

Phase unwrapping in digital holography based on Dv3p-IRA

Yuan Chen^{1,2,*}, Peichao Li³, Hongwei Ma^{2,3}, Sitian Li³, Guangming Zhang⁴, Zhili Feng³, Ming Dong^{2,3}

1College of Sciences, Xi'an University of Science and Technology, Xi'an, Shaanxi, 710054, China

2Shaanxi Key Laboratory of Mine Electromechanical Equipment Intelligent Detection and Control, Xi'an University of Science and Technology, Xi'an, Shaanxi, 710054, China

3College of Mechanical Engineering, Xi'an University of Science and Technology, Xi'an, Shaanxi, 710054, China

4General Engineering Research Institute, School of Engineering, Faculty of Engineering and Technology, Liverpool John Moores University, Byrom Street, L3 3AF, United Kingdom

**Email: chenyan1030@126.com*

Abstract: The accurate acquisition of the phase information is crucial for the three-dimensional shape reconstruction of an object. However, the inverse tangent operation in the phase extraction process will inevitably cause the phase wrapping phenomenon, resulting in phase discontinuity. Therefore, a digital holographic phase unwrapping method based on Dv3p-IRA (Deeplabv3plus-Inverted-Residual-Attention) is proposed in this paper. This method takes Deeplabv3+ as the basic framework, and adopts the encoder-decoder structure: the encoder achieves multi-scale feature extraction through dense block and convolutional block attention module-atrous spatial pyramid pooling (CBAM-ASPP), while the decoder achieves phase reconstruction through cross-layer feature fusion and up-sampling. Random matrix enlargement (RME) and Gaussian function superposition (GFS) methods are used to construct the dataset, Cross-Entropy loss and Dice loss are integrated as the loss function to optimize the network parameters, and a reflective off-axis digital holographic optical path system is constructed for experimental verification. The simulation and experimental results show that compared with other methods, Dv3p-IRA achieves higher accuracy, with smaller fluctuations in the error value range and more stable model performance. In addition, it can effectively realize the separation of object information and background, and the reconstructed phase shape has good smoothness and continuity. Therefore, the proposed method not only realizes high-precision phase recovery, but also effectively deals with the holograms with speckle noise, and shows significant advantages in phase region segmentation and noise robustness.

Keywords: Digital holography; Phase unwrapping; Semantic segmentation; Dense block; Dilated convolution

1. INTRODUCTION

Digital holography has been widely used in the fields of medical imaging, contour measurement and defect detection due to its ability to record the phase information of the object to be measured. The technique can obtain the complex amplitude distribution containing the amplitude and phase information of the measured object by numerically reconstructing the hologram [1-4]. However, due to the inverse tangent operation in the phase extraction process, the obtained phase value is limited to the range of $(-\pi, \pi]$. If the actual phase variation exceeds 2π , the phase jump phenomenon occurs, i.e. phase wrapping [5]. This phase discontinuity will seriously affect the accuracy of phase analysis, so phase unwrapping is crucial to obtain accurate phase information.

The traditional phase unwrapping methods have some limitations in practical application, such as high sensitivity to residual points, poor noise resistance, insufficient ability to maintain phase continuity, and so on [6-10], which seriously affect the accuracy of phase recovery. In recent years, with the rapid development of deep learning technology, phase unwrapping methods based on neural networks have gradually become a research hotspot. Such methods take advantage of the powerful feature extraction and learning ability of convolutional neural networks, which can adaptively extract multi-scale features from complex interference images and achieve the accurate phase

information recovery.

The Deep learning based phase unwrapping methods can be mainly divided into two categories: deep learning regression method and deep learning wrap-count method. The deep learning regression method achieves phase recovery by learning the direct mapping relationship between the wrapped phase and the true phase. Wang et al. [11] achieved one-step phase unwrapping by using the U-Net network to directly learn the nonlinear relationship between the wrapped phase and the true phase. Xie et al. [12] introduced residual connection into the U-Net3+ architecture, which significantly improved the accuracy of phase reconstruction. Zhao et al. [13] proposed the VDE-Net method through a weighted jumping edge attention mechanism to improve the network's perception of phase boundaries. However, these methods are more sensitive to noise and are prone to significant errors when dealing with noisy wrapped phases. In contrast, the deep learning wrap-count method transforms the phase unwrapping problem into a pixel classification problem [14], which recovers the true phase of the measured object by dividing the wrapped phase map into different phase regions and assigning different integer multiples of 2π to each region. Spoorthi et al. [15] proposed the PhaseNet method by classifying each pixel point into predefined categories and verified the performance of the proposed network by validating it under strong noise conditions. Zhang et al. [16] presented the edge-enhanced self-attentive network (EESANet), which enhances the global features and thus improves the classification accuracy of each pixel. Spoorthi et al. [17] proposed the PhaseNet2.0 method based on the DenseNet network to solve the problems such as error accumulation in phase unwrapping, integrating the residuals by minimizing the gradient difference and overcoming the category imbalance problem by using the L1 loss function, and the results showed that this method has high robustness to noise. However, the existing methods still have the limitation of insufficient classification accuracy when dealing with different wrap-count boundaries.

To address the limitations of the above methods, we propose a phase unwrapping method based on Dv3p-IRA by taking Deeplabv3+ [18] as the network framework. This method introduces multiple dense blocks with integrated attention mechanism and dilated convolutional blocks with different expansion rates to obtain different degrees of semantic information. The dataset is constructed using RME and GFS, and the cross-entropy loss function and Dice loss function are integrated as the loss function to solve the sample category imbalance problem. The reflective off-axis digital holographic optical path system is constructed to obtain experimental data to verify the effectiveness of the Dv3p-IRA method by simulation and experiment.

