A Novel Quality-Guided Two-Dimensional InSAR Phase Unwrapping Method via GAUNet

Hai Wang , Jun Hu , *Member, IEEE*, Haiqiang Fu , Changcheng Wang , *Member, IEEE*, and Zhenhai Wang

Abstract—Phase unwrapping (PU) has always been a critical and challenging step in interferometric synthetic aperture radar (InSAR) data processing. Inspired by existing research, i.e., the PGNet, we propose a novel quality-guided 2-D InSAR PU method via deep learning, and regard PU as a two-stage process. In the first stage, the ambiguity gradient is estimated using the proposed global attention U-Net (GAUNet) architecture, which combines the classic U-Net structure and global attention mechanism. Then, in the second stage, the classical PU framework (e.g., the L1- or L2-norm) is applied as a post-processing operation to retrieve the absolute phase. Since class imbalance is a key factor affecting the estimation of ambiguity gradient, different strategies based on four commonly used quality maps are adopted to deal with the problem. The quality map is not only input as additional information for the guidance of the training process, but also participates in the construction of loss function. As a result, GAUNet can pay more attention to the nonzero ambiguity gradients. By using the number of residues as the evaluation metric, we can choose the optimum strategy for the restoration of the absolute phase. In addition to the simulated interferograms, the proposed method is tested both on a real topographic interferogram exhibiting rugged topography and phase aliasing and a differential interferogram measuring the deformation from MW 6.9 Hawaii earthquake, all yield state-of-art performance when comparing with the widely used traditional 2-D PU methods.

Index Terms—2-D phase unwrapping (2-D PU), ambiguity gradient, class imbalance, deep learning (DL), interferometric synthetic aperture radar (InSAR), quality-guided.

I. INTRODUCTION

INTERFEROMETRIC synthetic aperture radar (InSAR) technology has grown to maturity in recent decades, and has been widely used in the reconstruction of digital elevation model (DEM) and the detection of ground deformation [1], [2]. The basic principle of InSAR is to separate the target signal from the phase difference of coregistered SAR image pairs (i.e., interferometric phase) to obtain the desired information of elevation or deformation. Due to the limitation of the InSAR

imaging hardware system, the obtained interferometric phase is wrapped in the range of $(-\pi, \pi]$. Therefore, discontinuous phase occurs at $\pm\pi$ even though the phase satisfies the irrotational property. Phase discontinuities can be removed by assigning the correct integer multiples of 2π to each pixel of the interferogram ($2k\pi$ with k represents the ambiguity number). The process of recovering the absolute phase from the obtained wrapped phase is called phase unwrapping (PU) [3], [4].

Because of the ill-posed relationship between the wrapped phase and the absolute phase, additional prior information or constraint will be necessary for obtaining a unique PU solution. Traditional 2-D InSAR PU methods are almost based on the constraint of phase continuity assumption. Phase continuity assumption, also known as Itoh's condition, requires that the absolute phase difference between adjacent pixels in the interferogram does not exceed π [5]. Under ideal conditions, the complex InSAR PU process can be achieved by simple path integration. However, in practical applications, due to the presence of noise and phase aliasing (caused by such as shadow, layover, and foreshortening) in the SAR image, phase continuity assumption in the areas with low coherence or steep terrain will often fail, resulting in unwrapping errors and negative impact on the final results. Therefore, PU is a key and thorny problem in InSAR data processing.

Plenty of 2-D PU algorithms have been developed to tackle the problem of PU in recent decades, which can be broadly classified into two categories: path-following algorithms and optimization-based algorithms. Path-following algorithms have high achievable accuracy in high-quality areas but are easy to form isolated islands in areas with low signal-to-noise ratio. Global optimization algorithms have stronger robustness, but usually induce unwrapping failure in areas with steep phase gradients and propagate the unwrapping errors on the whole image. That is because as a precondition to traditional PU methods, phase continuity assumption seems unreliable in many cases. Actually, failure of phase continuity assumption is the intrinsic cause of PU errors and is a key issue demanding prompt solution in all traditional PU methods.

In the past few years, deep learning (DL) has yielded unusually brilliant results in various fields including remote sensing [6], [7]. Especially, deep convolutional neural network (DCNN) has shown unparalleled advantages in many tasks such as target recognition, image classification, and image segmentation. However, the applications of DL in remote sensing mainly concentrate in optical remote sensing, whereas there are few applications in radar remote sensing especially InSAR [8], [9]. Recently, some researchers in InSAR field are actively exploring

the combination of DL and InSAR, attempting to draw support from DL to solve the problems existing in traditional InSAR technique [10].

Hitherto, DL-based PU methods have shown their elegant demeanor. We know that the phase continuity assumption is actually proposed by the designers of the PU algorithms based on their own experience. However, the latest researches such as AlphaGo have indicated that in some respects, the ability of artificial intelligence (AI) to sum up experience is stronger than that of humans. The biggest advantage of DL-based PU method is that it can get rid of the constraints of phase continuity assumption and can learn end-to-end directly through data-driven instead of model-driven method used in traditional 2-D PU. Therefore, the computer has the ability to autonomously extract features from the wrapped phase after appropriate training, and can obtain the true phase directly or through a certain post-processing process.

Several existing researches have proved the effectiveness of DL-based PU methods. Spoorthi *et al.* [11] proposed PhaseNet, which regards PU as a semantic segmentation problem, and is the first application of DCNN in PU. In [12], a fully connected neural network structure is applied. The network has the ability to extract the ambiguity number K from the wrapped phase after training, absolute phase can thus be recovered through a simple post-processing process. Zhang *et al.* [13] and Zhang *et al.* [14] adopted similar strategies as PhaseNet to solve the PU problem. Wang *et al.* [15] used the absolute phase instead of ambiguity number K as ground truth and proposed a DLPU structure, which can directly extract the absolute phase from the wrapped phase after training without any post-processing step. Spoorthi *et al.* [16] proposed PhaseNet 2.0 on the basis of PhaseNet, utilized DenseNet to implement PU tasks, which greatly increases the classes of ambiguity number that can be accurately predicted and enhances the network's performance by proposing a new residual loss.

The aforementioned works were all achieved by researchers in optical field. Compared with optical PU, the application scenarios of InSAR PU are usually more complicated due to the effect of phase aliasing and noise; therefore, it is ineffective to use ambiguity number or absolute phase as ground truth in InSAR PU situations. Zhou *et al.* [17] proposed PGNet in 2020, which is the first combination of DL and InSAR PU. In this article, phase gradients are used as ground truth, and DCNN is used to replace the phase continuity assumption to estimate the horizontal and vertical phase gradient. The 2-D InSAR PU is transformed into a three-class segmentation problem, which greatly improves the generalization performance and noise robustness of the algorithm. Conventional methods such as minimum cost flow (MCF) can be used for postprocessing to obtain the final unwrapping phase using the estimated phase gradients. Sica *et al.* [18] input the wrapped phase and its corresponding coherence simultaneously into the network for training and also use the wrap-count gradients as ground truth. Both of the two methods show state-of-art performance when comparing with the traditional PU methods. Zhou *et al.* [19] made a review on applications of AI in the PU domain, where the existing methods are well summarized and analyzed. However, to the best of our knowledge, there is no pertinent research focus on the generalization performance and class imbalance problem of DL-based PU method, and nobody has tested the performance of DL-based PU method on the interferogram with ground deformation caused by the earthquake.

In this article, we propose a novel quality-guided 2-D InSAR PU method based on a DCNN structure, which is referred to as global attention U-Net (GAUNet), to improve the accuracy of the ambiguity gradient estimation and ultimately to obtain a more accurate PU result. Specifically, compared with other existing works, we make the following new contributions.

1) Unlike the existing researches, we generated a multisource dataset with various noise levels for training process. Parameters from TerraSAR-X/TanDEM-X with X-band, Sentinel-1 with C-band, and ALOS-2 with L-band are used to generate the training data sets, and noise is added to the simulated phase based on constant mean coherence or coherence derived from the real processed interferograms.

2) To deal with the adverse effects of class imbalance, global attention mechanism is added to the U-Net architecture to build GAUNet architecture. Therefore, GAUNet can dynamically focus on certain parts of the input, where the distribution of nonzero ambiguity gradients is denser.

3) Four kinds of quality maps are not only evaluated as an input feature to the GAUNet, but also participate in the formulation of the loss function in this work.

4) Eventually, we demonstrate the performance of GAUNet on both simulated and real interferograms. Deserved to be mentioned, we verified the performance of DL-based PU method in the interferogram with the deformation caused by an earthquake for the first time.

The remainder of the article is organized as follows. In Section II, we introduce the principle and categories of traditional 2-D InSAR PU in detail, then review the mechanism of current DL-based PU methods. Materials and details of the proposed method are presented in Section III. In Section IV, the performance of GAUNet-PU is evaluated both on simulated and real SAR interferograms. Finally, Section V concludes this article.

II. PRINCIPLE AND RELATED WORK

In this section, we first give a detailed introduction of the principle and categories of the traditional 2-D PU method, then overview the implementation mechanism of the DL-based PU method.

A. 2-D Phase Unwrapping and Ambiguity Gradient

The relationship between the obtained wrapped phase φ and the unknown absolute phase ψ in an interferogram can be stated as follows:

$$\varphi(s) = \psi(s) - 2k(s)\pi, \text{ where } \varphi(\mathrm{s}) \in (-\pi, \pi] \, k(s) \in Z \quad (1)$$

where s denotes the pixel position, $k(s)$ is the unknown ambiguity number of the sth pixel, and Z denotes set of integers. From (1), we can find that 2-D PU is actually an ill-posed inverse problem. With one observation and two unknowns, the unwrapping phase can not be solved directly. As mentioned in Section I, phase continuity assumption can be used as a restriction for traditional 2-D PU. If phase continuity assumption
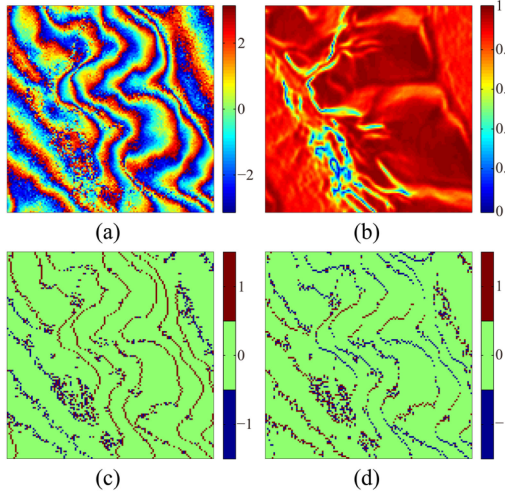
Fig. 1. Example of (a) wrapped phase and its corresponding: (b) coherence, (c) HG, and (d) VG.
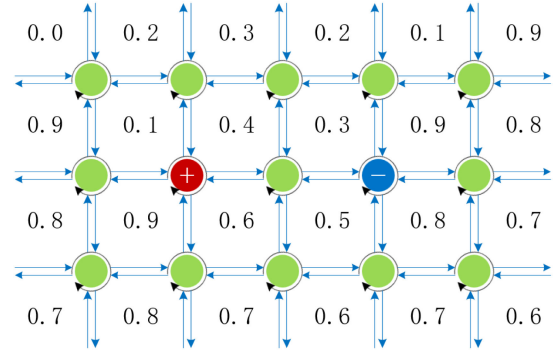

Fig. 2. Illustration of identify residues from the wrapped phase grid. The red circular represents the positive residue, whereas the blue circular represents the negative residue. Green circulars represent nodes without residue, i.e., the loop-integration value is 0.

holds at any position in the interferogram, PU can be easily realized by a simple integration process.

