

# Unifying temporal phase unwrapping framework using deep learning

XINMING GUO,<sup>1,2,3,†</sup> YIXUAN LI,<sup>1,2,3,†</sup>  JIAMING QIAN,<sup>1,2,3</sup>   
YUXUAN CHE,<sup>1,2,3</sup> CHAO ZUO,<sup>1,2,3</sup>  QIAN CHEN,<sup>1,3</sup>  EDMUND Y.  
LAM,<sup>4</sup>  HUAI WANG,<sup>5</sup> AND SHIJIE FENG<sup>1,2,3,\*</sup> 

<sup>1</sup>Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China

<sup>2</sup>Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China

<sup>3</sup>Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing, Jiangsu Province 210094, China

<sup>4</sup>Department of Electrical and Electronic Engineering, University of Hong Kong, Pokfulam Road, Hong Kong

<sup>5</sup>Suzhou Abham Intelligence Technology Co., Ltd, Suzhou, Jiangsu Province 215000, China

<sup>†</sup>These authors contributed equally to this work.

\*shijiefeng@njjust.edu.cn

**Abstract:** Temporal phase unwrapping (TPU) is significant for recovering an unambiguous phase of discontinuous surfaces or spatially isolated objects in fringe projection profilometry. Generally, temporal phase unwrapping algorithms can be classified into three groups: the multi-frequency (hierarchical) approach, the multi-wavelength (heterodyne) approach, and the number-theoretic approach. For all of them, extra fringe patterns of different spatial frequencies are required for retrieving the absolute phase. Due to the influence of image noise, people have to use many auxiliary patterns for high-accuracy phase unwrapping. Consequently, image noise limits the efficiency and the measurement speed greatly. Further, these three groups of TPU algorithms have their own theories and are usually applied in different ways. In this work, for the first time to our knowledge, we show that a generalized framework using deep learning can be developed to perform the TPU task for different groups of TPU algorithms. Experimental results show that benefiting from the assistance of deep learning the proposed framework can mitigate the impact of noise effectively and enhance the phase unwrapping reliability significantly without increasing the number of auxiliary patterns for different TPU approaches. We believe that the proposed method demonstrates great potential for developing powerful and reliable phase retrieval techniques.

© 2023 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

## 1. Introduction

Fringe projection profilometry (FPP), as an optical 3D shape measurement method, is widely used in machine vision, automated optical inspection, visual inspection, industrial quality control, biomedicine, and other research fields [1–5]. The simplest FPP system is comprised of a camera, a projector and a processing unit (computer). Under the control of the computer, a series of well-designed fringe patterns are projected onto the target object. The camera captures the corresponding deformed fringe patterns, which contain the phase information of the projected patterns. The computer then performs a decoding algorithm to extract the phase information from the deformed fringe patterns and maps it to the real-world 3D coordinates of the object by triangulation. In order to measure the surface shape of an object using phase information, a number of fringe projection techniques have been proposed, among which phase-shifting (PS) profilometry [6,7] and Fourier transform profilometry [8–10] are the two representative methods

currently used to obtain phase. Both techniques estimate the phase distribution by performing an arctangent calculation over the four quadrants of phasor space, resulting in the extracted phase having a  $2\pi$  jump. Therefore, it is necessary to implement phase unwrapping after phase retrieval to eliminate phase ambiguity [11–14].

At present, the common phase unwrapping methods can be divided into two categories: the spatial phase unwrapping (SPU) method [11,12] and the temporal phase unwrapping (TPU) method [13]. SPU calculates the relative fringe order of the center pixel on a single-wrapped phase map by analyzing the phase information of its neighboring pixels. This category of methods is characterized by the ability to perform phase unwrapping only using information about the phase of the surrounding pixels. However, it is difficult for the SPU method to obtain reliable unwrapping phase results for discontinuous surfaces (e.g., step surfaces) in general [15,16]. In order to solve these problems, TPU methods have been proposed to achieve pixel-by-pixel, robust absolute phase unwrapping by projecting additional patterns [17–21]. Since the temporal phase unwrapping method is performed pixel by pixel, it can accurately measure objects with discontinuous surfaces. Currently, there are typically three groups of TPU methods proposed in the literature: the multi-frequency (hierarchical) method [18,20,22–24], the multi-wavelength (heterodyne) method [15,25–29], and the number-theoretic method [30–34]. Zuo *et al.* [13] performed comparative studies of these three phase unwrapping methods by building a rigorous mathematical model and showed that the multi-frequency method has the best performance among these methods. Further, all three TPU methods demonstrated unreliable phase unwrapping performance as the frequency increases, and they are sensitive to noise.