2. DATASET CONSTRUCTION

When the deep learning wrap-count method is used for phase unwrapping, it is necessary to construct a dataset containing wrapped phases and wrap-counts. Due to the limitations of experimental conditions, it is difficult to obtain a large number of true phase data for network training, so a simulated dataset is generated using RME and GFS [19]. To ensure the randomness and uniformity of the data distribution, the initial matrix size of RME is set to [2,8], and the distribution of phase values is set according to a certain proportion, of which 40% is distributed in [10,30], 30% in [30,40], and the remaining 30% in [40,50], which is scaled up to a 128×128 size by linear interpolation [20,21]. The parameters of the GFS method are set as follows: the number of Gaussian functions ranges from [2,10], the phase height is [10,50], the standard deviation is [10,20], and the amplitude is 1. The number and centroids of the Gaussian functions are determined randomly, and the phase surfaces of the object light waves can be effectively simulated by superimposing

multiple randomly distributed Gaussian functions [22-25]. Based on the generated true phase data, the corresponding wrapped phases and wrap-counts are calculated to provide sufficient and well-distributed data samples for network training. The wrapped phase can be calculated by the following formula:

$$\varphi(x, y) = \arctan \left\{ \frac{\sin[\sigma(x, y)]}{\cos[\sigma(x, y)]} \right\} \quad (1)$$

where $\sigma(x, y)$ is the true phase and $\varphi(x, y)$ is the wrapped phase.

After obtaining the wrapped phase and the true phase, the wrap-count k can be obtained by:

$$k = \text{round} \left[\frac{\sigma(x, y) - \varphi(x, y)}{2\pi} \right] \quad (2)$$

where *round* denotes rounding, and k is an integer which represents the number of categories for semantic segmentation. Fig. 1 shows the flowchart for generating the simulated phase data.

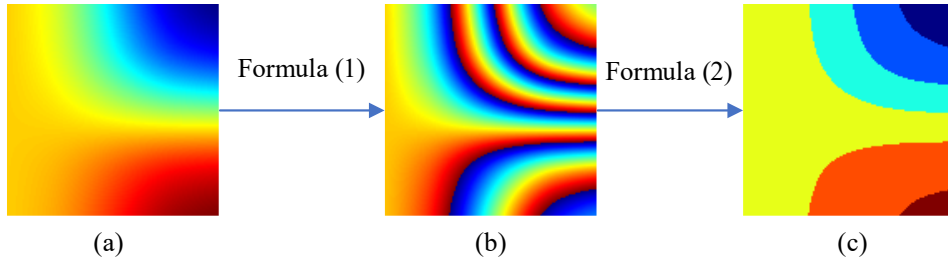


Fig. 1. The flowchart of simulated phase data generation: (a) true phase; (b) wrapped phase; (c) wrap-count.

To simulate the noise interference in the digital holographic measurement, Gaussian random noise with a standard deviation interval of [0,1] is added to the generated true phase, and then the corresponding wrapped phase is obtained by the inverse tangent function. The dataset constructed in this paper contains 20,000 groups of samples, and each group of samples consists of true phase, wrapped phase and wrap-count.

3. DV3P-IRA NETWORK MODEL

In this paper, the Dv3p-IRA network model is constructed for the phase unwrapping based on the Deeplabv3+ framework, which adopts the encoder-decoder structure to achieve semantic segmentation, and its network structure is shown in Fig. 2. The encoder is used to extract the multi-scale feature information of the image, which consists of Backbone and CBAM-ASPP. The Backbone is composed of multiple dense blocks mobile block component (MBC), which achieves down-sampling through four convolutional layers (layers 1,7, 9 and 12) with a step size of 2 to obtain different levels of image features (Low-Level Features 1-3). Among them, Low-Level Features 1-2 are directly input into the decoder for preliminary feature extraction, while Low-Level Feature 3 is input into the CBAM-ASPP module for deep feature enhancement. The CBAM-ASPP module combines dilated convolution with different expansion rates, point-by-point convolution and global average pooling, so that the model can obtain semantic information at different scales during feature extraction, and adds the CBAM to achieve weighted adjustment of channel and spatial resolution, which improves the perception ability of the model to key information. Finally, the five branches are fused to obtain Deep-Level Features.

The decoder combines the shallow, medium and deep features (X_1 , X_2 and X_3) through the cross-layer feature fusion mechanism, and recovers to the original resolution through up-sampling. The model achieves the synergistic optimization of local detail and global features, further enhancing the representation capability of the network.

