InSAR Phase Unwrapping by Deep Learning Based on Gradient Information Fusion

Liutong Li, Hong Zhang, Yixian Tang, Chao Wang, and Feng Gu

***Abstract*— Phase unwrapping (PhU) is an important step in interferometric synthetic aperture radar (InSAR) technology. At present, difficulties are encountered when using deep learning to solve the PhU problem because the fringe density of the actual interferogram varies, resulting in an imbalanced class of semantic segmentation. Deep learning cannot completely use gradient information, and it is difficult to address a large number of residues. In this letter, a PhU semantic segmentation model based on gradient information fusion and improved PhaseNet network is proposed to solve the problem of imbalanced classification and error propagation. 21 613 pairs of phase samples are constructed by using simulated and real Sentinel-1 InSAR Data. The experimental results show that the average classification accuracy of the method can reach 97%, and the mean square error is only 0.97. The average processing speed of 256 × 256 slices is only 0.5 s. Compared with the traditional methods and other deep learning methods, this method solves the problem of classification imbalance, and the use of fusion gradient information improves the efficiency of the algorithm as well as reduces the burden of network classification and the error propagation, showing increased robustness in the case of many residues and high fringe density.**

***Index Terms*— Deep learning, interferometric synthetic aperture radar (InSAR), phase unwrapping (PhU), semantic segmentation.**

I. Introduction

PHASE unwrapping (PhU) is a key step of interferometric synthetic aperture radar (InSAR) technology [1]. Under different terrain conditions, the density of fringes in the interferogram varies significantly, and numerous phase discontinuities are noted. This is the main problem faced by PhU.

Traditional PhU methods include integration path algorithms based on residues and quality maps [2], [3]. The main idea is to suppress PhU errors in low-quality regions from propagating along the integration path. Methods based on optimization theory with the goal of minimizing the phase gradient distance also exist, that is, the minimum Lp-norm (MLPN) algorithm [4], the minimum-cost Flow (MCF) algorithm [5], the Statistical-cost Network-flow PhU (SNAPHU) algorithm [6], and region recognition and expansion by combining optical images and coherence information [7]. The increase in robustness of traditional methods is usually accompanied by a decrease in computational efficiency, and it is difficult to deal with phase discontinuities under complex terrain conditions.

Manuscript received June 17, 2021; revised September 30, 2021; accepted November 4, 2021. Date of publication November 12, 2021; date of current version January 7, 2022. This work was supported by the National Natural Science Foundation of China under Grant 41930110. *(Corresponding author: Hong Zhang.)*

The authors are with the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and also with the College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: zhanghong@radi.ac.cn).

Digital Object Identifier 10.1109/LGRS.2021.3127318

In recent years, the deep learning methods have been introduced into PhU. Through semantic segmentation, the pixels belonging to the same period in the wrapped phase image are classified into the same category one by one, so as to solve PhU problems. Such methods represented by PhaseNet based on SegNet modules [8], CNNs with ResNet modules [9], [10], DeepLabV3+ [11], UNet [12], [13], or other models to achieve semantic segmentation. They often carry out experiments on simulated data and predict the period of the wrapped phase directly without considering the characteristics of actual InSAR data, so it is difficult to solve the classification imbalance caused by real complex terrain. Gradient information is ignored and it is difficult to deal with a large number of residues. The current improved methods include using the unwrapped phases as labels [13]–[15], learning gradient information and backpropagating it [16]–[18], integrating residues into the loss function [19], and using two networks combined with the coherence map to predict the gradient information in two directions for phase reconstruction [20].

However, these deep learning methods still exhibit the following shortcomings.

1) The gradient information of two directions is not merged in the learning network, and two networks are needed to predict the gradients in two directions, which increases the training cost and reduces the computational efficiency and will lead to gradient error propagation.
2) Networks, such as PhaseNet, are relatively shallow, and it is difficult to fit phase data with many residues.

To solve the above problems, this letter proposes a semantic segmentation model of PhU based on azimuth and range gradient information fusion and PhaseNet. The gradient information of range direction and azimuth direction is fused to form a nine-class classification problem, and PhU is performed through the corresponding phase reconstruction algorithm. Because it is a constant nine-class problem for any data, the classification imbalance problem is better solved. By increasing the depth of the PhaseNet network and adding residual modules to better maintain the phase structure characteristics, the gradient information in different directions can be predicted through one network, which reduces the training cost, improves the classification efficiency, and reduces the propagation of gradient error.