In recent years, deep learning techniques have been used in 3D imaging [35–43]. Feng *et al.* [35] demonstrated that trained deep neural networks can perform fringe analysis by rapidly predicting background images and estimating the numerator and denominator of the arctangent function, thereby improving the phase demodulation accuracy using a single fringe pattern. Qian *et al.* [36] proposed deep-learning-enabled geometric constraint and a phase unwrapping method for single-shot absolute 3D shape measurement. For SPU, many deep learning methods have emerged [37,44–48]. Wang *et al.* [37] compared in detail dataset generation methods under deep learning with methods involving deep learning for the SPU method. They recommended using a modified random matrix enlargement (RME) method to generate datasets. It is also demonstrated that both deep learning methods, dRG and dWC can achieve satisfactory results with experiments on severe noise, discontinuities and spectral aliasing. In the study of the deep learning-based TPU method, Yin *et al.* [39] applied deep learning to perform TPU to improve the efficiency of phase unwrapping. Compared with the traditional method, this method improves the performance of the phase unwrapping of the multi-frequency method. But this work was only developed to enhance the multi-frequency TPU method. The multi-wavelength and number-theoretic methods that are also widely used and are still quite sensitive to noise. Therefore, a unified framework is lacking to enhance the performance of these temporal phase unwrapping methods.

Inspired by previous work, we find that a generic neural network can be constructed to implement a variety of TPU methods. The input is a set of wrapped phase maps which is the same as the one used in the conventional TPU method. The high-quality fringe order pattern of the high-frequency phase is output after proper training. By constructing the multi-frequency method phase unwrapping model (DL-MF), the multi-wavelength method phase unwrapping model (DL-MW), and the number-theoretic method phase unwrapping model (DL-NT) based on the principles of the three typical methods proposed, we demonstrate that deep learning techniques can automatically implement TPU through supervised learning. The proposed method can reliably and efficiently recover object pixels under the influence of noise or in the presence of discontinuous surfaces. To validate the proposed method, we recover the absolute phase of various test objects by projecting fringe patterns at different frequencies. Experimental results show that the proposed method can directly and reliably unwrap 64 periods of high-frequency

phases, with the absolute phase of more than 95% of the valid pixels being correctly recovered. Our proposed method can effectively improve the performance of phase unwrapping and is less sensitive to noise than traditional methods. The present method can reduce the phase unwrapping error at least twice under the influence of noise.

The remainder of this paper is organized as follows. Section 2 introduces the basic principles of TPUs, deep learning training data and network architectures. Section 3 describes the experimental validation in detail as well as summarizes the comparative results. Conclusions are drawn in Section 4.

## 2. Principle

### 2.1. Principle of temporal phase unwrapping

In traditional TPU, a series of fringe patterns of different frequencies are projected, and the fringe order is then determined from the wrapped phase distribution. In this work, the distribution of distortion fringes captured by the camera can be expressed as:

$$I_n(x, y) = A(x, y) + B(x, y) \cos \left[ \phi(x, y) - \frac{2\pi n}{N} \right] \quad (1)$$

where  $A(x, y)$  is the average intensity of the sinusoidal fringe,  $B(x, y)$  is the modulation of the sinusoidal fringe,  $n$  is the phase shift index ( $n = 1, 2, \dots, N$ ),  $N$  is the number of periods of the fringe patterns, and  $\phi(x, y)$  is the phase of the object tested. Using least squares, the phase  $\phi(x, y)$  can be obtained by the following equation:

$$\phi(x, y) = \tan^{-1} \frac{\sum_{n=1}^N I_n(x, y) \sin(\frac{2\pi n}{N})}{\sum_{n=1}^N I_n(x, y) \cos(\frac{2\pi n}{N})} \quad (2)$$

Due to the truncation effect of the arctangent function, the obtained phase  $\phi(x, y)$  is wrapped within the range of  $(-\pi, \pi]$ . In order to obtain a continuous distribution of the phase, it is necessary to use a phase unwrapping method to remove the ambiguity in the wrapped phase. The phase unwrapping can be performed by using Eq. (3):

$$\Phi(x, y) = \phi(x, y) + 2\pi k(x, y) \quad (3)$$

where  $\Phi(x, y)$  is the unwrapped phase and  $k(x, y)$  is the fringe order. In order to obtain a highly accurate unwrapped phase distribution, we choose to predict the fringe order  $k(x, y)$ . According to Eq. (3),  $\Phi(x, y)$  can be obtained from a simple linear combination of  $\phi(x, y)$  and  $k(x, y)$ , so the crucial aspect of phase unwrapping is to quickly and accurately obtain  $k(x, y)$  for each pixel in the phase pattern.