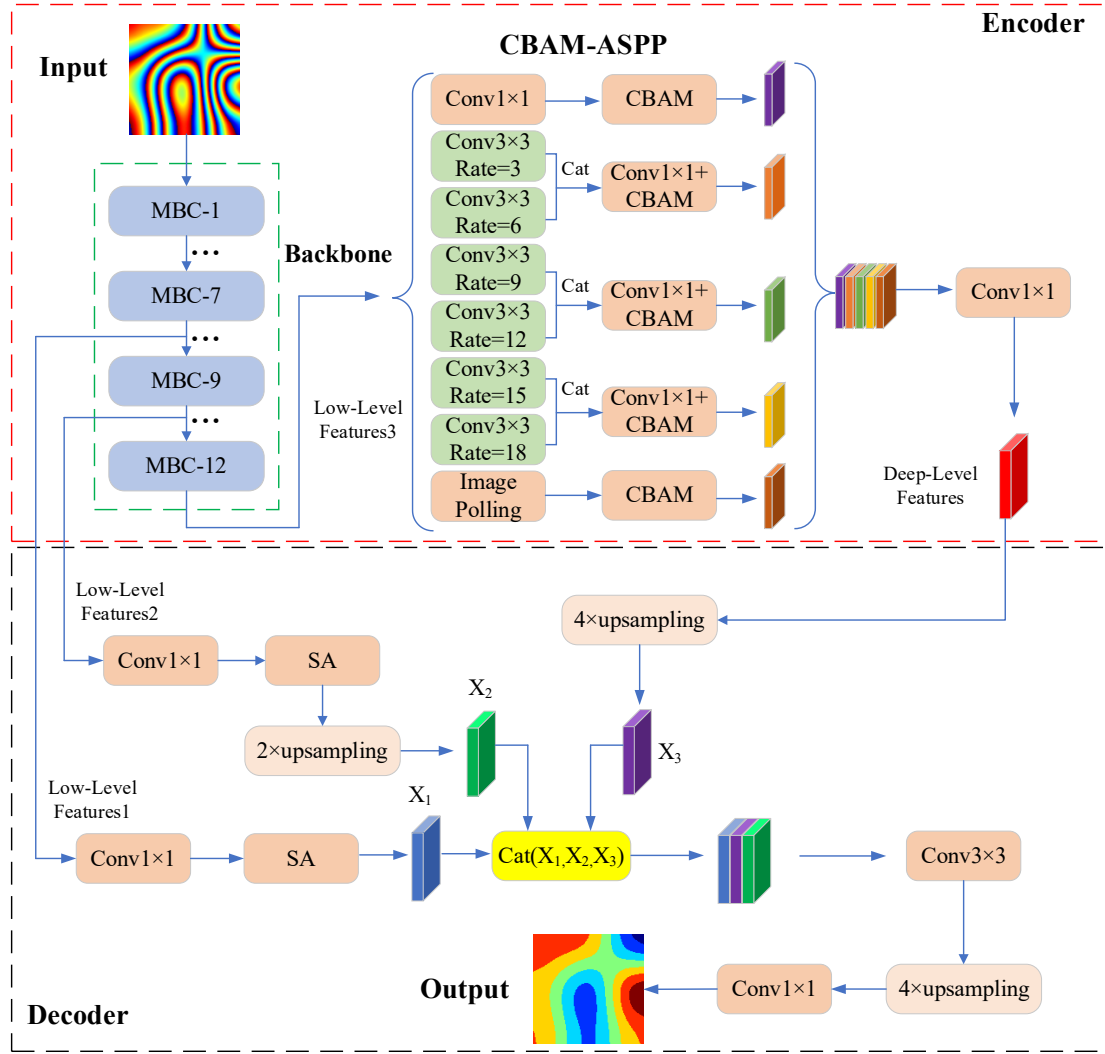


Fig. 2. Dv3p-IRA network structure.

3.1 Encoder

The Encoder structure is shown in Fig. 3, which consists of Backbone and CBAM-ASPP module. Among them, the MBC is composed of inverted residual block and squeeze-and-excitation (SE), which can significantly reduce the computational complexity of the model while ensuring the accuracy of network segmentation. The specific implementation process is described as follows: firstly, for the 128×128 wrapped phases as the input, the number of channels is increased by 1×1 convolution. The BN layer normalizes the data to speed up the model training and improve the stability, and the Hard-Swish activation function improves the nonlinear representation ability, and the coordination between the two improves the generalization ability of the model to reduce the risk of overfitting; then, the deep feature extraction is performed by 3×3 convolution, and the SE module is added to dynamically adjust the channel weights of the feature maps so that the model can adaptively focus on the important feature information; finally, the dimensionality reduction output is performed by 1×1 convolution. In the model design process, the number of iterations of the MBC module is set to 12 by experimental verification. Three different levels of features (Low-Level Features 1-2, Deep-Level Features) can be obtained through the coding process, which achieves a balance between computational efficiency and feature expression capability, and lays a solid foundation for feature fusion and reconstruction of the subsequent decoder.

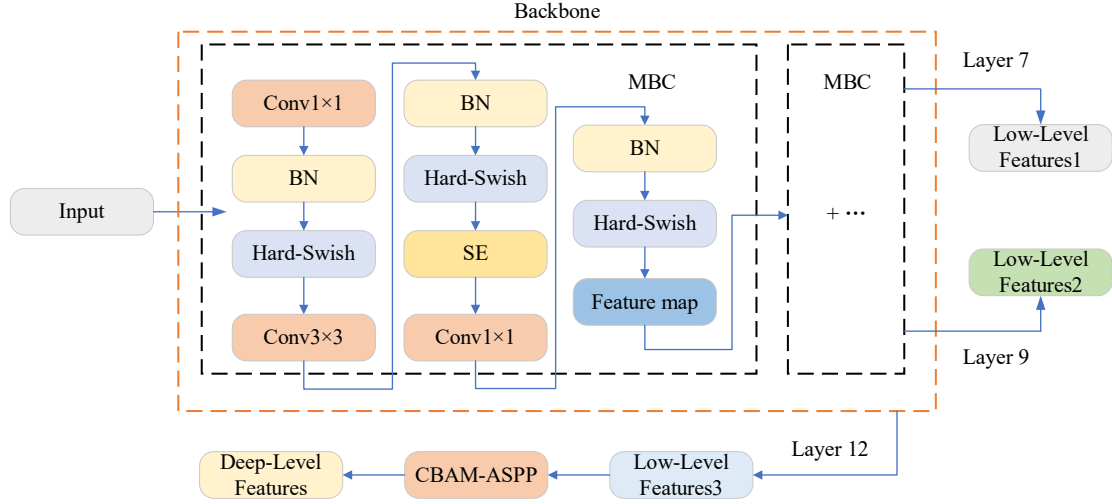


Fig. 3. Encoder structure.

3.2 Decoder

The Decoder structure is shown in Fig. 4. Firstly, the number of channels of the three features extracted by the Encoder is adjusted by 1×1 convolution, and the addition of the ReLu activation function can effectively maintain the nonlinear representation ability of the multi-scale features and mitigate the vanishing gradient problem. In addition, the spatial attention (SA) is introduced to improve the representation ability of important spatial information. Then, the Low-Level Features2 and the Deep-Level Features are uniformly adjusted to 32×32 resolution by 2 times and 4 times up-sampling respectively to obtain the feature maps X_1 , X_2 and X_3 , which are fused by 3×3 convolution. And the Dropout layer is introduced to improve the generalization ability of the model. Finally, the original image resolution is restored by 4 times up-sampling, and the number of output channels is adjusted to the number of categories by 1×1 convolution to output the network prediction results.