This letter is arranged as follows: The second part introduces the proposed method. The experimental results are presented and discussed in Section III. Finally, conclusions are presented in Section IV.

II. Proposed Method

PhU restores the period of the wrapped phase, as shown in the following equation:

$$\Phi(s) = \varphi(s) + 2k(s)\pi, \quad \varphi(s) = W(\Phi(s)) \\ \varphi(s) \in [-\pi, \pi], \quad k(s) \in \text{interger} \quad (1)$$

TABLE I
FUSION RULES OF GRADIENT INFORMATION

$\nabla_a(s_a)$	$\nabla_r(s_r)$	Fusion Value
0	0	0
0	1	1
0	-1	2
1	0	3
1	1	4
1	-1	5
-1	0	6
-1	1	7
-1	-1	8

where $\varphi(s)$ is the wrapped phase of the sth pixel, $\Phi(s)$ is the absolute phase, and $k(s)$ is the wrap count. W is the operation of dividing by 2π and taking the modulo. After our unwrapping algorithm reverses the wrap count, which is k, the unwrapping phase can be obtained according to (1).

A. Semantic Segmentation Process of PhU

The number of cycles of the wrapped phase is difficult to determine and varies greatly in different regions. Given that the period boundary is determined by the gradient, learning gradient information can better extract space distribution features of phases. However, if two networks are used to predict the information of azimuth and range gradient, respectively, it will increase the training cost and lower the prediction efficiency, and the accumulation and propagation of phase errors in two directions are inevitable. In this letter, the range and azimuth gradient information is fused. There are three gradient values in each direction, and nine values are generated after the fusion. Therefore, only one nine-class network is used, which improves computational efficiency and reduces training costs and error accumulation. Gradient distribution images in different directions are obtained according to the following equations:

$$\nabla_a(s_a) = \begin{cases} C(s_a+1) - C(s_a), & (0 \le s_a \le H-1) \\ C(s_a) - C(s_a-1), & (s_a = H) \end{cases} \tag{2}$$

$$\nabla_r(s_r) = \begin{cases} C(s_r+1) - C(s_r), & (0 \le s_r \le W-1) \\ C(s_r) - C(s_r-1), & (s_r = W) \end{cases} \tag{3}$$

where $\nabla_a(s_a)$ represents the gradient value of the azimuth direction to the s_a pixel, and $\nabla_r(s_r)$ represents the gradient value of the range direction to the s_r pixel. C represents the wrap count, and H and W are the maximum pixel positions in two directions. Assuming the continuity of the phase, the value range of $\nabla_a(s_a)$ and $\nabla_r(s_r)$ is $\{0, \pm 1\}$. Information fusion produces nine values, and the network is transformed into a nine-class classification problem. The rules of fusion are shown in Table I. There will be rare multiple cycles jumps and the network will produce some errors. However, this situation is rare, and the redundant equations from two gradient directions can weaken its influence on the global unwrapping effect to a certain extent.

The processing flow is shown in Fig. 1. Training set is generated through wrapping the absolute phase, gradient calculation, and gradient fusion. Taking the wrapped phase image as input, the fusion gradient information is predicted, and the wrap count map is inverted using the phase reconstruction algorithm and then combined with (1) to complete the PhU.

TABLE II
NETWORK STRUCTURE CONFIGURATION INFORMATION FOR PhU

Layer	Kernel	Size	Output
Conv_1+Relu	128	5×5	$256 \times 256 \times 128$
$\text{Conv}_{2,3}$+Relu	128	5×5	$256 \times 256 \times 128$
Max-pooling_1		2×2	$128 \times 128 \times 128$
$\text{Conv}_{4,5}$+Relu	128	5×5	$128 \times 128 \times 128$
Max-pooling_2		2×2	$64 \times 64 \times 128$
$\text{Conv}_{6,7}$+Relu	128	5×5	$64 \times 64 \times 128$
Max-pooling_3		2×2	$32 \times 32 \times 128$
Upsampling_1		2×2	$64 \times 64 \times 128$
$\text{Conv}_{8,9}$	128	5×5	$64 \times 64 \times 128$
Upsampling_2		2×2	$128 \times 128 \times 128$
$\text{Conv}_{10,11}$	128	5×5	$128 \times 128 \times 128$
Upsampling_3		2×2	$256 \times 256 \times 128$
$\text{Conv}_{12,13}$	128	5×5	$256 \times 256 \times 128$
Conv_{14}	9	1×1	$256 \times 256 \times 9$