Ambiguity gradient $\Delta_k(s, s-1)$ refers to the difference of ambiguity number between adjacent pixels according to (2). For 2-D InSAR PU problem, there are two kinds of ambiguity gradients in the azimuth and range (i.e., horizontal and vertical) directions

$$\Delta_k(s, s-1) = k(s) - k(s-1)$$

$$\text{where } k(s) = \text{round}((\psi(s) - \varphi(s))/2\pi). \quad (2)$$

Under the constraints of phase continuity assumption, ambiguity gradient can be defined as

$$\Delta_k(s, s-1) = \begin{cases} 1, & \varphi(\mathrm{s}) - \varphi(\mathrm{s}-1) < -\pi \\ 0, & |\varphi(\mathrm{s}) - \varphi(\mathrm{s}-1)| \leq \pi \\ -1, & \varphi(\mathrm{s}) - \varphi(\mathrm{s}-1) > \pi . \end{cases} \quad (3)$$

Fig. 1 shows an example of the wrapped phase and its corresponding coherence, horizontal ambiguity gradient (HG), and vertical ambiguity gradient (VG). It can be found that the distribution of nonzero gradients is closely related to the fringe pattern and the presence of noise. Coherence can reflect the distribution of nonzero gradients to some extent.

B. Path-Following Algorithms

The path-following algorithms utilize external information to choose the most reasonable integration path. Goldstein's branch-cut (BC) method is a typical path-following algorithm, which uses the residues to assist the choice of the integration path [20], [21]. Simply stated, residues are recognized by calculating a 2×2 loop-integration of the wrapped phase difference (see Fig. 2). Integration path via nonzero residues will cause unwrapping errors, which can transfer from the low-quality areas to the high-quality areas of the input interferogram. BC method can balance the positive and negative residues by setting reasonable branch cuts. Another path-following algorithm received almost as gospel is the quality-guided phase unwrapping (QGPU) method [22], [23]. The value of the quality map represents the

quality of the corresponding interferogram. It is believed that high-quality pixels are not likely to cause unwrapping errors. Therefore, researchers insist that high-quality pixels should be considered first when selecting the optimum integration path. That is to say, QGPU unwraps high-quality areas at the beginning and then gradually expands to low-quality areas.

C. Optimization-Based Algorithms

Optimization-based algorithms regard PU as a global optimization problem. It believes that a dependable PU result can be obtained by solving the global optimal solution of the certain constraint minimum energy function [24]. According to [4], in its general form the formulation of optimization-based algorithms can be expressed as follows:

$$\arg\min_{\psi(s)} \sum_{(s,s-1)} f(\psi(s) - \psi(s-1) - \widehat{\Delta}_\psi(s, s-1)). \quad (4)$$

Depending on the choice of PU model $f(\cdot)$, optimization-based algorithms can be divided into two sorts: minimum-norm algorithms and statistics-based algorithms. MCF [25] and least square [26] are two typical minimum-norm methods. Representative statistics-based methods include phase unwrapping via max flows (PUMA) [27] and statistical-cost, network-flow algorithm for phase unwrapping (SNAPHU) [28]. Some specific details of optimization-based algorithms will be elaborated later.

D. Deep Learning Based Algorithms

According to existing research [11]–[18], the key point of DL-based PU is to determine an appropriate learning object (i.e., the ground truth), so as to investigate the PU problem from DL perspective. Through clever transformation, PU can be converted into an image segmentation task, which is DL's area of expertise.

So far, three learning objects, i.e., the absolute phase, the ambiguity number, and the ambiguity gradient, can be utilized for DL-based PU. DL-based PU methods taking absolute phase as learning object can be regarded as one-stage processes, whereas those taking ambiguity number or gradient as learning object are two-stage processes, or in other words, additional steps are required for obtaining the unwrapped phase. As mentioned

earlier, using ambiguity gradient as ground truth seems more reliable for the complex scenarios of InSAR, and has the ability to process a larger size of interferogram.

III. MATERIALS AND PROPOSED METHOD

Inspired by the existing research [17], we regard 2-D InSAR PU as a two-stage programming approach. For the first stage, GAUNet based on DCNN is proposed to estimate the ambiguity gradient $\widehat{\Delta}_k(s, s-1)$. It is worth emphasizing that the true ambiguity gradient $\Delta_k(s, s-1)$ in (2) is assumed only three values (i.e., 0 or ± 1) due to the limitation of DCNN's capacity in dealing with the minority class. Hence, we replace ambiguity gradient larger than 1 with 1, and that smaller than -1 with -1. It is acknowledged that this range can cover most of the common situations [17], [18]. According to (2), the estimation of the ambiguity gradient can be regarded as a semantic segmentation problem with three classes $(0, \pm 1)$ [29]. Due to the unbalanced quantity and uneven distribution of the nonzero ambiguity gradients, class imbalance becomes a key factor restricting the results. We adopt several strategies to deal with the problem, which will be introduced in details in this section. For the second stage, the traditional optimization-based PU framework is employed as a postprocessing process to obtain the final unwrapped phase based on the ambiguity gradient estimated in the first stage. Therefore, (4) can be converted to the formulation as follows:

$$\arg \min_{k(s)} \sum_{(s, s-1)} f(k(s) - k(s-1) - \widehat{\Delta}_k(s, s-1)). \quad (5)$$

From (5), we can see that the reliability of the PU result is highly dependent on the correctness of $\widehat{\Delta}_k(s, s-1)$. If the estimation of $\widehat{\Delta}_k(s, s-1)$ is totally correct, the result of PU will be unambiguous. However, in low-quality areas, the ambiguity gradient estimation based on the phase continuity assumption has proved wildly inaccurate. In this case, errors will always exist in the PU solution whatever kind of $f(\cdot)$ is chose. Compared with the phase continuity assumption, the proposed method shows significant accuracy improvement in the estimation of ambiguity gradient under various noise levels; thus, we can obtain more robust PU results even using identical optimization framework.

A. Quality Map

QGPU is a very effective PU method, which can make a good balance between efficiency and accuracy. The result of QGPU is highly depend on the quality map. The range of value in the quality map is between 0 and 1, which can evaluate the goodness of each pixel in the corresponding interferogram. There are four most commonly used quality maps in InSAR PU, i.e., the correlation map, the phase derivative variance map, the maximum phase gradient map, and the pseudocorrelation map [3]. Now, we will briefly introduce the definitions of these quality maps. Among them, the correlation map is what we commonly referred to as coherence, and can be defined as follows:

$$\gamma_{(m,n)} = \frac{\Sigma u_{(i,j)} v_{(i,j)}^*}{\sqrt{\Sigma |u_{(i,j)}|^2 \Sigma |v_{(i,j)}|^2}} \quad (6)$$

where $u_{(i,j)}$ and $v_{(i,j)}$ represent the multilook averaging plural pixels performed in the $k \times k$ neighborhood centered at pixel (m, n) of the coregistered master-slave interferogram, and $v_{(i,j)}^*$ and $|v_{(i,j)}|$ represent the complex conjugate and the modulus of $v_{(i,j)}$, respectively.

The phase derivative variance is a frequently-used quality map, which is closest to the distribution of residues. It can be computed using (7). The results of selecting the phase derivative variance tend to be more reliable especially when the terrain of the application scenario is steep

$$Z_{m,n} = \frac{\sqrt{\sum \left(\Delta_{i,j}^x - \overline{\Delta_{m,n}^x}\right)^2} + \sqrt{\sum \left(\Delta_{i,j}^y - \overline{\Delta_{m,n}^y}\right)^2}}{k^2} \quad (7)$$

where k is the length of the smoothing window, (m, n) is the central pixel of the smoothing window, $\Delta_{i,j}^x$ and $\Delta_{i,j}^y$ are the partial derivative of the wrapped phase, and $\overline{\Delta_{m,n}^x}$ and $\overline{\Delta_{m,n}^y}$ are the mean values of the partial derivative in the smoothing window.

The maximum phase gradient measures the magnitude of the largest phase gradient (i.e., wrapped phase difference or partial derivative) of each pixel in the $k \times k$ neighborhood, and its definition is shown as follows:

$$G_{m,n} = \max\{\max\{|\Delta_{i,j}^x|\}, \max\{|\Delta_{i,j}^y|\}\}. \quad (8)$$

The pseudocorrelation is proposed to deal with the situation when the correlation coefficient is unknown, but there is a certain difference between them. Pseudocorrelation can be computed using the following equation:

$$|Z_{m,n}| = \frac{\sqrt{(\sum \cos \varphi_{i,j})^2 + (\sum \sin \varphi_{i,j})^2}}{k^2} \quad (9)$$

where k is the length of the smoothing window, (m,n) is the central pixel of the smoothing window, $\varphi_{i,j}$ is the wrapped phase, sum operations in (9) are taken in the smoothing window.

The corresponding phase derivative variance map, maximum phase gradient map, and the pseudocorrelation map of Fig. 1(a) are presented in Fig. 3. Through comparison, it is found that the phase derivative variance map [Fig. 3(a)] is quite similar to the correlation map [Fig. 1(b)]. The maximum phase gradient map [Fig. 3(b)] and the pseudocorrelation map [Fig. 3(c)] have a disposition to mark the area with steep terrain as low-quality, which will undoubtedly bring negative effect on the traditional 2-D PU method. However, nonzero ambiguity gradients tend to be very crowded in areas with steep terrain. Marking these areas as low-quality can make the DCNN pay more attention to them and may have a beneficial effect on the gradient estimation. Therefore, all four kinds of quality maps are tested in the following experiments.

B. Generation of Dataset

A high-quality dataset is one of the compelling requisites for DL-based method. Fortunately, unlike many other applications, there is a definitive input-output relationship between the wrapped phase and the ambiguity gradient in PU, as shown in (1) and (2). Therefore, we can generate our dataset effortlessly. The topographic component of interferometric phase is simulated to represent the true phase, which can be computed using the
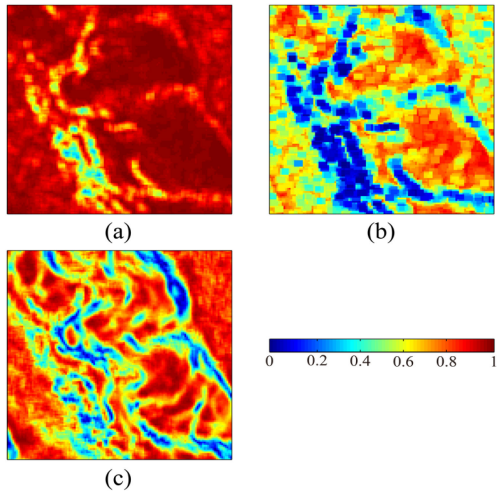
Fig. 3. Corresponding (a) phase derivative variance map, (b) maximum phase gradient map, and (c) pseudocorrelation map of Fig. 1(a).

TABLE I
SIMULATED PARAMETERS

Sensor	$\lambda(m)$	$B_\perp(m)$	$R(m)$	$\theta(°)$
TerraSAR-X/TanDEM-X	0.031	227.86	710344.5	46.3
Sentinel-1	0.055	159.60	876298.8	39.3
ALOS-2	0.236	316.73	793416.8	39.0

following equation:

$$\psi(m,n) = \frac{4\pi}{\lambda} \frac{B_\perp h(m,n)}{R \sin(\theta)} \qquad (10)$$

where $\psi(m,n)$ is the true phase of pixel (m,n), $B_\perp$ is the perpendicular baseline, λ is the wavelength, R is the orbit height, and θ is the incidence angle. Altitude $h(m,n)$ can be obtained by back-geocoding the existing DEM database, such as Shuttle Radar Topography Mission (SRTM) DEM or TanDEM-X DEM (TanDEM) [30], [31]. The final choice depends on the specific situation, mainly referring to the size of the holes in the DEM of the study areas. To enhance the noise robustness and generalization capability of our method, the acquisition parameters of the X-band TerraSAR-X and TanDEM-X satellites, the C-band Sentinel-1 satellite, and the L-band ALOS-2 satellite are all utilized to simulate the interferometric phase. The coverages of the data are shown in Fig. 4. Constant coherence and coherence derived from the real interferograms are used to add phase noise into the simulated wrapped phase, constituting two training data sets.

For data sets 1, the simulated parameters of three different SAR satellite sensors are listed in Table I. According to [32], the phase noise of pixel (m,n) is added by

$$\text{noi}(m,n) = \sqrt{\frac{1-\gamma(m,n)^2}{2L\gamma(m,n)^2}} \cdot \text{randn}(1) \qquad (11)$$

where $\gamma(m,n)$ is the coherence of pixel (m,n) and L represents the number of looks. Constant mean coherence ranging from 0.4 to 1 with 0.05 as the interval is used to add noise into the simulated interferometric phase.