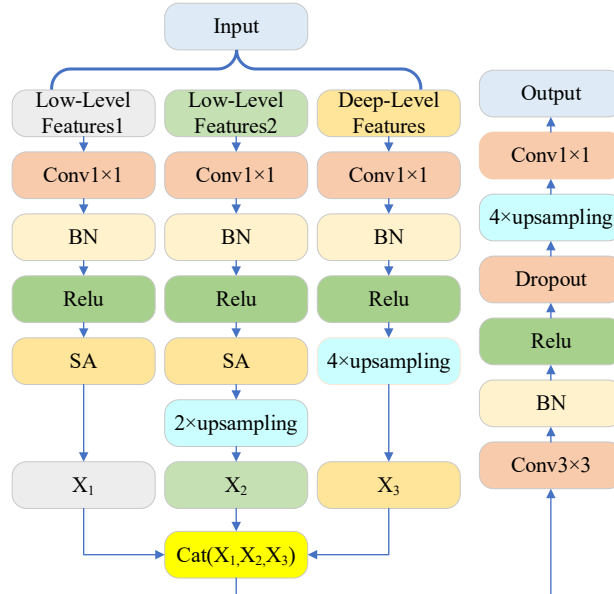


Fig. 4. Decoder structure.

3.3 Model Training Parameter Settings

In this paper, the Adam optimizer is used to train the model, the initial learning rate is set to 0.001, and iteratively updated by a multiple of 0.95 until 1×10^{-5} is reached to stop the iteration. The

training period is 100, and the number of samples in each batch is 16. The constructed dataset contains 20,000 sets of data, which are randomly divided into training, validation and test sets in a ratio of 7:2:1. Aiming at the common category imbalance problem in the semantic segmentation task [16], this paper combines the cross-entropy loss function and the Dice loss function, and dynamically adjusts the weights of the two during the training process, so as to ensure the classification accuracy while effectively alleviating the category imbalance problem, and further improve the model performance. The cross-entropy loss function is expressed by:

$$L_{cross} = - \sum_{i=1}^N y_i \log(p_i) \quad (3)$$

where N is the number of categories, y_i is the label value, and p_i is the probability of belonging to that category.

The Dice loss function is expressed as:

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^M p_i y_i}{\sum_{i=1}^M p_i \sum_{i=1}^M y_i} \quad (4)$$

where M is the total number of pixels. The combinational loss function can be expressed as:

$$L_{total} = \omega \times L_{cross} + (1 - \omega) \times L_{Dice} \quad (5)$$

where ω is the weight. It has been experimentally verified that $\omega=0.7$.

4. SIMULATION VERIFICATION AND RESULT ANALYSIS

The Dv3p-IRA model is trained and verified using the simulated dataset, and the results are shown in Fig.5.

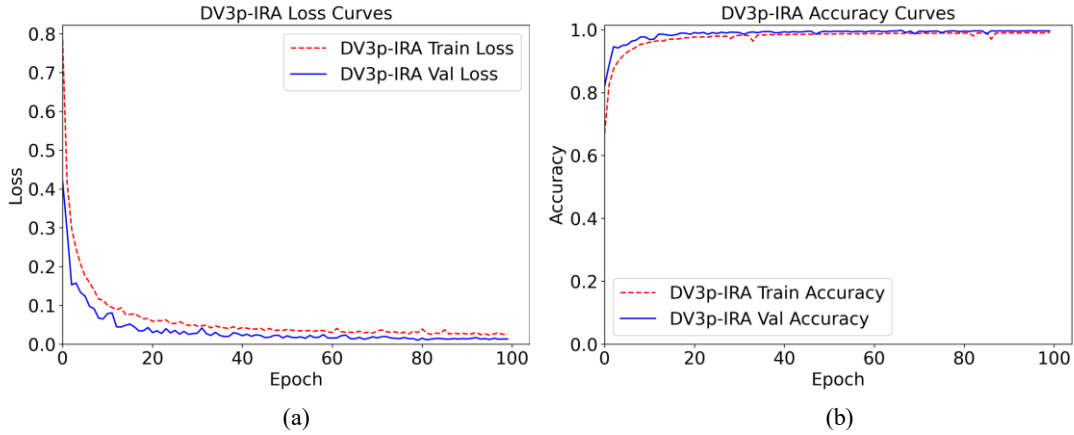


Fig. 5. Network training and validation results: (a) loss curves; (b) accuracy curves.

As can be seen from Fig. 5, the Dv3p-IRA model has reached convergence when the number of network iterations reaches 100. In particular, both the training loss and the validation loss decrease synchronously with the increase of the training period and gradually level off, indicating that the model is not overfitted. Both the training accuracy and validation accuracy gradually increase with the same trend. It shows that the model has good generalization ability.

