17. Cross-sections through denoised phase of airborne data.

TABLE VI
EVALUATION RESULTS OF AIRBORNE DATA

Iterferogram	Residues	RPSD
Noisy phase	33627	1.1860
Slope adaptive filter	263	0.3436
Improved Goldstein filter	74	0.3407
IPDnCNN method	**22**	**0.3258**

For traditional filtering methods, the phase denoising performance is greatly affected by the filter window, which cannot make full use of the information contained in the entire image. IPDnCNN extracts image features by training massive data through a non-linear network structure and estimate phase noise with the entire image’s pixels. IPDnCNN not only describes the noise more precisely but also makes full use of the entire image, so it achieves a clear improvement in residual reduction and edge preservation.

To compare computational costs for the three methods, Table VI shows the running time of each experiment in this section using different methods on a computer with Intel(R) Core(TM) i5-5200U@2.2GHz CPU.

TABLE VI
RUNNING TIME (S)

Method	Simulated data	ERS data	Airborn data
Slope adaptive filter	383	322	317
Improved Goldstein filter	479	0.368	466
IPDnCNN method	**85**	**48**	**125**

As shown, although data training takes a lot of time, once the network is well-trained, the proposed IPDnCNN is more efficient than traditional methods.

CONCLUSIONS

In this paper, the convolutional neural network is introduced to InSAR phase denoising. In contrast to the existing phase denoising methods which directly predict the complex phase fringes, the proposed method estimates phase noise first and then removes them from the noisy interferogram. The proposed IPDnCNN is constructed based on DnCNN. Sine and cosine values of the interferometric phase are used as the input to the network so that it can avoid misjudgment of phase fringe edges in noise detection. The loss function is redesigned and network training parameters are modified to deal with the phase noise reduction problem. Moreover, the proposed IPDnCNN increases the patch size and two convolution layers to utilize the phase information more effectively. As demonstrated by experimental results using both simulated and real SAR data, the proposed method has achieved the best performance in noise reduction while preserving fringe edges. Like other deep learning methods, the data training process is time-consuming, but a well-trained network can effectively improve the efficiency of data processing.

In the current work, phase noise and clean phase are used as training samples. As part of our future work, we will try to improve the IPDnDNN model to divide the InSAR interferogram into trip point, noise point, overlap mask point, shadow point, and so on, so that the overlap and shadow areas can be detected in advance to improve the quality of the denoised phase.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China under Grant 61471020 and Shanghai Aerospace Science and Technology Innovation Fund under Grant SAST2019-026.

IV. REFERENCES

[1] E. Rodriguez and J. M. Martin, "Theory and design of interferometric synthetic aperture radars," *IEE Proceedings F - Radar and Signal Processing*, vol. 139, no. 2, pp. 147-159, Apr. 1992.

[2] R. Bamler, P. Hartl, "Synthetic aperture radar interferometry", *Inverse Probl.*, vol. 14, no. 4, pp. R1-R54, 1998.

[3] D.C. Ghiglia and L.A. Romero. "Minimum Lp-norm two-dimensional phase unwrapping." *Journal of the Optical Society of America A*, vol. 13, no. 10, pp. 1999–2013, 1996.

[4] P.A. Rosen *et al.*, "Synthetic aperture radar interferometry," in *Proceedings of the IEEE*, vol. 88, no. 3, pp. 333-382, Mar. 2000.

[5] J.S. Lee, K. P. Papathanassiou, T. L. Ainsworth, M. R. Grunes and A. Reigber, "A new technique for noise filtering of SAR interferometric phase images," *IEEE Transactions on Geoscience and Remote.*

[6] Z. Li, Z. Bao, H. Li and G. Liao, "Image autocoregistration and InSAR interferogram estimation using joint subspace projection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 2, pp. 288-297, Feb. 2006.

[7] S. Zhang, J. Tang, M. Chen, S. Zhu and H. Yang, "Image Autocoregistration and Interferogram Estimation Using Extended COMET-EXIP Method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 12, pp. 4204-4218, Dec. 2010.

[8] L. Denis, F. Tupin, J. Darbon and M. Sigelle, "Joint Regularization of Phase and Amplitude of InSAR Data: Application to 3-D Reconstruction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3774-3785, Nov. 2009.

[9] H. Li *et al.*, "A Modification to the Complex-Valued MRF Modeling Filter of Interferometric SAR Phase," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 3, pp. 681-685, Mar. 2015.

[10] W. Ben Abdallah and R. Abdelfattah, "A Joint Markov Random Field Approach for SAR Interferogram Filtering and Unwrapping," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 9, no. 7, pp. 3016-3025, Jul. 2016.

[11] G. Vasile, E. Trouve, Jong-Sen Lee and V. Buzuloiu, "Intensity-driven adaptive-neighborhood technique for polarimetric and interferometric SAR parameters estimation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1609-1621, Jun. 2006.

[12] C. A. Deledalle, L. Denis and F. Tupin, "NL-InSAR: Nonlocal Interferogram Estimation," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 49, no. 4, pp. 1441-1452, Apr. 2011.

[13] G. Poggi, F. Sica, L. Verdoliva, G. Fornaro, D. Reale and S. Verde, "Non-local methods for filtering interferometric SAR datasets," 2012 *Tyrrhenian Workshop on Advances in Radar and Remote Sensing* (TyWRRS), Naples, 2012, pp. 136-139.

[14] J. Li, Z. Li, Z. Bao, Y. Hou and Z. Suo, "Noise Filtering of High-Resolution Interferograms Over Vegetation and Urban Areas With a Refined Nonlocal Filter," *IEEE Geoscience and Remote Sensing Letters,* vol. 12, no. 1, pp. 77-81, Jan. 2015.