B. Structure Design of the Network

The network used is based on the encoding-decoding structure, and the residual modules are used to help the network maintain the spatial structure information of the input image. Fig. 2 is the structure adopted by the network. Table II shows the specific configuration information of the network structure. The conv layers use filter windows kernels to extract image features. Output size of a conv layer is determined by size and number of kernels. Relu is an activation function that helps the network resist noise. The max-pooling layers are used to reduce parameters to suppress over fitting. For the semantic segmentation task, it is also necessary to use the upsampling layers to restore the feature layers to the original input size [21].

C. Phase Reconstruction Algorithm

The inversion of the wrap count map requires the following two steps.

1) *Information Decoding.* According to Table I, the azimuth and range gradient values can be directly obtained through the fused gradient information, forming a column vector $\boldsymbol{b}_{2n^2\times 1}$. The size of the input phase image slice is $n \times n$, and $n = 256$ in this study.
2) *Numerical Operations.* Storing the wrap count value in the row-major order in vector $\boldsymbol{x}_{n^2\times 1}$, the formula (4) is obtained [19]

$$\boldsymbol{A}_{2n^2\times n^2}\boldsymbol{x}_{n^2\times 1} = \boldsymbol{b}_{2n^2\times 1} \tag{4}$$

where $\boldsymbol{A}_{2n^2\times n^2}$ is the const coefficient matrix, and its structure is shown in the following equation:

$$\boldsymbol{A}_{2n^2\times n^2} = \begin{bmatrix} \Delta Kr_{n^2\times n^2} \\ \Delta Ka_{n^2\times n^2} \end{bmatrix}. \tag{5}$$

Here, $\Delta Kr_{n^2\times n^2}$ and $\Delta Ka_{n^2\times n^2}$ are the subcoefficient matrices. $\boldsymbol{x}_{\boldsymbol{n^2}\times 1}$ is the wrap count matrix obtained by unwrapping. The gradients in two directions of each pixel can be predicted and stored in $\boldsymbol{b}_{2n^2\times 1}$. Thus, the size of $\boldsymbol{A}$ is determined according to the matrix multiplication rule. Actually, $\Delta Kr_{n^2\times n^2}$ and $\Delta Ka_{n^2\times n^2}$ are used to describe the mapping relationship between $\boldsymbol{x}$ and $\boldsymbol{b}$ in the directions of range and azimuth. Fortunately, according to (2) and (3), the structure of $\boldsymbol{A}$ can be inferred, in which each row is composed of -1 and 1 at the corresponding gradient position, and the remaining positions are composed of 0. Now, $\boldsymbol{x}$ is the only unknown quantity in (4).

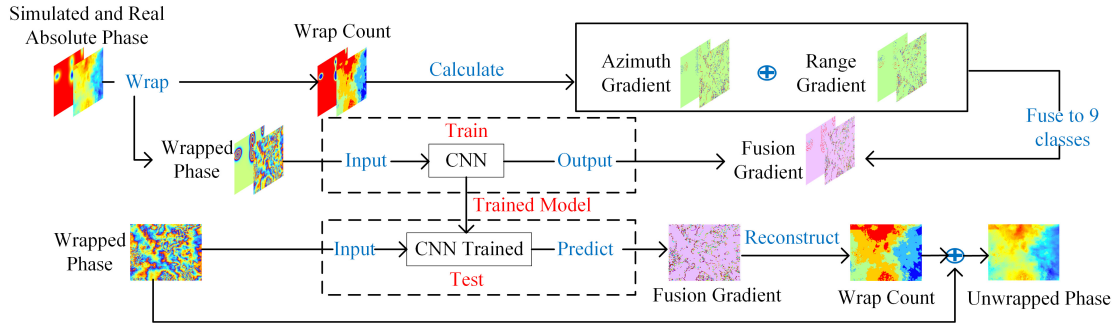

Fig. 1. Proposed method.