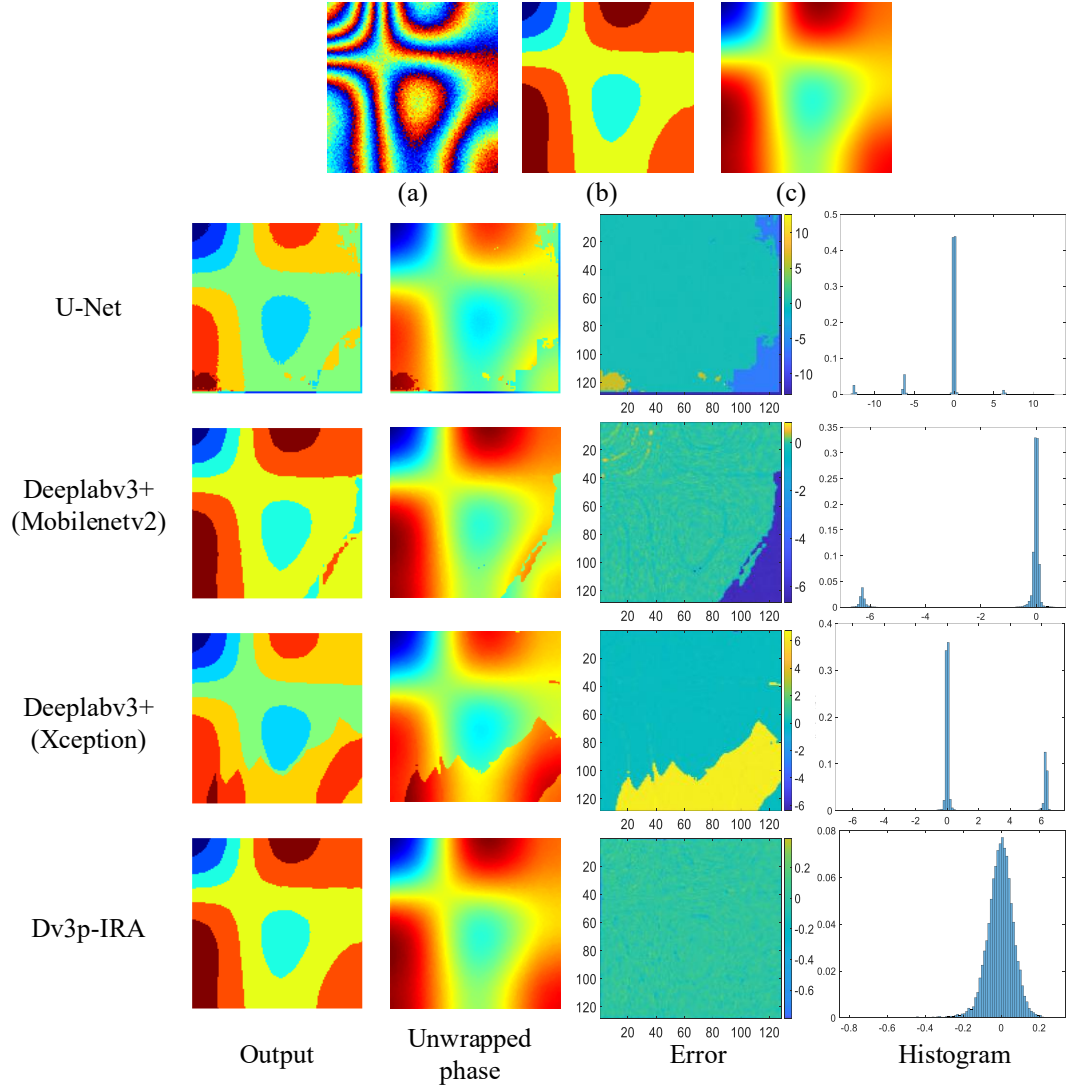


Fig. 6. Comparison of unwrapping results of different methods:(a) wrapped phase; (b) wrap-count; (c) true phase.

To verify the unwrapping performance of the method proposed in this paper, the Dv3p-IRA model is tested using the test set, and the unwrapping results are compared and analyzed with those of U-Net, Deeplabv3+Mobilenetv2 and Deeplabv3+Xception. The results are shown in Fig. 6. It can be seen that U-Net and Deeplabv3+Xception have obvious classification errors in phase unwrapping, and their prediction results have large misclassified regions, resulting in poor phase unwrapping. Statistical analysis of the errors shows that the error histograms of these two methods have bimodal distribution characteristics, with the peaks located at anomalous offset positions such as -12 and 6, respectively, reflecting the lack of model robustness. In contrast, the performance of Deeplabv3+Mobilenetv2 is improved to some extent, but it still suffers from local misclassification and fails to achieve accurate phase region segmentation. On the other hand, the Dv3p-IRA method proposed in this paper has optimal performance, its classification results are closer to the wrap-count, and the error interval is controlled within the range of only $[-0.6, 0.2]$. In addition, the distribution of the error histogram of this method is concentrated without significant bias, indicating that the prediction results of this model have high accuracy.

In order to quantitatively assess the model performance, four metrics are used to evaluate the test results, namely mean intersection over union (MIOU), precision, structural similarity index (SSIM)

and root mean square error (RMSE).

(1) MIOU is used to measure the agreement between the segmentation results predicted by the model and the ground truth, which can be calculated by averaging the intersection and concurrency ratios for each category with the expression:

$$MIOU = \frac{1}{k} \sum_{i=1}^k \frac{TP}{FN + FP + TP} \quad (6)$$

where k is the number of categories, TP denotes the case where the prediction is a positive case and is actually a positive case, FP denotes the case where the prediction is a positive case and is actually a negative case, and FN denotes the case where the prediction is a negative case and is actually a positive case.

(2) Precision is the proportion of predicted positive cases that are actually positive, which is expressed as:

$$Precision = \frac{1}{k} \sum_{i=1}^k \frac{TP}{TP + FP} \quad (7)$$

(3) SSIM is used to measure the similarity between two images in terms of luminance, contrast and structure [26], which is calculated by:

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (8)$$

where μ is the mean value of image, σ is the standard deviation, and c_1 and c_2 is a smaller constant to ensure that the denominator is non-zero.

(4) RMSE is used to measure the error between the predicted and the true values of the model [27], which is calculated by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

where n is the number of samples, y_i is the true value, and $\hat{y}_i$ is the predicted value.

The calculation results of the four metrics for the test results of the different network models are shown in Table 1. As can be seen from Table 1, the Dv3p-IRA model proposed in this paper has significant advantages over U-Net, Deeplabv3+Xception and Deeplabv3+ Mobilenetv2: the MIOU values are improved by 7.4%, 7.1% and 4.1% respectively, the Precision values are improved by 2.6%, 1.9% and 1.0% respectively, the SSIM values are increased by 16.0%, 12.1% and 8.6% respectively, and the RMSE values are reduced by 80.9%, 77.4% and 72.4% respectively. Therefore, the Dv3p-IRA model achieves optimal performance in all four metrics, demonstrating the effectiveness of the proposed method.

Table 1 The mean values of the four metrics for the test results of different network models.

Network models	MIOU	Precision	SSIM	RMSE
U-Net	0.9212	0.9678	0.8557	0.1646
Deeplabv3+Xception	0.9236	0.9748	0.8851	0.1391
Deeplabv3+Mobilenetv2	0.9496	0.9832	0.9138	0.1143
Dv3p-IRA	0.9890	0.9931	0.9923	0.0315

5. EXPERIMENTAL VERIFICATION AND RESULT ANALYSIS

5.1 Experimental Setup

To further verify the effectiveness of the Dv3p-IRA model, a reflective off-axis digital holographic optical path system is constructed, as shown in Fig. 7. The USA-1951 test target and the specimen with defects as shown in Fig. 8 are used as the experimental samples, and the holograms of the regions marked by the red box are collected for experimental verification. Two traditional unwrapping methods, i.e. DCT and GBC, and Deeplabv3+, which has better performance in the simulation results, are used to compare and analyze the test results with the Dv3p-IRA model proposed in this paper.