For data sets 2, the parameters obtained from real interferograms are used to simulate the interferometric phase referring to (10). Because of the large quantity, the specific details of the simulated parameters will not be explained here for simplicity. It is worth recalling that the length of the perpendicular baseline is between 5 and 320 m. Long spatial baselines show denser phase fringes and can represent the complex topographic interferograms, whereas short baselines result in sparse fringes, which are more consistent with the deformation interferograms. Coherence derived from the real interferograms is used to add phase noise according to (11).

We can wrap the simulated phase with different noise levels through a simple step, using

$$\varphi(m,n) = \text{angle}(\exp(j\psi(m,n))). \qquad (12)$$

Here, $\text{angle}(\cdot)$ is used to calculate the principal value of the argument of a complex number, this is what we call the wrapped phase. We cropped the simulated wrapped phase and its corresponding four kinds of quality maps to a fixed size of 128×128 as inputs, and cropped the HGs and VGs to the identical size as ground truth for guidance of the training process. Ultimately, we got a training-validation set comprising 164 726 samples (80% reserved for training).

C. Proposed Method

In this part, we first introduce the basic compositions of GAUNet, i.e., the U-Net structure and the global attention upsample (GAU) module, then give a detailed description of the GAUNet architecture, the loss function, and the training hyperparameters.

1) U-Net: Through observation and comparison, it can be found that there are many similarities between the process of estimating the ambiguity gradient from the wrapped phase and the medical segmentation task [33]. Analogous to medical images, the noisy wrapped phase images have blurred boundaries and complex gradients, more high-resolution information is thus required for accurate segmentation. On the other side, the semantics of the ambiguity gradient are relatively simple and clear; therefore, low-resolution information is also significant for target recognition.

U-Net is a classic medical semantic segmentation model. Since the unique U-shaped structure and skip-connection can combine low- and high-resolution information, U-Net is very appropriate for our task [34]. The architecture of classic U-Net is shown in Fig. 5.

2) Global Attention Upsample: GAU is an efficient attention-based decoder proposed in [36]. As mentioned above, it is very important to fuse low- and high-resolution information in our task.

The structure of GAU is shown in Fig. 6. By performing global average pooling, GAU can extract high-level global context to weight low-level features. Compared with the low-level features, high-level features have rich semantic information. Therefore, high-level features can provide guidance information for low-level feature maps to select category localization details, i.e., the localization of the nonzero ambiguity gradients in our task.
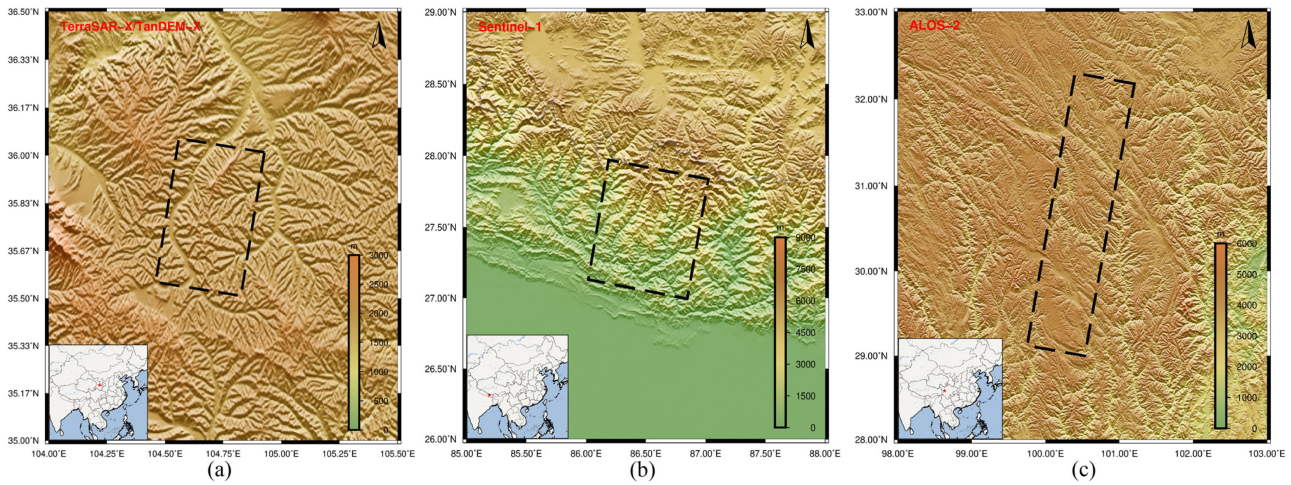
Fig. 4. Black dotted bordered rectangles denote the coverage of the data using for simulated InSAR system, red stars in the inset map show the location of the three study area. From left to right: (a) TerraSAR-X/TanDEM-X, (b) Sentinel-1, and (c) ALOS-2.
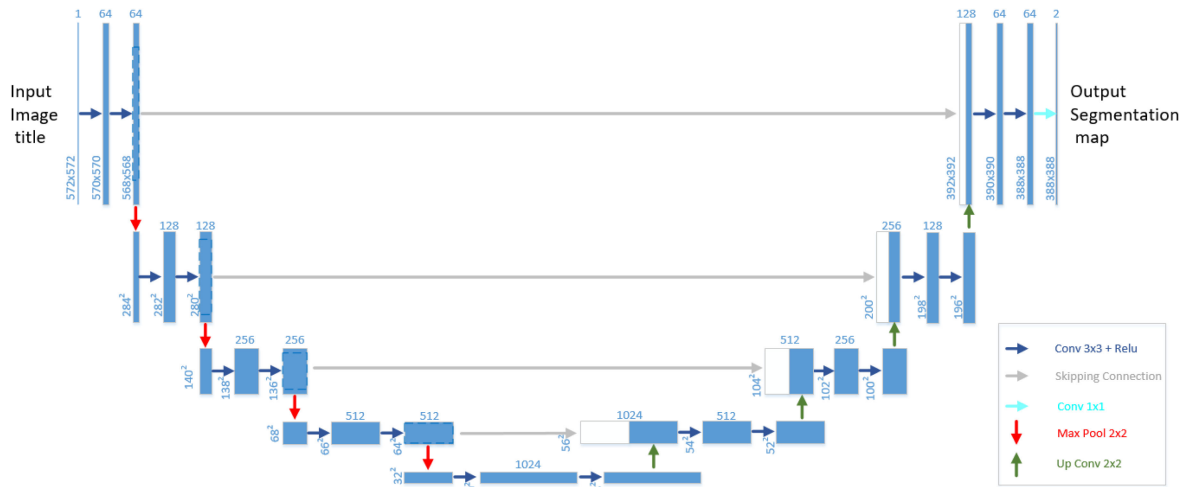

Fig. 5. Architecture of classic U-Net (modified from [34]). The network consists of a contracting path and an expanding path, exhibits U-shaped structure.
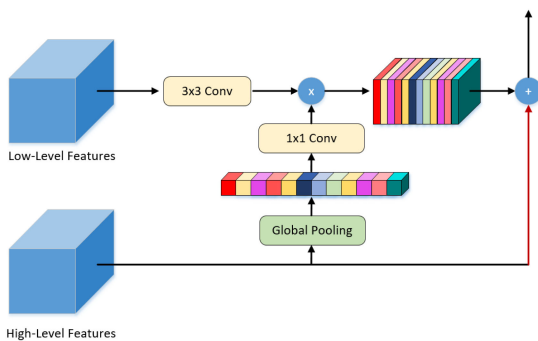

Fig. 6. Structure of GAU module (modified from [36]).

Specifically, a 3×3 convolution operation is performed on low-level features to reduce the number of channels in the DCNN feature map. The global context information generated from high-level features undergoes 1×1 convolution, batch normalization

(BN), and nonlinearity in turn, and then is multiplied by low-level features. Finally, the high-level features are added to the weighted low-level features and a gradual upsampling process is performed. GAU module can not only adapt to feature mapping at different scales with high efficiency, but also provide guidance information for mapping low-level features in a simple way.

3) GAUNet: By combining U-Net structure and GAU module, we proposed GAUNet specialized for ambiguity gradient estimation. As shown in Fig. 7, GAUNet is based on classic encoder-decoder architecture, and the overall structure of the network still maintains a U-like shape. The encoder module gradually reduces the resolution of the feature map and captures high-level semantic information while the decoder module restores spatial information step-by-step. Changes in the number of feature maps during the sampling process can be seen in Fig. 7.

As illustrated in Fig. 7, the encoder of GAUNet is consistent with the encoder of classical U-Net, containing four max-pooling layers sandwiched between five DoubleConv modules.
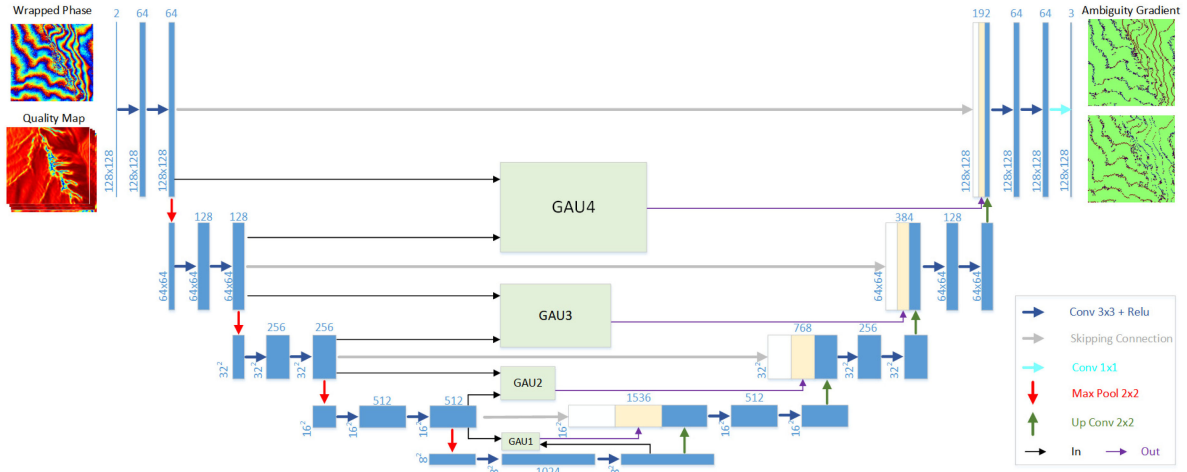
Fig. 7.   Network structure of the proposed GAUNet, which takes wrapped phase and its corresponding quality map as input and outputs HGs and VGs.

TABLE II
NETWORK CONFIGURATION FOR GAUNET

Module Name	Layer Name	Kernel size (Scale factor)	Stride	Padding
Double-Conv	Conv1+BN1+Relu1	3×3	1	1
	Conv2+BN2+Relu2	3×3	1	1
Pool	Max-pooling	2×2	-	-
UpConv	Upsample	2	-	-
	Conv1+BN1+Relu1	3×3	1	1

Through downsampling, the size of the feature map becomes one-sixteenth of the input size. The detailed configurations are shown in Table II.

We mainly modified the decoder subnetwork to specialized in treatment of the PU problem. First, we use a combination of upsample and 3×3 convolution layer, which is referred to as UpConv (detailed configurations are also shown in Table II) to replace the transposed convolution layer in the classic U-Net. Second, GAU module is used to provide global context information as guidance for low-level features, for indicating the positioning of pixels with nonzero ambiguity gradients. Other attention mechanisms, such as spatial attention mechanism, channel attention mechanism, and the combination of them [37]–[39], have been tried to introduce into our task. However, the experiments indicated that the results of ambiguity gradient estimation will be more inaccurate after the addition of these attention mechanisms. The specific representations are embodied in the decreases of mean IoU (MIoU) on validation set in terms of the evaluating indicators. MIoU on validation set is a generally used metric to evaluate the performances of different DCNNs after the training process. After that utilized skip connection, the upsampled features are concatenated with the corresponding features coming from the encoding path and the features generated by GAU module. Concatenated feature blocks are fed into DoubleConv module as the same as that in the encoding path, and the number of channels is changed back to the number before concatenating. The decoder also contains four UpConv modules. After four upsampling steps, the features are restored to the size of the original input image. The last DoubleConv module comprises an additional 1×1

convolutional layer, which compresses the output into three channels. Finally, through a three-way softmax activation, the final output is expressed as the desired pixel-wise class membership probabilities.