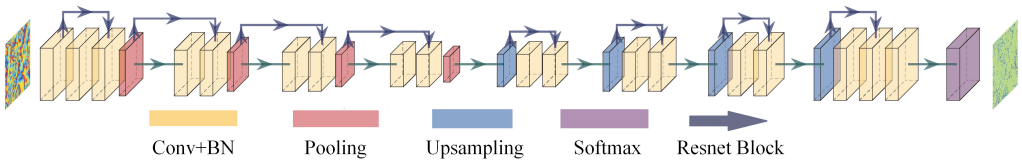

Fig. 2. Semantic segmentation network structure used in this letter.

TABLE III
IMBALANCE OF THE PHASE SAMPLE SET

Wrap count	1	2	3	4	5	6	7	8	9	>=10
Number of image slices	642	6578	7781	2851	1566	941	434	262	171	387
Proportions (%)	2.97	30.43	36.00	13.19	7.25	4.35	2.01	1.21	0.79	1.79

$\boldsymbol{A}_{2n^2\times n^2}$ is a sparse matrix. Numerical methods can be used to solve (4) to obtain the wrap count value.

III. Experimental Results

First, Sentinel-1 system parameters are used to simulate the data, and these data are used to let the network learn how to unwrap phase data. Next, the actual IW mode data are added to the sample set to improve the robustness and generalization ability of the network. The Sentinel-1 IW phase data is unwrapped by the SNAPHU method, then the unwrapped phase is rewrapped by (1), and the latter two form a data pair. The simulated data in the sample set accounts for about 10%, and the rest are real. The final phase dataset contains 21 613 data pairs composed of wrapped phases and corresponding unwrapped phases, of which 14 100 pairs are in the training set and 6613 pairs are in the test set.

Table III shows the imbalance of the phase sample set. According to statistics, the highest wrap count is even more than ten, that is, it is impossible to predict the wrap count for a new phase data. The traditional CNN PhU algorithms take the wrap count as the label, which is difficult to deal with the changeable wrap count under the condition of fixed classification ability. In order to make our network more stable in the face of imbalance data or new data, we use the fused gradient information to transform the network into a nine-class semantic segmentation problem. In this way, no matter how the fringe characteristics of phase data are, there is no need to worry about the insufficient classification ability of our network.

The network is trained to convergence state under the hardware conditions of a 3.6 GHz CPU and an 8G RTX 2080 GPU. The loss function adopts cross entropy loss which is used to accelerate the update of parameters of the CNN.

To verify the effectiveness and robustness of the method, the proposed method was compared with the MLPN method [5], the SNAPHU method [7], the PhaseNet method [8], and the two CNNs method [20]. Fig. 3 shows the unwrapped results of different methods. If residues in a phase image accounts for more than 20% of the image, the propagation of phase error is difficult to suppress, and some traditional algorithms (such as SNAPHU) cannot even complete unwrapping for stuck in a local area. In Fig. 3, residues in rows 1, 2, 7, and 8 account for more than 20%, rows 3 and 4 represent 10%, and rows 5 and 6 represent less than 1%. Rows 2, 4, 6, and 8 of Fig. 3 show the ground truth of the wrap count in different cases and the error maps between the equivalent wrap count predicted by each method and the true value. In error maps, green indicates that the prediction of the pixel is correct. The method proposed in this study has high unwrapping accuracy under different phase qualities and can suppress the propagation of PhU errors in low-quality regions.

The accuracy (ACC) of the classification of the wrap count, mean square error (MSE), and the average processing time of a single slice measure the performance of the algorithm

$$\mathrm{ACC}=\frac{\mathrm{Class}_{\mathrm{true}}}{\mathrm{Class}_{\mathrm{true}}+\mathrm{Class}_{\mathrm{false}}},\quad \mathrm{MSE}=\frac{\sum_{i=1}^{S}(\mathrm{UW}_i-\mathrm{AP}_i)^2}{S}. \tag{6}$$

Here, $\mathrm{Class}_{\mathrm{true}}$ represents the number of pixels that are correctly classified for the wrap count in a single slice, and $\mathrm{Class}_{\mathrm{false}}$ represents the number of pixels that are incorrectly classified. UW_i represents the PhU value of the ith pixel, AP_i represents the corresponding true value, and S represents the total number of pixels in a single slice.