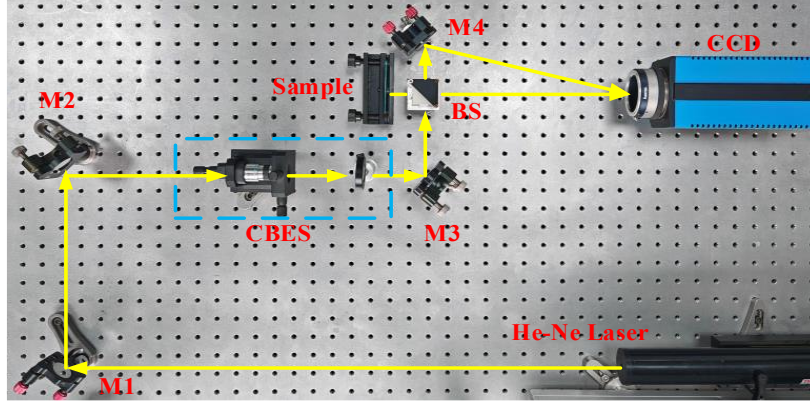


Fig. 7. Experimental setup.

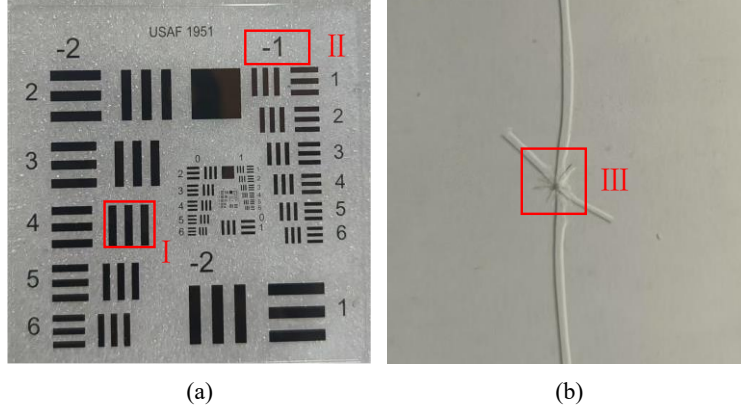


Fig. 8. Test samples: (a) USA-1951 test target; (b) specimen with defects.

5.2 Analysis of Experimental Results

The phase unwrapping results of the four methods for the hologram of region I are shown in Fig.9. It can be seen that the traditional DCT and GBC methods have obvious speckle noise interference, which causes the reconstructed phase map to show global phase error and makes it difficult to achieve the effective separation of the object phase information from the background. The Deeplabv3+ method improves the phase quality to some extent, but the right side of the object fringe information is lost due to the misclassification of some pixel points. In contrast, the Dv3p-IRA method proposed in this paper has the optimal performance, which not only separates the object fringe information from the background, but also has the highest integrity and clarity of the phase information, and has significant advantages in both noise robustness and phase fidelity.

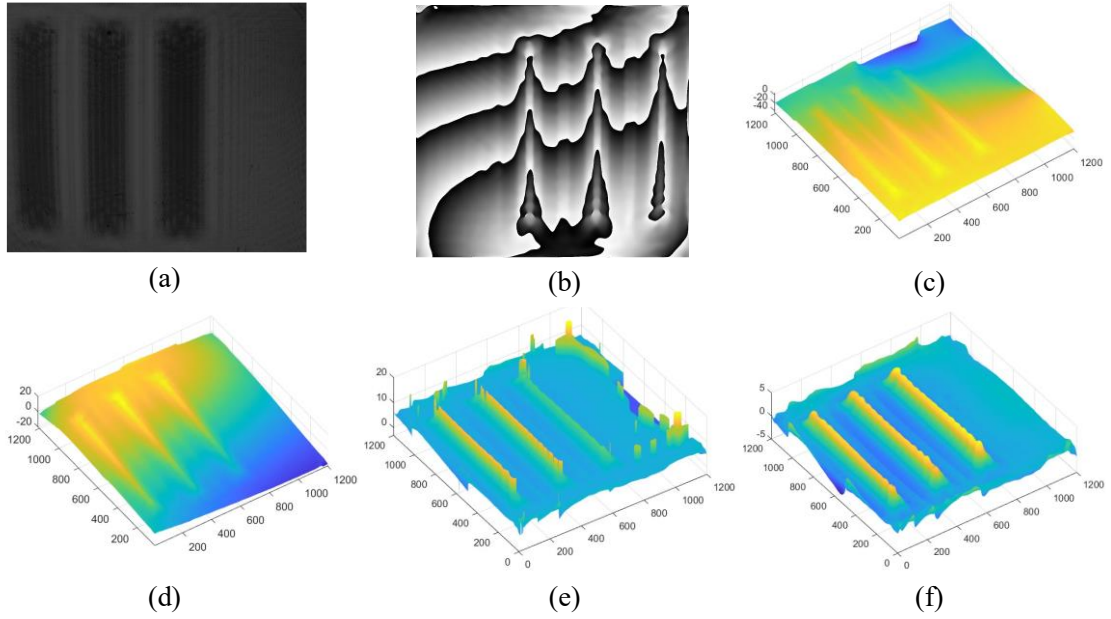


Fig. 9. Test results of four methods for region I: (a) hologram; (b) wrapped phase; (c) DCT; (d) GBC; (e) Deeplabv3+; (f) Dv3p-IRA.

The phase unwrapping results of the four methods for the hologram of region II are shown in Fig.10. It can be seen that the DCT and GBC methods can only reconstruct the fuzzy contour of the “-1” pattern, and the important morphological features are seriously lost. The GBC method generates obvious phase voids due to the effect of residual points, which seriously affects the phase unwrapping effect. Although Deeplabv3+ can recover the basic contour of the object, phase perturbation occurs in the edge region, affecting the reconstruction quality. In contrast, the Dv3p-IRA method proposed in this paper shows good noise suppression ability, and the reconstructed phase shape is smooth and continuous with details preserved, which is significantly better than the other three methods.