We construct three deep learning-based phase unwrapping models with the help of three representative traditional phase unwrapping physical models: the multi-frequency method phase unwrapping model (DL-MF), the multi-wavelength method phase unwrapping model (DL-MW), and the number-theoretic method phase unwrapping model (DL-NT). What these three methods have in common is to unwrap the phase with the help of one (or more) additional wrapped phase patterns with different fringe periods [49,50]. We use the simplest dual-frequency here as an example, and it should be noted that the method in this paper can be extended to three or more frequencies. In the following, the two wrapped phases are denoted as  $\phi_h$  and  $\phi_l$ , and the fringe wavelengths are denoted as  $\lambda_l$  and  $\lambda_h$  ( $\lambda_h < \lambda_l$ , the subscripts  $h$  and  $l$  denote high and low frequencies respectively), and the corresponding unambiguous phases are  $\Phi_h$  and  $\Phi_l$ , and it is not difficult to obtain the following relation:

$$\frac{\Phi_h(x, y)}{\Phi_l(x, y)} = \frac{\lambda_l}{\lambda_h} \quad (4)$$

### 2.1.1. Multi-frequency (hierarchical) temporal phase unwrapping (MF)

The multi-frequency method uses a continuous unambiguous phase as the auxiliary phase and then finds the fringe order of each pixel of the wrapped phase according to the multiplicative relationship between the auxiliary phase and the frequencies corresponding to the wrapped phase to be unwrapped. According to Eq. (3), the relationship between the unambiguous phase and the wrapped phase can be expressed as:

$$\begin{cases} \Phi_h(x, y) = \phi_h(x, y) + 2\pi k_h(x, y) \\ \Phi_l(x, y) = \phi_l(x, y) + 2\pi k_l(x, y) \end{cases} \quad (5)$$

where  $k_h$  and  $k_l$  are the fringe orders corresponding to the high and low frequencies respectively. In the dual-frequency TPU, the low-frequency phase distribution  $\phi_l$  is assisted by the use of a set of single-frequency fringes, so it is not necessary to perform a phase unwrapping of  $\phi_l$ , that is,  $\Phi_l = \phi_l$ , then the fringe order  $k_h$  for each pixel can be determined according to Eqs. (4) and (5):

$$k_h(x, y) = \text{Round} \left[ \frac{(\lambda_l / \lambda_h) \phi_l(x, y) - \phi_h(x, y)}{2\pi} \right] \quad (6)$$

where  $\text{Round} [\ ]$  means obtaining the closest integer value.

### 2.1.2. Multi-wavelength (heterodyne) temporal phase unwrapping (MW)

The basic principle of the multi-wavelength method is to demodulate the phase information from the beat frequency signal generated by the interference of the reference signal and the measurement signal. The basic operation is to obtain an unambiguous phase by making a difference between the wrapped phases of two different frequencies, and then use the same operation as the hierarchical method to unwrap the high-frequency wrapped phase. Generally, this unambiguous phase can be referred to as  $\Phi_{eq}$ , and then  $\Phi_{eq}$  can be resolved by the following expression:

$$\Phi_{eq} = \text{mod}(\phi_h - \phi_l, 2\pi) \quad (7)$$

The wavelengths corresponding to:

$$\lambda_{eq} = \frac{\lambda_l \lambda_h}{\lambda_l - \lambda_h} \quad (8)$$

In order to enable  $\Phi_{eq}$  to be a continuous unambiguous phase, it is generally necessary to ensure that  $\lambda_{eq} \geq W$ ,  $W$  is the resolution of the projector in the direction of phase change. After the unambiguous phase  $\Phi_{eq}$  and the wavelength  $\lambda_{eq}$  are obtained, the fringe order of the high-frequency wrapped phase can be determined by the following equation:

$$k_h(x, y) = \text{Round} \left[ \frac{(\lambda_{eq} / \lambda_h) \Phi_{eq}(x, y) - \phi_h(x, y)}{2\pi} \right] \quad (9)$$

Furthermore, the dual-wavelength TPU approach can be extended to three or more wavelengths, further increasing the equivalent wavelength. Since the reference phase is generated by the wrapped difference of two phase functions, this TPU method is also called the hierarchical method.