As shown in Fig. 7, the wrapped phase and its corresponding quality map (any one of four) are fed into the network as channels 1 and 2, respectively. As an additional input, the quality map can assist GAUNet to focus on the areas with denser nonzero gradients distribution automatically. The output is the horizontal or vertical components of the ambiguity gradient. It is noticed that compared with simultaneously obtaining HG and VG through a single training process, disjoint training procedures show higher accuracy of ambiguity gradient estimation. This can be ascribed to the highly nonisotropic SAR geometry and has been demonstrated in the experiments. Therefore, in our method, GAUNet is trained separately to extract HG and VG. Ultimately, two sets of different network parameters are estimated for each guiding strategy.

4) Loss Function and Training Process: An important starting point for solving the problem of class imbalance in classification or segmentation tasks is the loss function [40]. In this article, a weighted combination of quality-guided focal loss and dice loss is used to counter the class imbalance in ambiguity gradient estimation.

Focal loss is designed to solve the problem of unbalanced samples in one-stage target detection [41]. In essence, focal loss is a loss that can solve the class imbalance and mine hard samples by down-weighting the contribution of easy samples during training. The basis of focal loss is the cross entropy (CE) loss, which can be defined as follows:

$$L_{CE} = -\sum_{t=1}^{M} y_t \log(p_t) \qquad (13)$$

where M represents the number of classes, which is three in our task. y_t is a one-hot vector that takes the value of either 0 or 1. If the predicted value of ambiguity gradient is the same as ground truth, y_t takes 1; otherwise, y_t takes 0. p_t represents the membership probability of class t. On the foundation of the CE loss, focal loss adds a modulating factor $(1 - p_t)^{\gamma}$ and can be

written as

$$L_{\text{Focal}} = -\alpha_t (1 - p_t)^\gamma L_{\text{CE}}. \tag{14}$$

The modulating factor can adjust the corresponding loss value according to the difficulty of the sample. When a pixel is misclassified and p_t is small, the modulating factor is near 1 and the loss is unaffected. However, as $p_t \to 1$, the modulating factor goes to 0 and the loss for well-classified examples is down-weighted. The focusing parameter γ smoothly adjusts the rate at which easy examples are down-weighted. When $\gamma = 0$, L_{Focal} is equivalent to L_{CE}, and as γ is increased, the effect of the modulating factor is likewise increased. Different from [41], we dynamically adjust γ according to the value of the corresponding pixel in the quality map instead of setting it to a fixed value. In detail, γ is obtained by mapping the value of each pixel in the quality map, and the experiments have shown that by mapping the value of γ to [0,2], the best performance can be achieved. α_t is the weight of different classes. According to [41] and our experimental result, it can be found that the value of α_t will not have a great impact on the result in the multiclassification task, we thus set it to 1.

Dice coefficient is a commonly used index in image segmentation, especially in medical image segmentation field [42], [43]. According to the Lee Raymond Dice command, the Dice coefficient is a metric function to measure the similarity of sets. It is usually used to calculate the pixels between two samples. The formula of dice loss is defined as follows [42]:

$$L_{\text{Dice}} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} = 1 - \frac{2\text{TP}}{2\text{TP} + \text{FN} + \text{FP}}. \tag{15}$$

For our segmentation task, X and Y represents the ground truth and the predicted gradient images, respectively. TP, FP, and FN represent the number of true positive, false positive, and false negative samples, respectively. Dice loss is suitable when the samples are extremely unbalanced. If dice loss is used under normal circumstances, it will have an adverse effect on backpropagation and make the training process unstable. Therefore, dice loss is usually weighted with other loss functions. Performance analysis demonstrates that using a combination of focal loss and dice loss performs better than using one of them or other tested loss functions in our ambiguity gradient estimation task. So our final loss can be given by

$$L_{\text{Total}} = \alpha L_{\text{Focal}} + (1 - \alpha) L_{\text{Dice}}. \tag{16}$$

By experimental verification, α is set to 0.6. The proposed GAUNet is implemented in Pytorch. We trained on an NVIDIA RTX 3090 with 24 GB of GPU memory for 200 epochs, 80% of the generated dataset is used for training, and the remaining 20% is used for validation to prevent overfitting. The initial learning rate is set to $5e - 5$, and the batch size is set to 64. Parameters are initialized using Xavier initialization, and Adam optimizer is used to update and calculate the parameters that affect model training to approximate or reach the optimal value, thereby minimizing the loss function [44]. We use the cosine annealing strategy as scheduler to adaptively adjust the learning rate. As aforementioned, HG and VG are estimated separately, so we finally get two sets of different network parameters for each guiding strategy. By directly inputting the cropped wrapped phase (128×128) and its corresponding quality map into the trained GAUNet, the targeted HG and VG can be well estimated.

D. Absolute Phase Reconstruction

Estimating HG and VG is the first step of the 2-D PU task. As a consequence of the noise and phase aliasing, it is nearly impracticable to get a completely correct estimation of the ambiguity gradient. Therefore, the absolute phase cannot be reconstructed by a simple flood-fill integration process. According to [45], almost all the existing 2-D PU methods, such as MCF and LS, can be used to obtain the absolute phase with the estimated ambiguity gradients. In this study, the L1-Norm is chosen as $f(\cdot)$ to reconstruct the absolute phase; therefore, the problem is transformed into a nonlinear minimization process with integer variables. According to [25], (5) can be written as

$$\arg \min_{k(s)} \sum_{(s,s-1)} c(s, s-1) \cdot |k(s) - k(s-1) - \widehat{\Delta}_k(s, s-1)| \tag{17}$$

where $\widehat{\Delta}_k(s, s-1)$ is the ambiguity gradient estimated by GAUNet. By the way, all of the four kinds of quality maps described earlier can be used as $c(s, s-1)$ to weight the optimized process. The objective is to solve the minimum cost flow on the network, or more generally, to minimize the weighted number of locations where the estimated ambiguity gradient differs from the true ambiguity gradient. As an effective PU method based on L1 norm, minimum cost flow (MCF) has been developed and perfected in practice since it was proposed. Currently, MCF can be solved effectively by several algorithms according to [46]. Since MCF can reduce the time and space complexity of PU, it is feasible to be used as the postprocessing step in GAUNet-PU. According to the estimated ambiguity gradients under different guidance strategies, we can calculate the corresponding number of residues and select the best guidance strategy for subsequent absolute phase reconstruction. The overall flow of GAUNet-PU is shown in Fig. 8.

IV. PERFORMANCE ANALYSIS

In this section, we first utilize ablation experiments to clarify the effectiveness of our three contributions in dealing with the problem of class imbalance. Next, we explore the performance improvements of GAUNet with different guiding strategies by comparing the number of residues under various noise levels. Then, the performance evaluation of GAUNet-PU methodology is implemented through three experiments, which are performed on simulated topographic interferogram, real topographic interferogram exhibiting rugged terrain, and real differential interferogram measuring surface deformation of Hawaii earthquake, respectively. In all of the three experiments, the performance of GAUNet-PU is compared with several representative traditional 2-D PU methods.

A. Performance Improvements Brought by Three Contributions

As aforementioned, to deal with the negative impact of class imbalance, GAU module is added to the U-Net to build GAUNet architecture, and the quality map is not only fed into the network
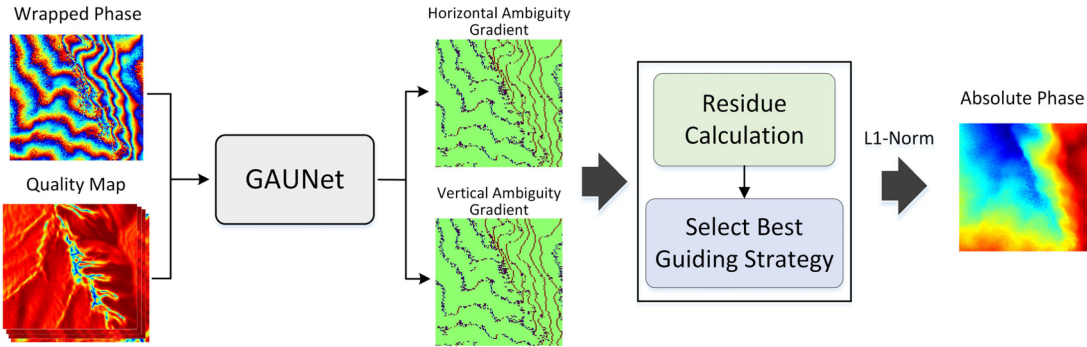
Fig. 8.   Overall flow of GAUNet-PU. Best guiding strategy is selected based on the number of residues, and L1-Norm is applied to restore the absolute phase.

TABLE III
ABLATION OF GAUNET

Contribution			MIoU	
A	B	C	Horizontal	Vertical
			0.9415	0.9406
✓			0.9493	0.9472
	✓		0.9470	0.9463
✓	✓		0.9579	0.9567
	✓	✓	0.9493	0.9469
✓	✓	✓	**0.9595**	**0.9572**

together with its corresponding wrapped phase, but also used for the formulation of the loss function. By adopting the methods of ablation experiment, the performance improvements brought by these three contributions are clarified in this part.

To analyze the performance of GAUNet architecture, MIoU is utilized as the evaluation metric [47]. IoU, also known as Jaccard index, refers to the intersection and the union ratio between the predicted segmentation and the ground truth, has been used as a standard measuring metric in multifarious semantic segmentation tasks. By calculating the IoU of each semantic class, and then calculating the average over classes, we can get the MIoU, which is defined as

$$\text{MIoU} = \frac{1}{k+1} \sum_{0}^{k} \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \qquad (18)$$

where TP, FP, and FN represent the number of true positive, false positive, and false negative samples, respectively, and k takes the value of 2 in our task.

The ablation experiments of GAUNet are implemented on a test dataset consisting of 2439 samples of 128×128 pixels simulated by acquisition parameters from three satellites (1300 with noise added by constant coherence and 1139 with noise added by real derived coherence), and the results are illustrated in Table III. Contribution A represents the addition of GAU module, contribution B represents the input of quality map (take coherence for example), and contribution C represents the participation of quality map in the loss function. Through analysis, it can be found that all three contributions have positive effects on ambiguity gradient estimation, whereas contributions A and B can yield greater improvements.

B. Strategies Analysis Under Various Noise Conditions

To verify and compare the performances of GAUNet under the guidance of different quality maps, in this part we use simulated data under various noise levels (noises added by fixed coherence and coherence derived from the real processed interferograms, respectively) to test five trained GAUNet with different strategies.

To begin with, GAUNet is tested in synthetic data with the noise added by constant coherence from 0.4 to 1 with 0.05 as interval; therefore, the robustness of five implemented strategies for varying noise is demonstrated. For each noise level, 1000 samples of 128×128 pixels simulated by acquisition parameters from three satellites are applied. Note that the samples used for test are independent from the previously used training sets. In this experiment, two evaluation metrics, i.e., the MIoU and the kappa coefficient, are used to evaluate the performances of GAUNet with different guidance strategies [47], [48]. The MIoU has been covered previously, here we only provide details regarding the kappa coefficient. Kappa coefficient is a metric for consistency test, and can also be used to measure the effect of classification. Kappa coefficient is derived from the confusion matrix, which is a square matrix with the size of the classification number, and the element at position (i, j) in confusion matrix represents the number of pixels whose ground truth is i and predicted to be j. Kappa coefficient can be calculated as follows:

$$\text{kappa} = \frac{p_0 - p_c}{1 - p_c}$$

$$\text{where } p_0 = \frac{\sum_{i=1}^{r} x_{ii}}{N}, \ p_c = \frac{\sum_{i=1}^{r} (x_{i+} \times x_{+i})}{N^2} \qquad (19)$$

where r represents the class number of ambiguity gradient, i.e., $r = 3$ in GAUNet. x_{ii} represents the number of correctly predicted pixels of each class, x_{i+} and x_{+i} represent the number of pixels belonging to the ith class in prediction map and ground truth map, respectively, and N is the total number of pixels in the wrapped phase images. Actually, p_0 is the metric of accuracy, which is the most basic and simplest evaluation metric but is not so effective in dealing with the imbalanced classes. Kappa coefficient can punish the bias of the model to replace the metric of accuracy in the case of class imbalance.