Table IV shows the quantitative analysis results of the experiment. Fig. 4 shows the variation of the average accuracy of different methods with the increase of the proportion of residues. It can be seen from Table IV and Fig. 4 that our method has the highest wrap count classification accuracy and the lowest mean square error. In addition, the proposed

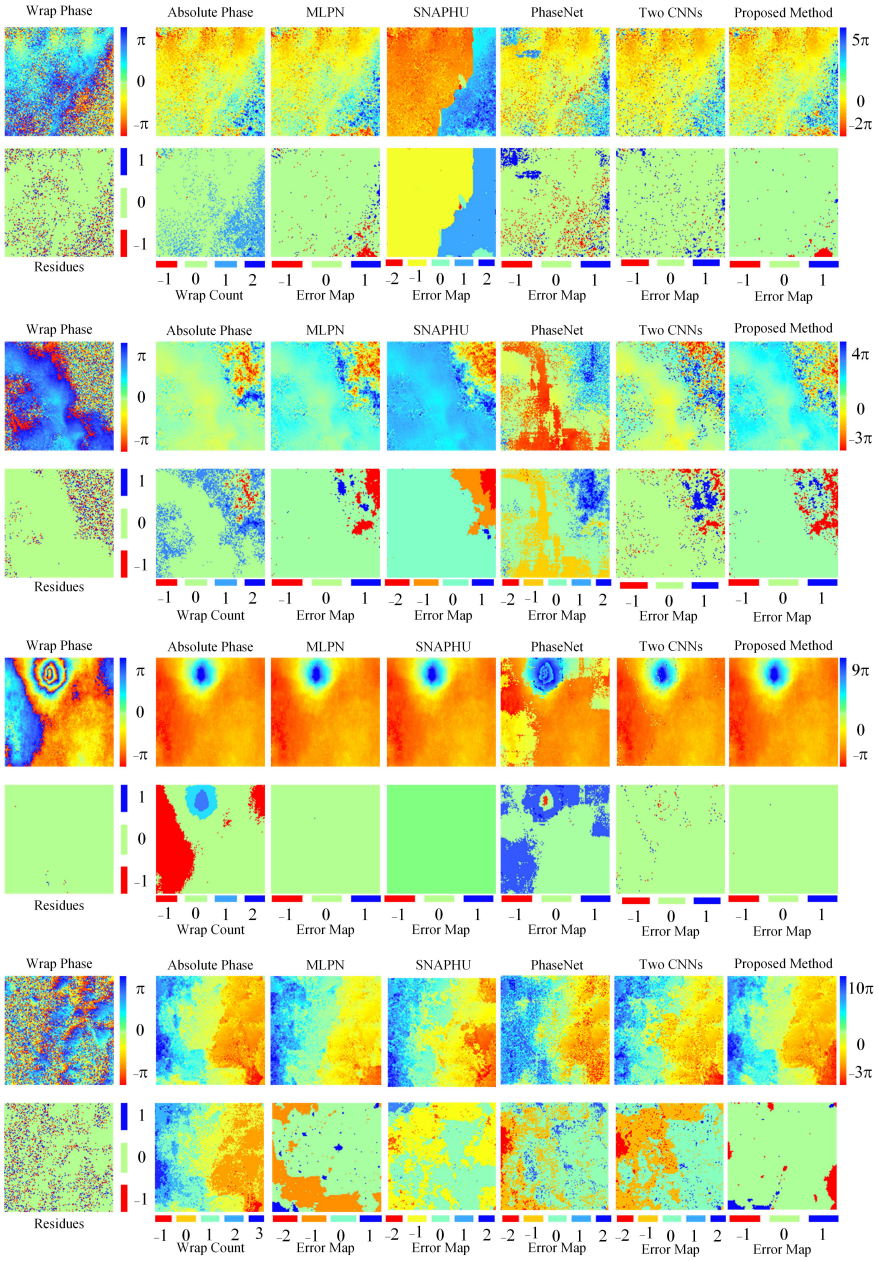

Fig. 3. PhU results in four different cases. The four cases from top to bottom include a large number of residues, a medium number of residues, a small number of residues with dense fringes, and a large number of residues with dense fringes. In the residue maps, −1, 0, and 1 represent negative residues, non-residues, and positive residues, respectively. The wrap count maps express the ground truth of the wrap count. The error maps refer to the error of the wrap count predicted by various methods, which is obtained by subtracting the predicted values from the true values.

TABLE IV

QUANTITATIVE COMPARISON OF UNWRAPPING EFFECTS ON THE TEST SET

Method	ACC(%)	MSE(rad^2)	Times(s)
Our method	97.00	0.9699	0.5081
MLPN	96.01	0.9909	0.8756
SNAPHU	96.08	1.3956	1.1284
PhaseNet	83.31	7.3440	0.1098
Two CNNs	96.70	1.4756	1.1704

method can achieve good unwrapped results in the case of different proportions of residues and also shows good unwrapping ability in the case of high fringe density. MLPN algorithm has high unwrapped accuracy and robustness, but its computational efficiency is low. The average unwrapping accuracy of the SNAPHU method is high, but its robustness is poor. In the case of more residues, the SNAPHU method will be trapped in the local area and unable to complete the accurate unwrapping of the whole slice sometimes. The PhaseNet has the best time efficiency, but it does not consider the gradient information, and it is difficult to deal with images with large phase residues. In the two CNN methods, the gradient error accumulation and propagation caused by the phase reconstruction is more serious. In conclusion, the proposed method has