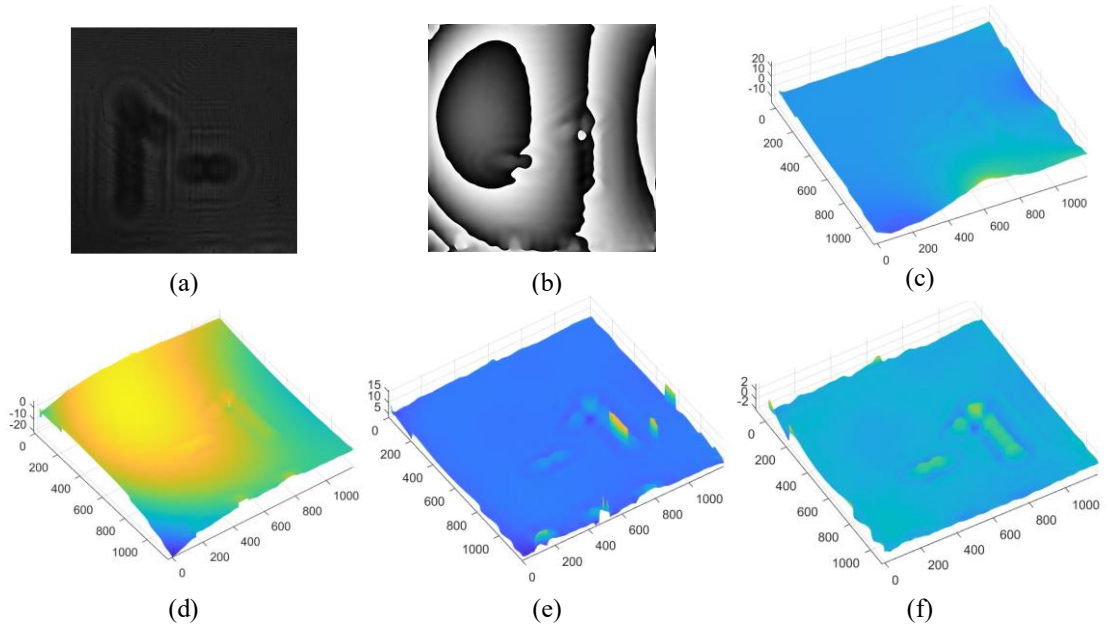


Fig. 10. Test results of four methods for region II: (a) hologram; (b) wrapped phase; (c) DCT; (d) GBC; (e) Deeplabv3+; (f) Dv3p-IRA.

The phase unwrapping results of the four methods for the hologram of region III are shown in Fig. 11. It can be seen that the GBC and DCT methods generate significant phase fluctuations in the boundary region, resulting in the loss of phase information at the edges. Although the Deeplabv3+ method improves the quality of the phase reconstruction, there are pixel classification errors in the lower left corner of the phase map, and the phase information at the central defect is lost. In contrast, the Dv3p-IRA method proposed in this paper shows better phase reconstruction performance, resulting not only in smoother processing of the information in the edge region of the phase map, but also in better reconstruction effect at the defects.

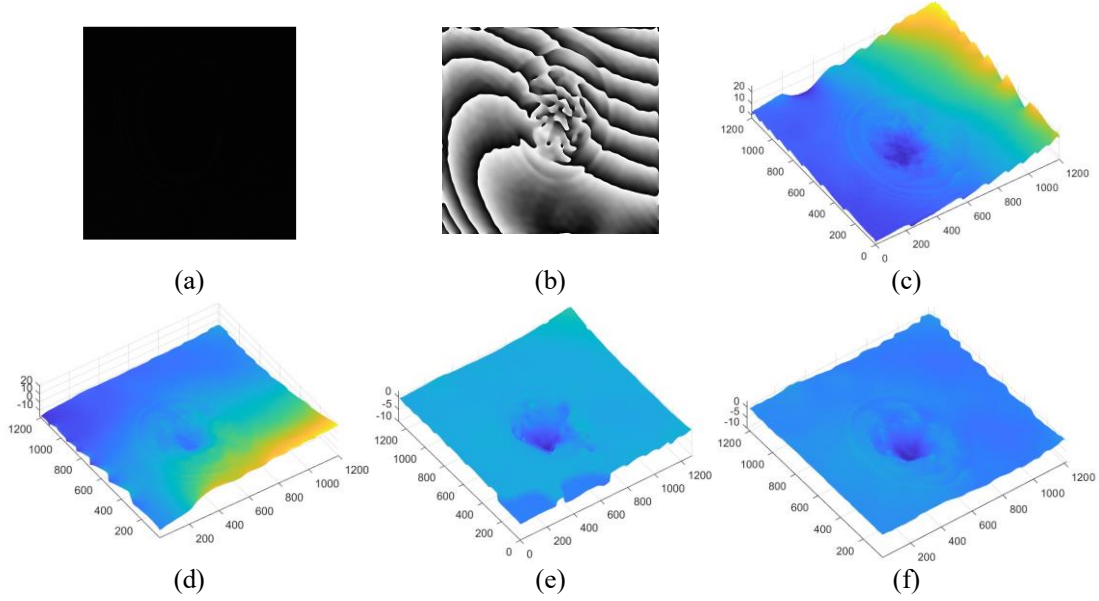


Fig. 11. Test results of four methods for region III: (a) hologram; (b) wrapped phase; (c) DCT; (d) GBC; (e) Deeplabv3+; (f) Dv3p-IRA.

The above three sets of experimental results show that the DCT and GBC unwrapping methods have insufficient noise resistance and are difficult to achieve effective separation of object phase and background information; although the deep learning based unwrapping method is significantly better than the two traditional methods, Deeplabv3+ still has phase anomalies due to misclassified pixel points. In contrast, the Dv3p-IRA method proposed in this paper has excellent noise resistance and higher unwrapping accuracy, which can accurately extract object phase information.