The trained GAUNet is directly applied, so that the parameters do not change during the test process. The performance evaluation results under various coherence levels with different
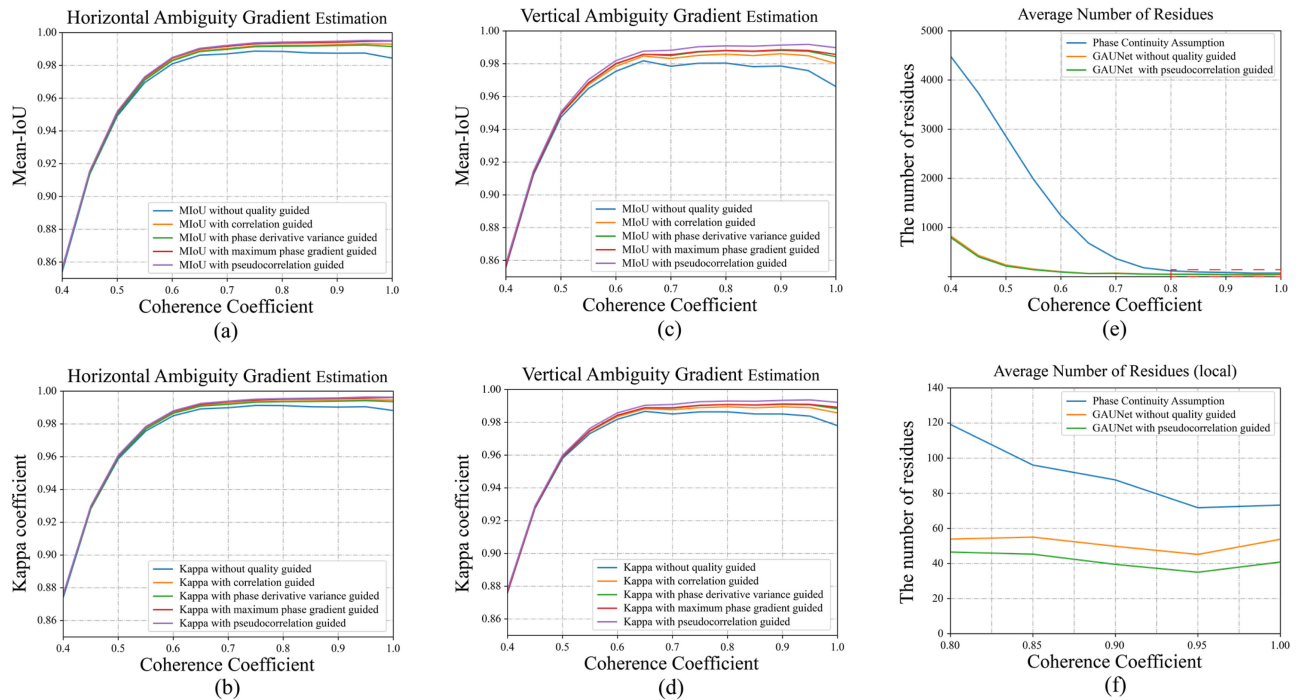
Fig. 9. (a) and (b) exhibit MIoU and kappa coefficient of HG estimation with different guiding strategies under various coherence coefficient levels, respectively. (c) and (d) exhibit those of VG estimation, respectively. (e) shows the comparison of the average number of residues obtained by different methods under various coherence coefficient levels, and the local enlarged drawing of red rectangle dashed box in (e) are shown as (f).

guidance strategies are shown in Fig. 9. Through comparison, it can be found that different quality maps have different degrees of influence on the ambiguity gradient estimation process. In the situation of low coherence, the help brought by the quality map is not obvious. This is expected since the entire input image is very noisy, and almost all of the pixels need to be paid attention to. With the increase of the coherence coefficient, quality maps begin to show their advantages. In general, GAUNet guided by pseudocorrelation shows the best performance, and GAUNet guided by maximum phase gradient seems to perform better than the ones guided by correlation and by phase derivative variance here. Because constant images of coherence are used, the correlation map cannot provide local information for each individual sample actually, can only tell the GAUNet how much attention it needs to provide for this sample at the scale of the whole data sets. The phase derivative variance map resembles the correlation map by comparing Figs. 1(b) and 3(a), yielding relatively similar performance. As mentioned earlier, the pseudo-correlation map and the maximum phase gradient map tend to mark the areas with steep terrain as low quality; thus, they can guide GAUNet to focus on them and achieve better performance, especially in the situation of higher coherence. Fig. 9(e) reveals the relationship between the coherence coefficient and the average number of residues obtained by phase continuity assumption and GAUNet with two diverse guiding strategies. As expected, all three methods perform well in the application scenario of high coherence coefficient, but it is very easy to find that the ability of GAUNet in dealing with low coherence coefficient is far stronger than that of phase continuity assumption. In other words, GAUNet has stronger noise robustness. For the sake of clarity, the local enlargement of Fig. 9(e) with

the coherence coefficients ranging from 0.8 to 1.0 is shown in Fig. 9(f). Drawing support from it, the performance comparison of the three methods in the case of high coherence coefficient can be observed more intuitively. It is worth pointing out that the result presented in Fig. 9 is a universal result based on a large number of test samples, and the best strategy in specific application scenarios may have subtle differences. However, the overall conclusion will never change, that is, the ability of GAUNet to estimate the ambiguity gradient is better than that of phase continuity assumption, and the quality map can provide effective guidance for this process.

GAUNet is then tested by synthetic InSAR data with noise added by coherence extracted from the real processed interferograms. In this way, the simulated data are more consistent with the real cases, ergo the corresponding test results are more convincing. Using the system parameters together with the coherence information obtained from the real single-pass interferograms of the three satellites, 1193 samples of 128×128 pixels are simulated for this purpose. The real coherence has a wider variation range, and the value of many regions is even lower than 0.2, which greatly increases the difficulty of ambiguity gradient estimation. Fig. 10(a) and (b) illustrate the testing results of GAUNet using different guiding strategies on simulated data with real coherence. These results indicate that they do have somewhat of a different behavior when comparing with the results shown in Fig. 9(a)–(d). One of the most obvious differences is that the performance improvement of GAUNet under the guidance of correlation is greatly increased, and the optimum guiding strategy changes from pseudocorrelation to correlation. This is because coherence is no longer a constant over the whole correlation map. Instead, as the basis of noise addition, it can accurately reflect the quality of each pixel and
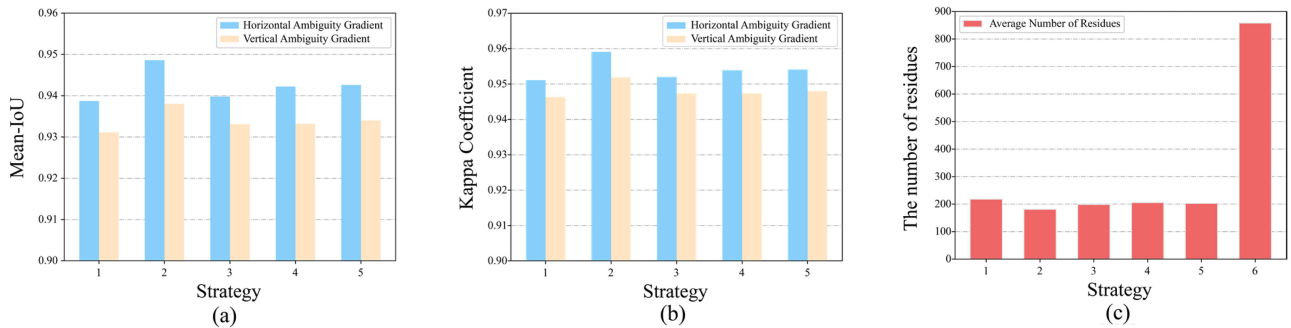
Fig. 10. Strategies 1–5 represent GAUNet without quality guided, GAUNet with correlation guided, GAUNet with phase derivative variance guided, GAUNet with maximum phase gradient guided, and GAUNet with pseudocorrelation guided, respectively. Strategy 6 represents phase continuity assumption. (a) and (b) demonstrate MIoU and kappa coefficient of horizontal and VG estimations under real coherence coefficient, respectively. (c) shows the comparison of the average number of residues obtained by phase continuity assumption and five GAUNet-based methods under real coherence coefficient.

preferably help GAUNet to deal with the presence of high-density nonzero ambiguity gradients. The other three quality maps can still provide certain guidance for the estimation of ambiguity gradient under the condition of true coherence, and the promotion of performance is basically consistent with the case of using constant coherence. Although the pseudocorrelation and the maximum phase gradient tend to divide the steep terrain areas into low-quality areas, which is conducive to the gradient estimation, they also inevitably make wrong judgments on the quality of some pixels. Hence one can see that, same as the traditional quality-guided methods, an ideal quality map is very important for our method, poor quality maps may have a negative impact on the results. Fig. 10(c) shows the average number of residues obtained by phase continuity assumption and GAUNet with five different guiding strategies, and we can still clearly see the incomparable superiority of GAUNet in estimation of the ambiguity gradient. As mentioned earlier, the results obtained in each individual application scenario may differ slightly from the general results depicted in Fig. 10. However, we believe that the real situation should be quite similar to the results shown in Fig. 10. It is acknowledged that coherence is an intermediate product of InSAR data processing, which is very convenient to obtain. Therefore, in practical application, we recommend using coherence to guide GAUNet for InSAR PU for the sake of simplicity. But in order to ensure the preciseness of the subsequent experiments, we will first determine the optimal guiding strategy for each example according to the number of residues, then compare the best result with other state-of-art methods to prove the effectiveness of the proposed method.

C. Performance Evaluation With Simulated Datasets

In this section, we compare the performance of GAUNet-PU with serval typical PU methods, including QGPU [23], MCF [25], and PUMA [27], on the synthetic interferograms with constant coherence and real coherence, respectively. The code of QGPU used in this article comes from [3], MCF method is implemented in GAMMA software [49], and the code of PUMA is supplied by its designers. In this and the subsequent experiments, QGPU is guided by phase derivative variance, the triangulation mode of MCF (Delaunay triangulation mode or

filled triangular mesh) and the parameters of PUMA are determined by the best performance. The first row in Fig. 11 illustrates a series of wrapped phase images with mean coherence from 0.4 to 1, and the corresponding residue images obtained by GAUNet and phase continuity assumption are shown in rows 2 and 3. It is noteworthy that the residue images obtained by GAUNet are determined based on different quality maps, guiding strategy minimizing the number of residues is used to guide the follow-up unwrapping process for each coherence level. The wrapped phase images are simulated with parameters from ALOS-2, with 128×128 pixels in size. We can find that with the decrease of the mean coherence, the phase fringes become blurred; it is thus very challenging for accurate estimation of the ambiguity gradient. The pattern of phase in the observed area is not exactly complicated except for three NW-SE oriented phase jump; therefore, the residues estimated by the two methods are mainly distributed around the jump regions. By comparison, we can clearly see that the number of residues obtained by GAUNet is much lesser than that obtained by phase continuity assumption for each coherence level. The specific statistic is shown in Table IV. Through statistics analysis, we learn that the total number of residues obtained by GAUNet is only 14.1% of that obtained by phase continuity assumption.