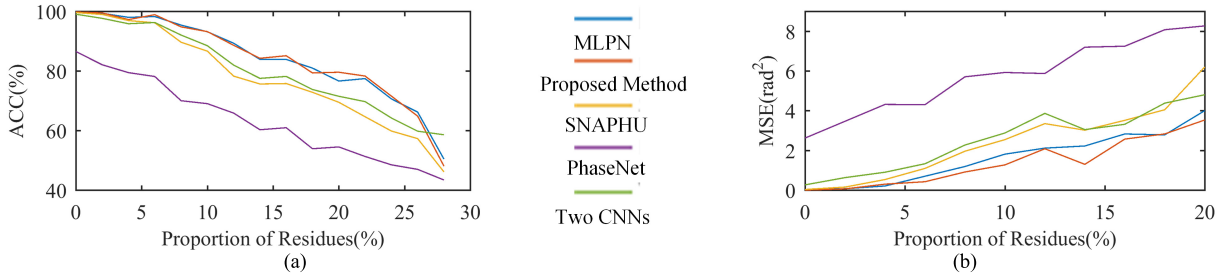

Fig. 4. Robustness evaluation of different methods. (a) Relationship between the proportion of residues and the ACC of different methods. (b) Relationship between the proportion of residues and the MSE of different methods.

high precision, high computational efficiency, and good anti-noise performance.

IV. Conclusion

This letter proposes a deep learning InSAR PhU method based on gradient information fusion to address the large number of residues and the unbalanced cycles in actual interferogram data. The use of gradient information can help the network fix the number of classification categories and better solve the problem of imbalance of fringe distribution; moreover, the fusion of gradient information simplifies the method process, reduces training and prediction costs, and simultaneously helps the phase reconstruction algorithm reduce the accumulation of gradient errors and improves the robustness of the algorithm. According to the experimental results, when faced with actual intricate InSAR phase data, the proposed method exhibits excellent performance in terms of computational efficiency, unwrapping accuracy, and adaptability to a variety of residue problems.

Furthermore, work should explore reasonable mosaic strategies to minimize the impact of local low-quality images on the global unwrapping effects. Furthermore, consideration of how the coherence coefficient affects the CNN especially when the coherence coefficient value is so low in large areas that there are too many noises is of great significance for the rational use of prior knowledge in InSAR.

Acknowledgment

The authors would like to thank the European Space Agency for providing Sentinel-1 images and Sentinel Application Platform (SNAP).