[15] C.-A. Deledalle, V. Duval, and J. Salmon, “Non-local Methods with Shape-Adaptive Patches (NLM-SAP),” J. Math. Imag. Vis., vol. 43, no. 2, pp. 103–120, Jun. 2012.

[16] C. Lopez-Martinez and X. Fabregas, "Modeling and reduction of SAR interferometric phase noise in the wavelet domain," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 12, pp. 2553-2566, Dec 2002.

[17] X. Zha, R. Fu, Z. Dai and B. Liu, "Noise Reduction in Interferograms Using the Wavelet Packet Transform and Wiener Filtering," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 3, pp. 404-408, Jul. 2008.

[18] Y. Bian and B. Mercer, "Interferometric SAR Phase Filtering in the Wavelet Domain Using Simultaneous Detection and Estimation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 4, pp. 1396-1416, Apr. 2011.

[19] R. M. Goldstein and C. L. Werner, “Radar interferogram filtering for geophysical applications,” Geophys. Res. Lett., vol. 25, no. 21, pp. 4035– 4038, Nov. 1998.

[20] M. Jiang *et al*., "The Improvement for Baran Phase Filter Derived From Unbiased InSAR Coherence," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 7, pp. 3002-3010, Jul. 2014.

[21] M. Jiang *et al*., "A hybrid method for optimization of the adaptive Goldstein filter." *ISPRS Journal of Photogrammetry and Remote Sensingg, vol. 98, pp. 29-43, Dec. 2014*.

[22] I. Baran, M. P. Stewart, B. M. Kampes, Z. Perski and P. Lilly, "A modification to the Goldstein radar interferogram filter," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 41, no. 9, pp. 2114-2118, Sept. 2003.

[23] Z. Li *et al.,* "Improved filtering parameter determination for the Goldstein radar interferogram filter." *Isprs Journal of Photogrammetry & Remote Sensing*, vol. 63, no. 6, pp. 621-634, Nov. 2008.

[24] E. Trouve, J. M. Nicolas and H. Maitre, "Improving phase unwrapping techniques by the use of local frequency estimates," IEEE *Transactions on Geoscience and Remote Sensing*, vol. 36, no. 6, pp. 1963-1972, Nov 1998.

[25] Y. Wang, X. Zhu and R. Bamler, "Retrieval of phase history parameters from distributed scatterers in urban areas using very high resolution SAR data", *ISPRS J. Photogramm. Remote Sens.*, vol. 73, pp. 89-99, Sep. 2012.

[26] Z. Suo, Z. Li and Z. Bao, "A New Strategy to Estimate Local Fringe Frequencies for InSAR Phase Noise Reduction," IEEE *Geoscience and Remote Sensing Letters*, vol. 7, no. 4, pp. 771-775, Oct. 2010.

[27] Z. Ding, Z. Wang, S. Lin, T. Liu, Q. Zhang and T. Long, "Local Fringe Frequency Estimation Based on Multifrequency InSAR for Phase-Noise Reduction in Highly Sloped Terrain," IEEE *Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1527-1531, Sept. 2017.

[28] Q. Feng, H.Xu, Z.Wu, Y.You, W.Liu and S.Ge, "Improved Goldstein Interferogram Filter Based on Local Fringe Frequency Estimation," *Sensors*, 2016.

[29] H. A. Zebker and K. Chen, "Accurate estimation of correlation in InSAR observations," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 2, pp. 124-127, Apr. 2005.

[30] J. S. Lee, K. W. Hoppel, S. A. Mango and A. R. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 5, pp. 1017-1028, Sept. 1994.

[31] K. Isogawa, T. Ida, T. Shiodera and T. Takeguchi, "Deep Shrinkage Convolutional Neural Network for Adaptive Noise Reduction," *IEEE Signal Processing Letters*, vol. 25, no. 2, pp. 224-228, Feb. 2018.

[32] A.Suksmono and A. Hirose. "InSAR image restoration by using stochastic complex-valued neural network." *Kes* 2002.

[33] Saba, Tanzila, A. Rehman, and G. Sulong. "An intelligent approach to image denoising." *Journal of Theoretical & Applied Information Technology,* vol. 7, no. 1, pp. 63-6, 2010.

[34] H. C. Burger, C. J. Schuler and S. Harmeling, "Image denoising: Can plain neural networks compete with BM3D?" *IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, 2012, pp. 2392-2399.

[35] K. Dabov, A. Foi, V. Katkovnik, K. O. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering", *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 80-2095, Aug. 2007.

[36] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," IEEE *Transactions on Image Processing*, vol. 26, no. 7, pp. 3142-3155, Jul. 2017.

[37] A. Qayyum *et al*. "Scene classification for aerial images based on CNN using sparse coding technique." *International Journal of Remote Sensing*, vol. 38, no. 8-10, pp. 2662-2685, Mar.2017.

[38] Y. Chen and T. Pock, "Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration," 769-776 *Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1256-1272, Jun. 2017.

[39] V. Jain and H. S. Seung, "Natural image denoising with convolutional networks." *International Conference on Neural Information Processing Systems*, pp. 769-776, 2008.

[40] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition", *Int. Conf. on Learning Representations*, 2015.

[41] K. He, X. Zhang, S. Ren and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification", Proc. IEEE *Int. Conf. Comput. Vis*., pp. 1026-1034, Dec. 2015.

[42] G. Franceschetti, A. Iodice, M. Migliaccio and D. Riccio, "A novel across-track SAR interferometry simulator," IEEE *Transactions on Geoscience and Remote Sensing*, vol. 36, no. 3, pp. 950-962, May 1998

[43] A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet classification with deep convolutional neural networks", Proc. Adv. Neural Inf. Process. Syst., vol. 25, pp. 1097-1105, 2012.