Based on the wrapped phase images presented in Fig. 11 and the estimated ambiguity gradient, we use L1-norm model to attain the final GAUNet-PU results, as shown in Fig. 12. For comparison, we also calculate the PU results using QGPU, MCF, and PUMA. Although using quality maps as weighted factors may improve the accuracy of GAUNet-PU, for the sake of fairness, we unweight all PU methods used in this and the follow-up experiments. In order to ensure the facticity of the mean coherence, no filtering operation is performed before PU. The same reference point is chosen to ensure that all PU methods are in the same scale. Through visual comparison, it can be found that except for QGPU, the other three methods can basically obtain robust PU results when the mean coherence is greater than 0.6. As a member of the path-following PU family, QGPU shows the general problem of these methods in this experiment, that is, the islanding effect. When the mean coherence is less than 0.6, the results obtained by MCF and PUMA begin to show obvious unwrapping errors. Especially when the mean coherence is 0.4, the results obtained by all three traditional methods seem not
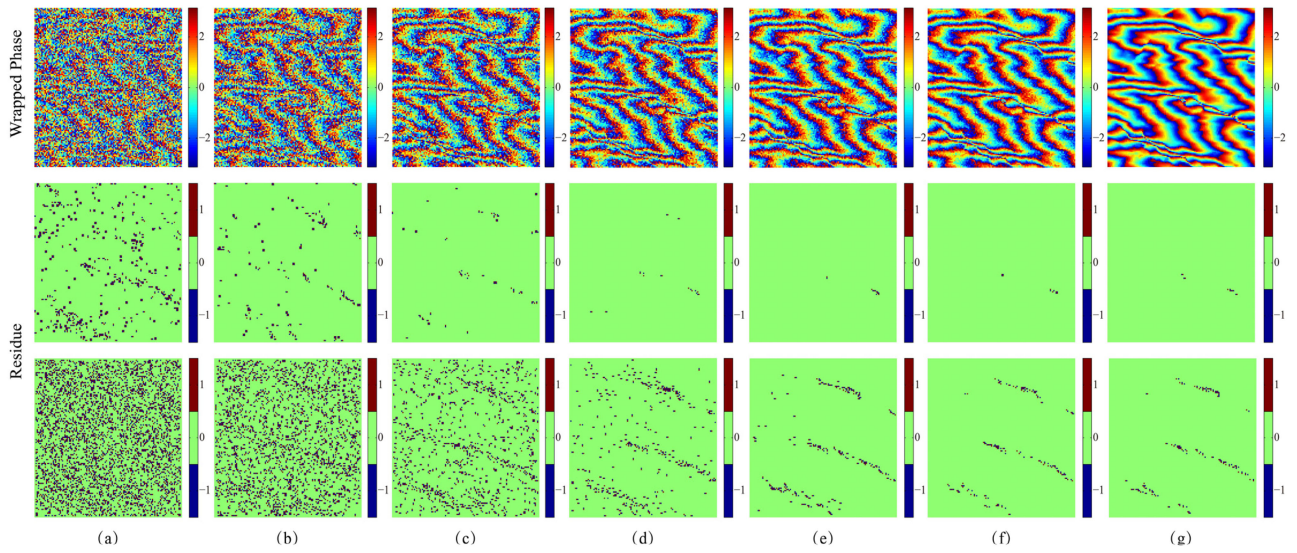
Fig. 11.   Row 1: wrapped phase images with constant mean coherence (a) 0.4, (b) 0.5, (c) 0.6, (d) 0.7, (e) 0.8, (f) 0.9, and (g) 1.0. Row 2: corresponding residue images obtained by GAUNet. Row 3: corresponding residue images obtained by phase continuity assumption.

TABLE IV
NUMBER OF RESIDUES

Mean Coherence	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Total
GAUNet	976	321	95	22	12	12	12	1450
Phase Continuity Assumption	4472	3090	1532	596	273	153	143	10259
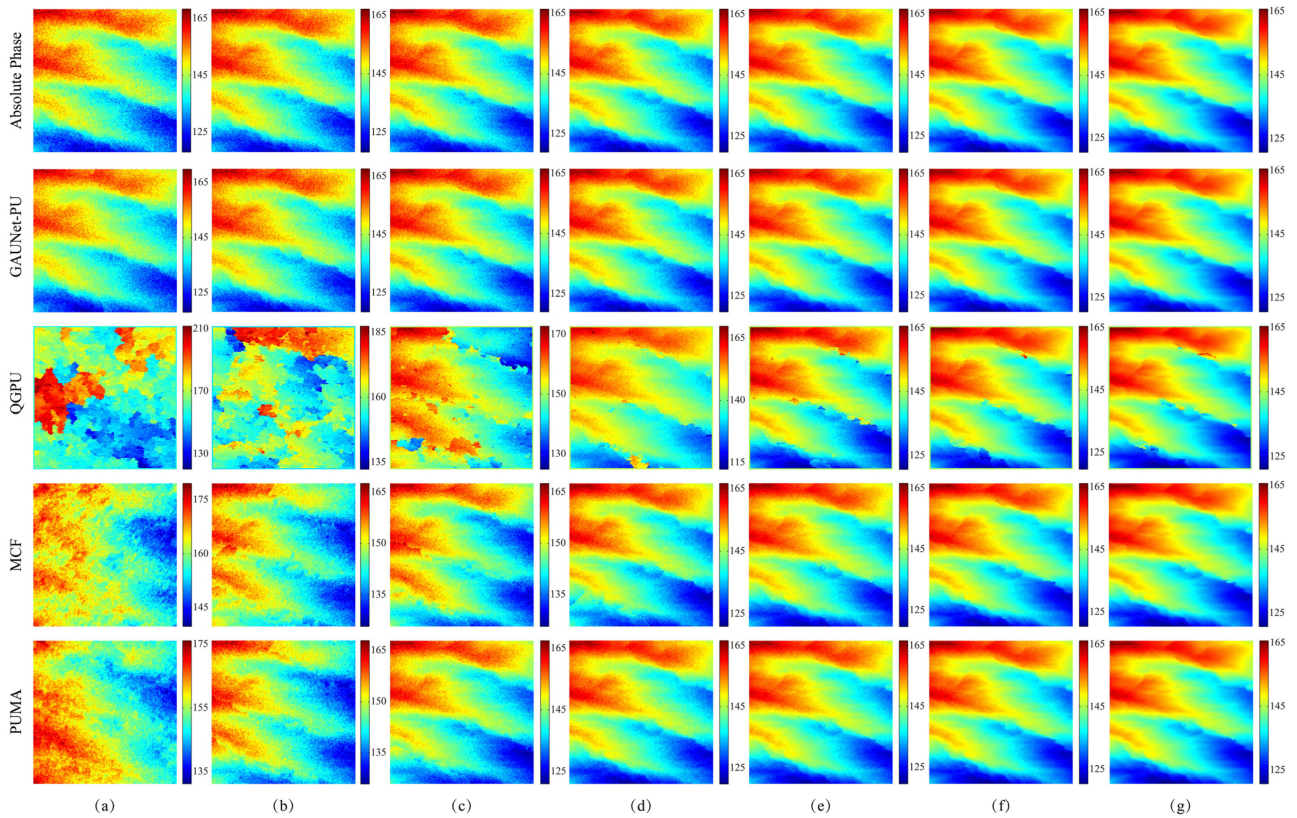
Fig. 12.   Row 1: corresponding true absolute phase of row 1 in Fig. 11. Row 2: absolute phase obtained by GAUNet. Row 3: absolute phase obtained by QGPU. Row 4: absolute phase obtained by MCF. Row 5: absolute phase obtained by PUMA.

TABLE V
RMSE BETWEEN TRUE PHASE AND ESTIMATED PHASE OF FOUR PU METHODS

Mean Coherence	0.4	0.5	0.6	0.7	0.8	0.9	1.0	Overall
GAUNet-PU	1.728	0.856	0.347	0.104	0.070	0.042	0.049	0.457
QGPU	25.789	18.811	11.767	2.332	1.091	0.928	0.954	8.810
MCF	19.278	6.943	5.234	1.795	0.245	0.264	0.250	4.858
PUMA	15.994	5.928	1.915	0.098	0.070	0.069	0.085	3.451
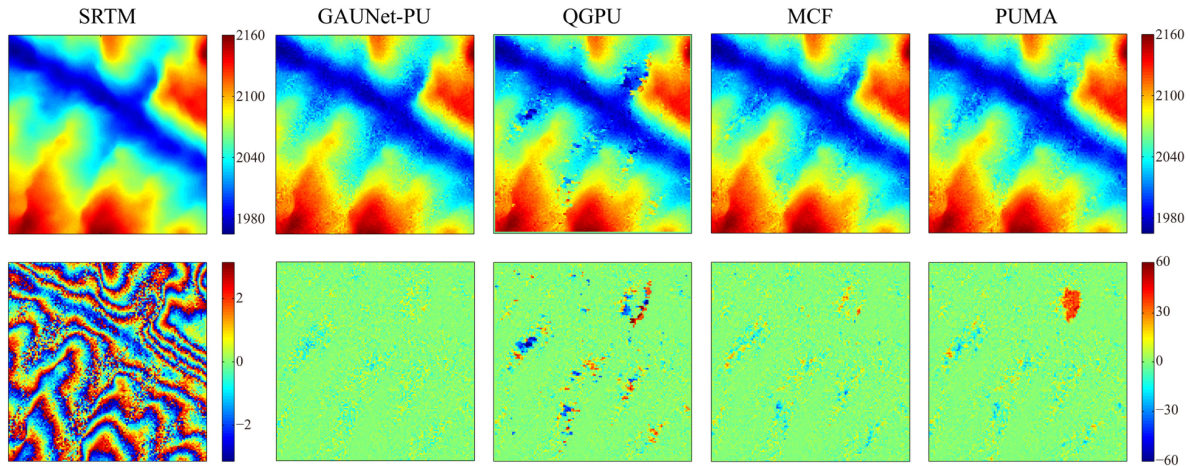
Fig. 13. Column 1: SRTM DEM and synthetic wrapped phase with real coherence. Columns 2–5: estimated DEM and corresponding error images between the estimated DEM and the reference SRTM DEM obtained by GAUNet-PU, QGPU, MCF, and PUMA.

credible at all. This is because the phase continuity assumption almost fails when the wrapped phase is too noisy, and it is difficult for the traditional methods to deal with the ambiguity gradient with too many errors effectively. In contrast, GAUNet can estimate the ambiguity gradient accurately even when the coherence is low; therefore, our method exhibits strong noise robustness and is very applicable in situations with high noise levels. For quantitative comparison, we calculate the root mean square error (RMSE) between the predicted unwrapped phase and the true phase, as illustrated in Table V. With respect to all tested methods, GAUNet-PU shows the best performance, getting an overall lowest mean RMSE. It can be found that even if the same optimization model is used, since the estimation of ambiguity gradient is more accurate, GAUNet-PU can perform better than MCF at each coherence level. When the mean coherence is high (≥ 0.7), GAUNet-PU and PUMA both perform very well, but in the situations of low mean coherence, GAUNet-PU is significantly better.

Using the parameters obtained from the TerraSAR-X/TanDEM-X interferogram, a wrapped phase sample with 128×128 pixels in size is simulated, to evaluate the performance of GAUNet-PU on cases with real coherence. The first column in Fig. 13 shows the reference SRTM DEM and the synthetic wrapped phase. The DEMs estimated by GAUNet-PU, QGPU, MCF, and PUMA are illustrated in row 1, columns 2 to 5 of Fig. 13. To reveal the discrepancy more clearly, the error images between the estimated DEM and the reference SRTM DEM are calculated, as shown in row 2, columns 2 to 5 of Fig. 13. It is worth mentioning that the colorbars of the four error images are in the same range. Through observation, it can be found that

TABLE VI
RMSE BETWEEN THE ESTIMATED DEM AND THE REFERENCE DEM

PU Methods	RMSE
GAUNet-PU	4.336
QGPU	7.819
MCF	4.744
PUMA	5.818

the error image obtained by GAUNet is the smoothest overall. In the high-quality regions, all the tested methods seem to achieve reliable PU results. However, in the regions with low coherence, due to the influence of noise, the estimated DEMs show certain errors when compared with the reference SRTM DEM. In general, the results obtained by GAUNet-PU and MCF are more robust. Since the ambiguity gradient estimated by GAUNet is more correct than that estimated by phase continuity assumption, GAUNet-PU can avoid many local unwrapping errors when compared with MCF although the same optimization model is applied. Besides the obvious unwrapping errors existing in the upper right corner of the image, PUMA has a similar performance with MCF. QGPU produces extensive unwrapping errors in all the low-quality areas. This can be ascribed to the fact that the error information containing in ambiguity gradient estimated by phase continuity assumption results in too many residues, thus QGPU cannot choose an effective integration path.