References

[1] 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, doi: 10.1109/MGRS.2018.2873644.

[2] 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, Jul./Aug. 1988, doi: 10.1029/RS023i004p00713.

[3] Q. L. Li, C. Y. Bao, J. S. Zhao, and Z. G. Jiang, "A new fast quality-guided flood-fill phase unwrapping algorithm," in *Proc. 3rd Annu. Int. Conf. Inf. Syst. Artif. Intell.*, vol. 1069, J. Tseng and I. Kotenko, Eds., 2018, Art. no. 012182.

[4] Z. Y. Suo, Z. F. Li, and Z. Bao, "A new strategy to estimate local fringe frequencies for InSAR phase noise reduction," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 771–775, Oct. 2010, doi: 10.1109/LGRS.2010.2047935.

[5] 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, doi: 10.1109/36.673674.

[6] 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, Opt. Image Sci.*, vol. 18, no. 2, pp. 338–351, Feb. 2001, doi: 10.1364/JOSAA.18.000338.

[7] H.-X. Bi and Z.-Q. Wei, "A new phase unwrapping method based on region recognition and region expansion," *Int. J. Remote Sens.*, vol. 37, no. 22, pp. 5287–5303, Nov. 2016, doi: 10.1080/01431161.2016.1214298.

[8] 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, doi: 10.1109/LSP.2018.2879184.

[9] G. Dardikman and N. T. Shaked, "Phase unwrapping using residual neural networks," in *Proc. Comput. Opt. Sens. Imag.*, 2018, Paper CW3B.5.

[10] G. Dardikman, N. A. Turko, and N. T. Shaked, "Deep learning approaches for unwrapping phase images with steep spatial gradients: A simulation," in *Proc. IEEE Int. Conf. Sci. Electr. Eng. Isr. (ICSEE)*, Dec. 2018, pp. 1–4, doi: 10.1109/ICSEE.2018.8646266.

[11] T. Zhang *et al.*, "Rapid and robust two-dimensional phase unwrapping via deep learning," *Opt. Exp.*, vol. 27, no. 16, pp. 23173–23185, Aug. 2019, doi: 10.1364/OE.27.023173.

[12] J. Di, K. Wang, Y. Li, and J. Zhao, "Deep learning-based holographic reconstruction in digital holography," in *Proc. Digit. Hologr. Three-Dimensional Imag.*, 2020, Paper HTu4B.2.

[13] 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, May 2019, doi: 10.1364/OE.27.015100.

[14] G. Dardikman-Yoffe, D. Roitshtain, S. K. Mirsky, N. A. Turko, M. Habaza, and N. T. Shaked, "PhUn-Net: Ready-to-use neural network for unwrapping quantitative phase images of biological cells," *Biomed. Opt. Exp.*, vol. 11, no. 2, pp. 1107–1121, Feb. 2020, doi: 10.1364/BOE.379533.

[15] C. C. Wu *et al.*, "Phase unwrapping based on a residual en-decoder network for phase images in Fourier domain Doppler optical coherence tomography," *Biomed. Opt. Exp.*, vol. 11, no. 4, pp. 1760–1771, Apr. 2020, doi: 10.1364/BOE.386101.

[16] F. Yang, T.-A. Pham, N. Brandenberg, M. P. Lutolf, J. Ma, and M. Unser, "Robust phase unwrapping via deep image prior for quantitative phase imaging," 2020, *arXiv:2009.11554*.

[17] J. Liang, J. Zhang, J. Shao, B. Song, B. Yao, and R. Liang, "Deep convolutional neural network phase unwrapping for fringe projection 3D imaging," *Sensors*, vol. 20, no. 13, p. 3691, Jul. 2020, doi: 10.3390/s20133691.

[18] 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, doi: 10.1109/TGRS.2020.2965918.

[19] 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, doi: 10.1109/TIP.2020.2977213.

[20] F. Sica, F. Calvanese, G. Scarpa, and P. Rizzoli, "A CNN-based coherence-driven approach for InSAR phase unwrapping," *IEEE Geosci. Remote Sens. Lett.*, early access, Oct. 21, 2020, doi: 10.1109/LGRS.2020.3029565.

[21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017, doi: 10.1145/3065386.