### 2.1.3. Number-theoretic temporal phase unwrapping (NT)

The number-theoretic method uses the wavelengths  $\lambda_h$  and  $\lambda_l$  of two sets of sinusoidal fringes of mutual quality to ensure that the wrapped phase pair  $(\phi_h, \phi_l)$  is unique within the pixel range  $[0, LCM(\lambda_h, \lambda_l)]$ , where  $LCM()$  denotes the least common multiple function. After the unique

wrapped phase pair  $(\phi_h, \phi_l)$  is determined, the fringe orders of the two phase patterns  $(k_h, k_l)$  can be determined by searching their wrapped phases. If  $LCM(\lambda_h, \lambda_l)$  (called the unambiguous range) can be exceeded by the lateral resolution of the projection pattern, the phase ambiguity of the whole field can be eliminated. For a projection pattern with  $W \times H$  resolution, the two different wavelengths  $\lambda_h$  and  $\lambda_l$  should be chosen to satisfy the following inequalities to exclude phase ambiguity:

$$LCM(\lambda_h, \lambda_l) \geq W \quad (10)$$

Since the two sets of fringe patterns have different wavelengths ( $\lambda_h$  and  $\lambda_l$ ), their absolute phase patterns also conform to the expression in Eq. (4). To determine the two fringe orders ( $k_h$  and  $k_l$ ) via the two unwrapped phase maps, we rewrite Eq. (4) as:

$$\frac{\Phi_h(x, y)}{\Phi_l(x, y)} = \frac{f_h}{f_l} \quad (11)$$

where  $f_h$  and  $f_l$  denote the fringe spatial frequencies corresponding to high frequency gratings and low frequency gratings. Assuming that the projection is a vertical grating and the lateral resolution of the projector is  $W$ , the relationship between frequency and wavelength is:

$$f = W/\lambda \quad (12)$$

where  $\lambda$  denotes the wavelength of the projected grating and is expressed in pixels.  $f$  denotes the spatial frequency of the projected grating. Combining Eqs. (5) and (11), we can obtain the following relationship:

$$\frac{f_h \phi_l - f_l \phi_h}{2\pi} = k_h f_l - k_l f_h \quad (13)$$

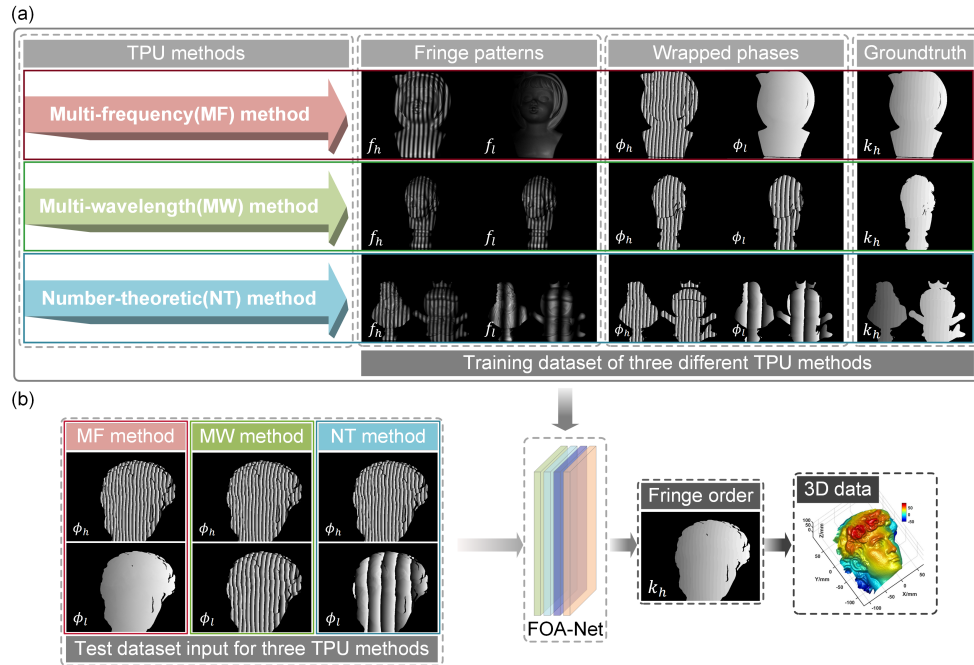
The fringe order  $(k_h, k_l)$  can be determined from a pre-computed look-up table (LUT). When obtaining two phase patterns at a given location, we compute their weighted difference  $(f_h \phi_l - f_l \phi_h) / 2\pi$ , round the values to the nearest integer, and then use the pre-computed LUT to determine the fringe order pair  $(k_h, k_l)$ .