To quantitatively evaluate the performance of GAUNet-PU, we calculate the RMSEs between the estimated DEMs and the reference DEM, as illustrated in Table VI. Seen from the numerical point of view, the RMSEs of the four PU methods are
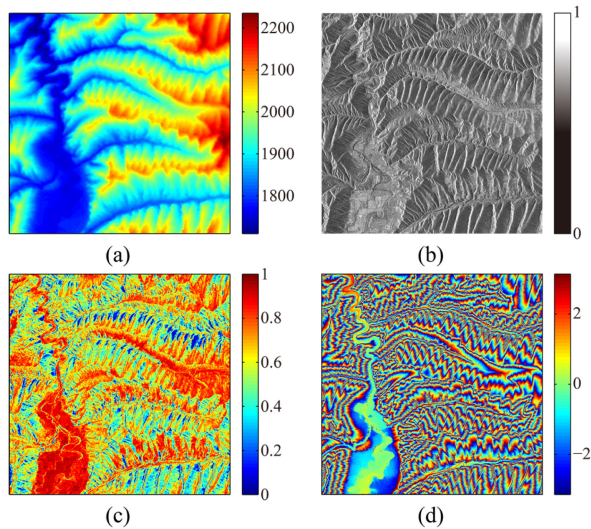
Fig. 14. (a) Reference SRTM DEM. (b) SAR intensity image. (c) Coherence map. (d) Flattened real interferogram.
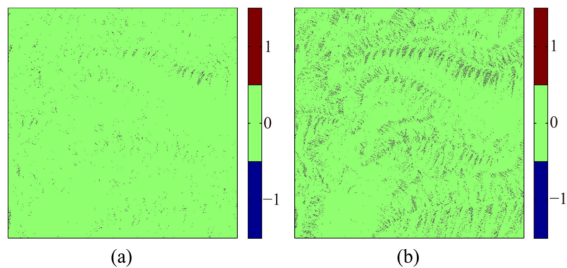

Fig. 15. (a) Residue image estimated by GAUNet. (b) Residue image estimated by phase continuity assumption.

not particularly different due to the high quality of most areas on the test image, but GAUNet-PU still achieves the lowest RMSE.

D. Experiments With Topographic Interferogram of Dingxi, China

To test the practical applicability of the proposed method, the experiment with the real topographic interferogram is implemented in this part. A TerraSAR-X/TanDEM-X interferogram covering Dingxi City, Gansu Province, China, acquired on August 22, 2013, is applied in this experiment. Fig. 14 shows the corresponding reference SRTM DEM, SAR intensity image, coherence map, and the flattened real interferogram of the study area. After 4:4 multilook processing, the size of the interferogram is 1000×1000 and its spatial resolution is approximately 12 m because the satellite adopts Stripmap Mode. As shown in Fig. 14(c), due to the special side-looking imaging mechanism of SAR system, phase aliasing often occurs in areas with big topographic relief, resulting in local low coherence. Under such circumstances, the ambiguity gradient estimated by phase continuity assumption is not credible. Fig. 15(a) and (b) are the residue images of Fig. 14(d) obtained by GAUNet and phase continuity assumption, respectively. By visual contrast, we can clearly see that the residues obtained by GAUNet are significantly lower than that of phase continuity assumption.

TABLE VII
RMSE BETWEEN THE ESTIMATED DEM AND THE REFERENCE DEM

PU Methods	RMSE
GAUNet-PU	20.105
QGPU	36.443
MCF	20.508
PUMA	20.956

Through statistical analysis, the number of residues obtained by GAUNet (6515 residues) is only 16.25% of that obtained by phase continuity assumption (40 078 residues). This proves the effectiveness of GAUNet in real scenarios.

The first row in Fig. 16 illustrates the DEMs estimated by GAUNet-PU, QGPU, MCF, and PUMA, respectively. Except for the obvious difference of QGPU, it is difficult for us to distinguish the results of the other three methods with naked eyes. For a more intuitive and clear comparison, we compute the error images between the estimated DEMs and the reference SRTM DEM, as shown in the second row of Fig. 16. On account of a large number of mistakes in ambiguity gradient estimated by phase continuity assumption, QGPU can not choose an appropriate integration path, thus exhibit quite obvious unwrapping errors. It can be found that the results of GAUNet-PU, MCF, and PUMA are still relatively similar. However, compared with the other two methods, GAUNet-PU can avoid some local unwrapping errors in certain low coherence regions. By calculating the RMSE between the estimated DEM and the reference SRTM DEM (see Table VII), the PU accuracy can be compared quantitatively.

Considering the similarity of the DEMs and RMSEs obtained by GAUNet-PU, MCF, and PUMA, it is inferred that the dominant error source is the inaccuracy of the reference SRTM DEM. Due to the difference of noise robustness, the RMSEs of the four methods present certain discrepancies. Even none of the tested methods can completely avoid the local unwrapping errors, GAUNet performs better than all the traditional PU methods, resulting in the lowest RMSE. Due to the constraints of data, we conduct the test in Dingxi, where the terrain is not particularly undulating. However, we believe that GAUNet-PU will show greater advantages in the case of more undulating terrain or lower mean coherence.

E. Experiments With Differential Interferogram of 2018 Hawaii Earthquake, USA

Traditional 2-D PU methods usually fail to unwrap the severely decorrelated areas in the interferograms, which leads to the inability to obtain or underestimate the coseismic deformation in the near-field of the earthquake. On 4 May 2018, a thrust-faulting M_W 6.9 earthquake occurred on Hawaii Island, USA [50]. This earthquake is a volcanic earthquake caused by the volcanic activity of the Kilauea Volcano Group on Hawaii Island. Fig. 17 demonstrates the distribution of earthquakes with magnitudes greater than 3.0 that occurred on the island of Hawaii from May 1 to 10, 2018.

To test the ability of GAUNet in dealing with the differential interferogram, two ALOS-2 data obtained on January 20, 2018 and May 26, 2018 (corresponding to the region of the red
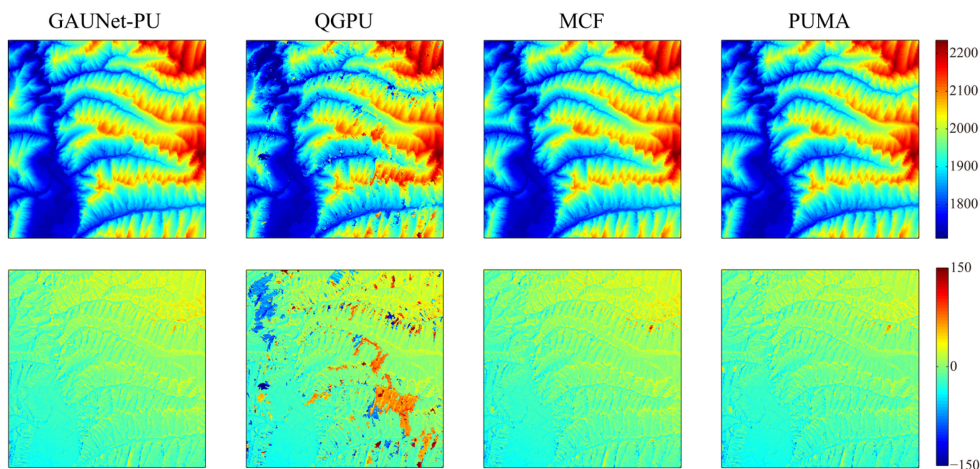
Fig. 16. Row 1: corresponding DEM of Fig. 13(d) obtained by GAUNet-PU, QGPU, MCF, and PUMA, respectively. Row 2: error images between the estimated DEM and Fig. 13(a).
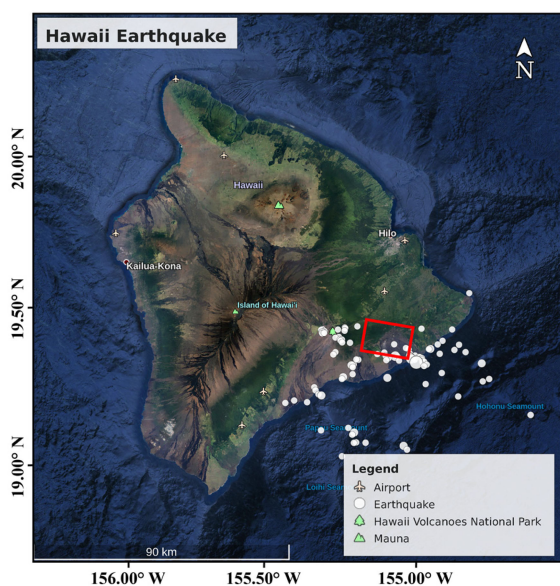

Fig. 17. Earthquakes with magnitudes greater than 3.0 that occurred on the island of Hawaii from May 1 to 10, 2018.
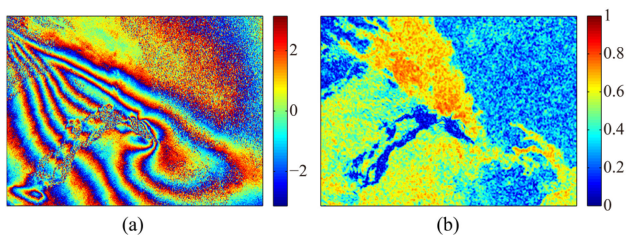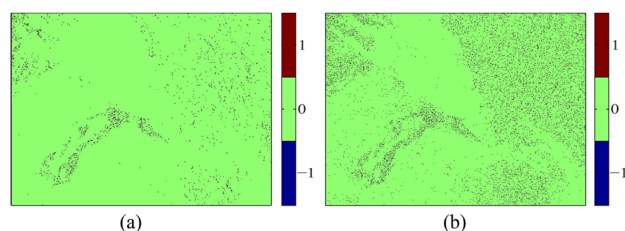

Fig. 19. (a) Residue image estimated by GAUNet. (b) Residue image estimated by phase continuity assumption.

Fig. 20. (a) Absolute phase recovered by GAUNet-PU. (b) Absolute phase recovered by MCF.

Fig. 18. (a) Deformation interferogram of 2018 Hawaii earthquake, USA. (b) Corresponding coherence image of (a).

rectangle in Fig. 17), are implemented to generate deformation interferogram of the Hawaii earthquake, as shown in Fig. 18(a). After 2:9 multilook processing, the size of the interferogram is 512×696. From the corresponding coherence image shown in Fig. 18(b), this can seem like a challenging task. In particular,

the interferogram covers the East Rift Zone of Hawaii Island, which causes serious decorrelated effects in local areas under the earthquake function. Even multilook processing is applied, it is difficult to obtain correct ambiguity gradient information using phase continuity assumption. Although our training sets do not contain any samples of deformation interferograms, as depicted in Fig. 19, the number of residues obtained by GAUNet (5098 residues) is still significantly less than that of phase continuity assumption (16 510 residues). This indicates the powerful generalization performance of the proposed method. Since the coherence of the fracture region is almost close to zero, neither of the methods can estimate the ambiguity gradient effectively. However, in other regions of the interferogram, the distribution of residues obtained by GAUNet is much sparser than that obtained by phase continuity assumption. The unwrapping phase obtained by GAUNet-PU and MCF are shown in Figs. 20(a) and (b), respectively. The overall trend is relatively consistent, but

there are obvious differences in local areas. For the lacking of real reference, we can not compare the PU results quantitatively. However, since the same optimization model is applied and the ambiguity gradient information estimated by GAUNet is more accurate, we believe that the unwrapping phase obtained by GAUNet-PU will be more in line with the real situation than that obtained by MCF. From Fig. 20(b), we can clearly observe the PU errors existing in the upper left corner and right area of the image where they are characterized by lower coherence. In contrast, GAUNet-PU can get smoother unwrapping phase and avoid many local unwrapping errors, which proves the effectiveness of the proposed method once again.