6. CONCLUSIONS

In this paper, a phase unwrapping method based on Dv3p-IRA was proposed. The simulated datasets are constructed using the RME and GFS methods for training and testing of network models. The simulation results show that the Dv3p-IRA method outperforms other deep learning methods in terms of segmentation accuracy, error distribution and histogram. In addition, compared to the U-Net, Deeplabv3+Xception and Deeplabv3+Mobilenetv2 methods, the results obtained by the Dv3p-IRA method have been improved the MIOU values by 7.4%, 7.1% and 4.1% respectively, the precision values by 2.6%, 1.9% and 1.0% respectively, the SSIM values by 16.0%, 12.1% and 8.6% respectively, while the RMSE values have been reduced by 80.9%, 77.4% and 72.4% respectively. To further verify the effectiveness of the proposed method, a reflective off-axis digital holographic optical path system is constructed to collect the holograms of USA-1951 test target and a specimen with defects for experimental verification. The results show that the Dv3p-IRA method can effectively suppress the speckle noise and accurately recover the true phase of the object, which

significantly improves the accuracy and robustness of phase unwrapping.

Funding. National Natural Science Foundation of China (Grant No. 52175518).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

- [1] Cheremkhin P A, Evtikhiev N N, Kozlov A V, et al. An optical-digital method of noise suppression in digital holography[J]. *Journal of Optics*, 2022, 24(11): 115702.
- [2] Huang L, Yan L, Chen B. Phase restoration of digital holographic microscopy with an adaptive reliability mask for phase unwrapping in microstructure testing[J]. *Optics and Lasers in Engineering*, 2021, 138: 106416.
- [3] Petrov V, Pogoda A, Sementin V, et al. Advances in digital holographic interferometry[J]. *Journal of Imaging*, 2022, 8(7): 196.
- [4] Picart P, Hassad S. Imaging acoustic field with multicolor digital holography[C]. *Holography, Diffractive Optics, and Applications XII*. SPIE, 2022, 12318: 244-249.
- [5] Park S, Kim Y, Moon I. Automated phase unwrapping in digital holography with deep learning[J]. *Biomedical Optics Express*, 2021, 12(11): 7064-7081.
- [6] Wang X, Fang S, Zhu X, et al. Phase unwrapping based on adaptive image in-painting of fringe patterns in measuring gear tooth flanks by laser interferometry[J]. *Optics Express*, 2020, 28(12): 17881-17897.
- [7] Zhao M, Huang L, Zhang Q, et al. Quality-guided phase unwrapping technique: comparison of quality maps and guiding strategies[J]. *Applied Optics*, 2011, 50(33): 6214-6224.
- [8] Flynn T J. Two-dimensional phase unwrapping with minimum weighted discontinuity[J]. *Journal of the Optical Society of America A*, 1997, 14(10): 2692-2701.
- [9] Asundi A, Wensen Z. Fast phase-unwrapping algorithm based on a gray-scale mask and flood fill[J]. *Applied Optics*, 1998, 37(23): 5416-5420.
- [10] Zhao Z, Zhang H, Ma C, et al. Comparative study of phase unwrapping algorithms based on solving the Poisson equation[J]. *Measurement Science and Technology*, 2020, 31(6): 065004.
- [11] Wang K, Li Y, Kemao Q, et al. One-step robust deep learning phase unwrapping[J]. *Optics Express*, 2019, 27(10): 15100-15115.
- [12] Xie X, Tian X, Shou Z, et al. Deep learning phase-unwrapping method based on adaptive noise evaluation[J]. *Applied Optics*, 2022, 61(23): 6861-6870.
- [13] Zhao J, Liu L, Wang T, et al. VDE-Net: a two-stage deep learning method for phase unwrapping[J]. *Optics Express*, 2022, 30(22): 39794-39815.
- [14] Zhang J, Tian X, Shao J, et al. Phase unwrapping in optical metrology via denoised and convolutional segmentation networks[J]. *Optics Express*, 2019, 27(10): 14903-14912.
- [15] Spoorthi G E, Gorthi S, Gorthi R K S S. PhaseNet: A deep convolutional neural network for two-dimensional phase unwrapping[J]. *IEEE Signal Processing Letters*, 2018, 26(1): 54-58.
- [16] Zhang J, Li Q. EESANet: edge-enhanced self-attention network for two-dimensional phase unwrapping[J]. *Optics Express*, 2022, 30(7): 10470-10490.
- [17] Spoorthi G E, Gorthi R K S S, Gorthi S. PhaseNet 2.0: Phase unwrapping of noisy data based on deep learning approach[J]. *IEEE transactions on Image Processing*, 2020, 29: 4862-4872.
- [18] Li D, Xie X. Deep learning-based phase unwrapping method[J]. *IEEE Access*, 2023, 11:

85836-85851.

- [19] Chen Y, Wang Q, Zhang G, et al. Phase unwrapping in digital holography based on SRDU-net[J]. Optics Communications, 2024, 573: 131055.
- [20] Wang K, Kemao Q, Di J, et al. Deep learning spatial phase unwrapping: a comparative review[J]. Advanced Photonics Nexus, 2022, 1(1): 014001.
- [21] Qin Y, Wan S, Wan Y, et al. Direct and accurate phase unwrapping with deep neural network[J]. Applied Optics, 2020, 59(24): 7258-7267.
- [22] Perera M V, De Silva A. A joint convolutional and spatial quad-directional LSTM network for phase unwrapping[C]. Iccasp 2021-2021 IEEE international conference on acoustics, speech and signal processing. IEEE, 2021: 4055-4059.
- [23] Wu C, Qiao Z, Zhang N, et al. Phase unwrapping based on a residual en-decoder network for phase images in Fourier domain Doppler optical coherence tomography[J]. Biomedical Optics Express, 2020, 11(4): 1760-1771.
- [24] Vengala K S, Paluru N, Subrahmanyam Gorthi R K S. 3D deformation measurement in digital holographic interferometry using a multitask deep learning architecture[J]. Journal of the Optical Society of America A, 2021, 39(1): 167-176.
- [25] Sumanth K, Ravi V, Gorthi R K. A multi-task learning for 2D phase unwrapping in fringe projection[J]. IEEE Signal Processing Letters, 2022, 29: 797-801.
- [26] Camacho A S, Velez-Zea A, Barrera-Ramírez J F. Unwrapping of phase-only holographic data using a convolutional neural network[J]. Optics Communications, 2025, 577: 131395.
- [27] Zhang Z, Wang X, Liu C, et al. Efficient and robust phase unwrapping method based on SFNet[J]. Optics Express, 2024, 32(9): 15410-15432.