$$(k_h, k_l) = \text{Round} \left[ \frac{f_h \phi_l - f_l \phi_h}{2\pi} \right] \quad (14)$$

## 2.2. Unifying framework based on deep learning

The purpose of a data-driven deep learning network has been to apply a large number of input values (samples) and groundtruth values (targets/labels) to train a model which predicts output values that can be infinitely similar to the groundtruth values. In this work, the supervised learning is used. The objective of supervised learning is to find a mapping function to map the input variables to the output variables. The machine is trained with "labeled" training data, and the machine predicts the output based on that data. Figure 1 shows the flow chart of the present method. We took two wrapped phases of different frequencies as input, used each of the three TPU methods mentioned previously to unwrap the high-frequency wrapped phase to  $k_h(x, y)$  and set it as the label of the network. We followed the image segmentation network structure MultiResUNet [51] to build a Fringe Order Analysis Network (FOA-Net) for fringe order estimation tasks. This network takes U-Net as the basic skeleton, which makes the structure of our network simple and allows the loss of the network to converge faster and more consistently, effectively improving the prediction accuracy of the network. It is important to note that our proposed deep neural network-based approach is capable of learning features of both spatial and temporal phase information, which is beneficial for enhancing the unwrapping performance. The diagram and the data processing process of the proposed network is shown in Fig. 1.

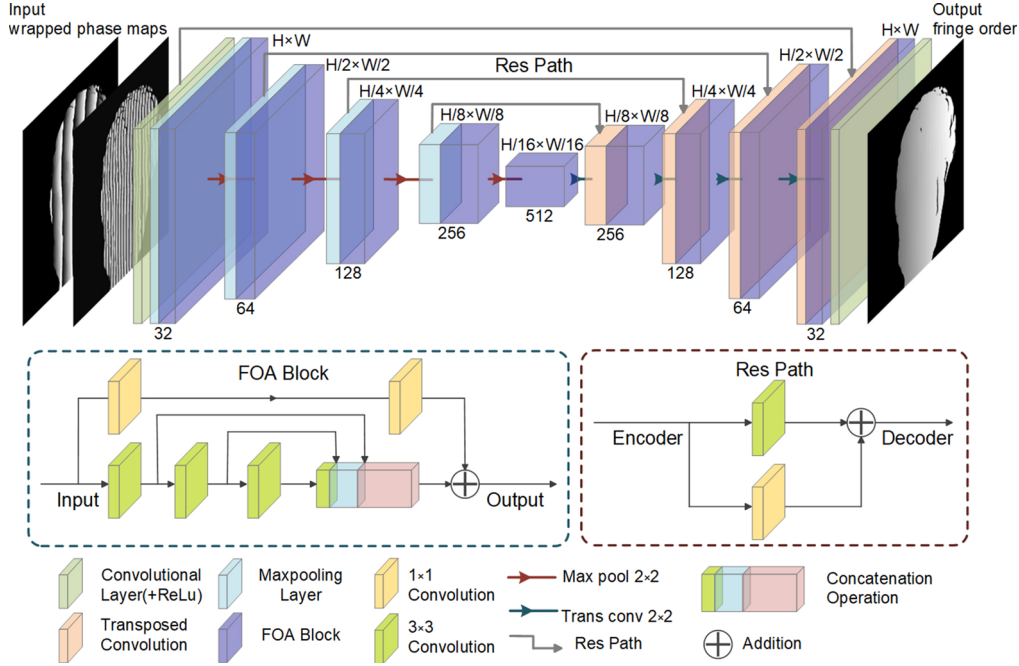
The input data of our proposed network are two wrapped phases of different frequencies (the same as the multi-frequency method based on dual-frequency phases) and the output data of



**Fig. 1.** Data processing process. (a) The preparation of the training dataset that is generated by three TPU methods; (b) TPU process based on deep learning model. The wrapped phase is obtained from the captured three sets of fringe images as the network input, and the fringe order  $k_h$  is output directly by FOA-Net. Finally, the 3D reconstruction results are obtained after the phase-to-height mapping.