V. Conclusion

In this article, we proposed a novel quality-guided GAUNet-PU algorithm for 2-D PU in InSAR interferogram. PU is regarded as a two-stage process. The first stage is the estimation of ambiguity gradient, whereas the second stage is the restoration of absolute phase by utilizing the gradients obtained in stage 1. We have innovated mainly for the first stage, committing to deal with the problem of class imbalance in ambiguity gradient estimation and thus improve the accuracy of the final unwrapping results. Specifically, GAU module is added to the classic semantic segmentation network—U-Net, and the quality map is not only input into the network as additional information, but also used to modify the loss function. Undergo the modification, the network can focus on the areas with denser nonzero gradient distribution. After training on the generated multisource dataset, GAUNet can estimate the ambiguity gradient more conveniently and accurately when comparing with the phase continuity assumption. Experiments show that the quality map can provide certain guidance for GAUNet, and GAUNet-PU demonstrates state-of-the-art performance on both synthetic and real data, including topographic and deformation interferograms.

Because there are no open source code and parameter settings, it is quite difficult to reproduce the existing DL research [17], [18]. Despite all this, it will be necessary to compare the performance with other DL-based PU methods in future studies. To ensure fairness, all methods tested are unweighted in this article. But actually, the quality map can play its third role in GAUNet-PU, that is, the weighted coefficient of PU. Moreover, the second stage of the proposed method is based on the original version of MCF. In fact, many academics have made further improvements on MCF, which can be added to our method. In addition, according to Fig. 19(a), we can find that for the case of extremely low coherence, it is challenging to estimate the correct ambiguity gradient even using DL. Therefore, a necessary mask is still required in this circumstance. Being directed against the unique characteristics of the interested signal (e.g., topography or deformation), more reliable quality maps or other InSAR products such as DEM can be explored as guidance in GAUNet to achieve higher PU accuracy in topography mapping or deformation monitoring. In addition, multi-temporal InSAR (MT-InSAR) has received widespread attention in recent years, but parameter estimation based on the wrapped phase is still a critical step. GAUNet-PU can be introduced into the processing of time-series InSAR to aid the parameter estimation in MT-InSAR techniques.

References

[1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geosci. Remote Sens. Mag.,* vol. 1, no. 1, pp. 6–43, Mar. 2013.

[2] F. Xue, X. Lv, F. Dou, and Y. Yun, "A review of time-series interferometric SAR techniques: A tutorial for surface deformation analysis," *IEEE Geosci. Remote Sens. Mag.,* vol. 8, no. 1, pp. 22–42, Mar. 2020.

[3] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. New York, NY, USA: Wiley, 1998.

[4] H. Yu, Y. Lan, Z. Yuan, J. Xu, and H. Lee, "Phase unwrapping in InSAR: A review," *IEEE Geosci. Remote Sens. Mag.,* vol. 7, no. 1, pp. 40–58, Mar. 2019.

[5] K. Itoh, "Analysis of the phase unwrapping algorithm," *Appl. Opt.*, vol. 21, no. 14, pp. 2470–2470, 1982.

[6] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature,* vol. 521, no. 7553, pp. 436–444, 2015.

[8] X. X. Zhu *et al.*, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.,* vol. 5, no. 4, pp. 8–36, Dec. 2017.

[9] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS-J. Photogramm. Remote Sens.,* vol. 152, pp. 166–177, 2019.

[10] X. Zhu *et al.*, "Deep learning meets SAR: Concepts, models, pitfalls, and perspectives," *IEEE Geosci. Remote Sens. Mag.*, to be published, doi: 10.1109/MGRS.2020.3046356.

[11] G. E. Spoorthi, S. Gorthi, and R. K. S. S. Gorthi, "PhaseNet: A deep convolutional neural network for two-dimensional phase unwrapping," *IEEE Signal Process. Lett.,* vol. 26, no. 1, pp. 54–58, Jan. 2019.

[12] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440.

[13] T. Zhang *et al.*, "Rapid and robust two-dimensional phase unwrapping via deep learning," *Opt. Exp.*, vol. 27, no. 16, pp. 23173–23185, 2019.

[14] J. Zhang, X. Tian, J. Shao, H. Luo, and R. Liang, "Phase unwrapping in optical metrology via denoised and convolutional segmentation networks," *Opt. Exp.*, vol. 27, no. 10, pp. 14903–14912, 2019.

[15] K. Wang, Y. Li, Q. Kemao, J. Di, and J. Zhao, "One-step robust deep learning phase unwrapping," *Opt. Exp.*, vol. 27, no. 10, pp. 15100–15115, 2019.

[16] G. E. Spoorthi, R. K. S. S. Gorthi, and S. Gorthi, "PhaseNet 2.0: Phase unwrapping of noisy data based on deep learning approach," *IEEE Trans. Image Process.,* vol. 29, pp. 4862–4872, 2020.

[17] L. Zhou, H. Yu, and Y. Lan, "Deep convolutional neural network-based robust phase gradient estimation for two-dimensional phase unwrapping using SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4653–4665, Jul. 2020.

[18] F. Sica, F. Calvanese, G. Scarpa, and P. Rizzoli, "A CNN-based coherence-driven approach for InSAR phase unwrapping," *IEEE Geosci. Remote Sens. Lett.,* to be published, doi: 10.1109/LGRS.2020.3029565.

[19] L. Zhou, H. Yu, and Y. Lan, "Artificial intelligence in interferometric synthetic aperture radar phase unwrapping: A review," *IEEE Geosci. Remote Sens. Mag.,* vol. 9, no. 2, pp. 10–28, Jun. 2021.

[20] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Sci.*, vol. 23, no. 4, pp. 713–720, 1988.

[21] J. M. Huntley, "Noise-immune phase unwrapping algorithm," *Appl. Opt.*, vol. 28, no. 16, pp. 3268–3270, 1989.

[22] D. J. Bone, "Fourier fringe analysis: The two-dimensional phase unwrapping problem," *Appl. Opt.,* vol. 30, no. 25, pp. 3627–3632, 1991.

[23] T. J. Flynn, "Consistent 2-D phase unwrapping guided by a quality map," in *Proc. Int. Geosci. Remote Sens. Symp.*, 1996, vol. 4, pp. 2057–2059.

[24] D. C. Ghiglia and L. A. Romero, "Minimum Lp-norm two-dimensional phase unwrapping," *J. Opt. Soc. Amer. A,* vol. 13, no. 10, pp. 1999–2013, 1996.

[25] M. Costantini, "A novel phase unwrapping method based on network programming," *IEEE Trans. Geosci. Remote Sens.,* vol. 36, no. 3, pp. 813–821, May 1998.

[26] D. C. Ghiglia and L. A. Romero, "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods," *J. Opt. Soc. Amer. A,* vol. 11, no. 1, pp. 107–117, 1994.

[27] J. M. Bioucas-Dias and G. Valadao, "Phase unwrapping via graph cuts," *IEEE Trans. Image Process.,* vol. 16, no. 3, pp. 698–709, Mar. 2007.

[28] C. W. Chen and H. A. Zebker, "Two-dimensional phase unwrapping with use of statistical models for cost functions in nonlinear optimization," *J. Opt. Soc. Amer. A,* vol. 18, no. 2, pp. 338–351, 2001.

[29] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," 2020, *arXiv:2001.05566.*

[30] T. G. Farr *et al.*, "The shuttle radar topography mission," *Rev. Geophys.*, vol. 45, no. 2, pp.1–33, 2007.

[31] G. Krieger *et al.*, "TanDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Trans. Geosci. Remote Sens.,* vol. 45, no. 11, pp. 3317–3341, Nov. 2007.

[32] E. Rodriguez and J. M. Martin, "Theory and design of interferometric synthetic aperture radars," in *Proc. IEE Radar Signal Process.*, 1992, vol. 139, no. 2, pp. 147–159. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/ip-f-2.1992.0018

[33] T. Lei, R. Wang, Y. Wan, X. Du, H. Meng, and A. K. Nandi, "Medical image segmentation using deep learning: A survey," 2020, *arXiv:2009.13120.*

[34] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2015, pp. 234–241.

[35] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science,* vol. 313, no. 5786, pp. 504–507, 2006.

[36] H. Li, P. Xiong, J. An, and L. Wang, "Pyramid attention network for semantic segmentation," 2018, *arXiv:1805.10180.*

[37] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.

[38] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 3–19.

[39] O. Oktay *et al.*, "Attention U-net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999.*

[40] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Deep Learning in Medical Image Analysis and Multimodal Learning For Clinical Decision Support*. Berlin, Germany: Springer, 2017, pp. 240–248.

[41] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2980–2988.

[42] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis.*, 2016, pp. 565–571.

[43] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: Achievements and challenges," *J. Digit. Imag.*, vol. 32, no. 4, pp. 582–596, 2019.

[44] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980.*

[45] H. Yu and Y. Lan, "Robust two-dimensional phase unwrapping for multibaseline SAR interferograms: A two-stage programming approach," *IEEE Trans. Geosci. Remote Sens.,* vol. 54, no. 9, pp. 5217–5225, 2016.

[46] R. K. Ahujia, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[47] G. Csurka, D. Larlus, F. Perronnin, and F. Meylan, "What is a good evaluation measure for semantic segmentation?," 2013, vol. 27, Art. no. 10.5244.

[48] M. L. McHugh, "Interrater reliability: The Kappa statistic," *Biochemia Medica,* vol. 22, no. 3, pp. 276–282, 2012.

[49] C. Werner, U. Wegmüller, T. Strozzi, and A. Wiesmann, "Gamma SAR and interferometric processing software," Citeseer.

[50] C. Liu, T. Lay, and X. Xiong, "Rupture in the 4 May 2018 MW 6.9 earthquake seaward of the Kilauea east rift zone fissure eruption in Hawaii," *Geophys. Res. Lett.,* vol. 45, no. 18, pp. 9508–9515, 2018.

Hai Wang was born in Qingdao, Shandong, China, in 1997. He received the B.S. degree in geographical information science from the China University of Mining and Technology, Xuzhou, China, in 2019. He is currently working toward the postgraduate degree in photogrammetry and remote sensing with Central South University, Changsha, China.

His research interests include phase unwrapping in interferometric synthetic aperture radar (InSAR) and learning algorithms on InSAR data processing and applications.

Jun Hu (Member, IEEE) received the M.Eng. and Ph.D. degrees in geodesy and surveying engineering from Central South University, Changsha, China, in 2008 and 2013, respectively.

From 2013 to 2014, he was a Postdoctoral Fellow with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong. He is currently a Full Professor with the Department of Geomatics and Remote Sensing, School of Geosciences and Info-Physics, Central South University. He is the author of more than 60 papers in international peer-reviewed journals. His research interests include mapping multidimensional and high-precision deformations under complicated environment and its applications in geophysical fields.

Haiqiang Fu received the bachelor's degree in remote sensing science and technology from Southwest Jiaotong University, Chengdu, China, in 2011, and the master's and Ph.D. degrees in geodesy and survey engineering from Central South University, Changsha, China, in 2014 and 2018, respectively.

He is an Associate Professor with the School of Geosciences and Info-Physics, Central South University. His research interests include polarimetric synthetic aperture radar (SAR) interferometry and its applications for monitoring forest parameters and extracting the underlying topography over forest areas.

Changcheng Wang (Member, IEEE) received the Ph.D. degree in photogrammetric and remote sensing from the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China, in 2008.

He is currently a Full Professor with the School of Geosciences and Info-Physics, Central South University, Changsha, China. His research interests include synthetic aperture radar (SAR), interferometric SAR (InSAR), and polarimetric InSAR methods and applications for geosciences.

Zhenhai Wang received an undergraduate degree in remote sensing science and technology in 2018 from Central South University, Changsha, China, where he is currently working toward the doctoral degree in surveying and mapping science and technology.

After studying for a master's degree for one year, he started his doctoral studies, so he doesn't have a master's degree. His current research interests include ground-based synthetic aperture radar error correction, and combined space-borne and ground-based radar monitoring of surface deformation.