the network are fringe order maps of the high-frequency wrapped phase. The basic structure of the FOA-Net is shown in Fig. 2. Similar to MultiResUNet, the FOA-Net replaces the two  $3 \times 3$  convolutions in the classical U-Net with  $3 \times 3$  and  $7 \times 7$  convolutional operations merged in parallel with  $5 \times 5$  convolutional operations, using a multi-resolution idea to replace the traditional convolutional layers. The input tensor of size  $(H, W, 1)$  is for extracting spatial features from different scales. In addition, an MultiRes block is constructed here that instead of using  $3 \times 3$ ,  $5 \times 5$  and  $7 \times 7$  filters in parallel, it decomposed the  $5 \times 5$  and  $7 \times 7$  filters into a series of  $3 \times 3$  filters for the concatenation operation. In the FOA block, we add two  $1 \times 1$  convolutional layers and a residual connection in order to better extract the detailed features in the input. And added residual connection, which we call FOA block. In order to reduce the differences between the encoder-decoder features, instead of simply concatenating the feature maps from the encoder stage to the decoder stage, we first pass them through a chain of convolutional layers with residual connections and then concatenate them with the decoder features. This connection is called ResPath. Specifically,  $3 \times 3$  filters are used in the convolutional layers and  $1 \times 1$  filters in the residual connections. The input feature map of the encoder is connected to the decoder feature map by ResPath. The input tensors are successively processed by a stack of Conv2D layers, Maxpooling2D layers, ConvTranspose2D layers, and the FOA block. Each Conv2D layer represents a convolution operation, which extracts patches from its input feature map and applies the same transformation to all of these patches, producing an output feature map. The role of the MaxPooling2D layer is to aggressively downsample feature maps, consisting of extracting windows from the input feature maps and outputting the max value of each channel. Usually, max pooling is done with  $2 \times 2$  windows and stride 2 to downsample the feature maps by a factor of 2.

Thus, the size of composite image input  $H \times W$  tends to shrink as the network becomes deeper. The ConvTranspose2D layer can be seen as the inverse process of the convolution operation, recovering the image size before convolution depending on the size of the convolution kernel and output.



**Fig. 2.** The FOA-Net network architecture.

The loss function we select in this neural network is mean squared error (MSE), which is used to compare these predictions with the targets and generate a loss value. The sequence of two convolutional layers is replaced with the FOA block introduced in the previous section. For each FOA block, specify a parameter  $V$  that controls the number of filters in the convolutional layers within that block. In order to maintain a comparable relationship between the number of parameters in the original U-Net network and the defined blocks, the value of  $V$  as follows:

$$V = \alpha \times U \quad (15)$$

where  $U$  is the number of filters in the corresponding layer of the U-Net network and  $\alpha$  is used as a scale coefficient to control the number of filters in the convolutional layers in the FOA-block. In the U-Net network, the number of filters along the horizontal direction are 32, 64, 128, 256 and 512, which are also the values of  $U$  in Eq. (15). In order to improve the accuracy of the network output results and without increasing the number of parameters significantly, we choose  $\alpha = 1.67$  according to our experimental results. This also improves the performance of our proposed network without compromising accuracy. This network structure uses three successive convolutional layers with  $V/6$ ,  $V/3$  and  $V/2$  filters. Similar to U-Net, the value of  $V$  is doubled after each pooling or deconvolution operation.

In addition to introducing the FOA blocks, this structure also replaces the ordinary shortcut connections with the proposed ResPath, so that some convolution operations can be applied to the feature mapping propagated from the encoder to the decoder. It is assumed that the strength of the semantic gap between the encoder and decoder feature mappings may decrease as one moves towards the internal shortcut path. Thus, it is also possible to gradually reduce the number

of convolution blocks used along the ResPath and use 4, 3, 2 and 1 convolution blocks along each of the four ResPaths. In addition, in order to account for the number of feature mappings in the encoder-decoder, 32, 64, 128, and 256 filters are used in each of the four ResPath blocks. The Rectified linear unit (ReLU) function is a common activation function that avoids gradient disappearance when propagating backward. Therefore, all convolutional layers used in our proposed network are activated by the ReLU activation function with batch normalization.

In addition, in order to enhance the learning ability of the network, we mask the invalid points of the training data image by setting the appropriate modulation threshold using the modulation function  $B(x, y)$  and the mask function  $Mask(x, y)$ .

$$B(x, y) = \frac{2}{N} \sqrt{M_h(x, y)^2 + D_h(x, y)^2} \quad (16)$$

$$Mask(x, y) = \begin{cases} B(x, y), & B(x, y) > Thr \\ 0, & B(x, y) \leq Thr \end{cases} \quad (17)$$

$$M_h(x, y) = \sum_{n=1}^N I_n^c(x, y) \sin\left(\frac{2\pi n}{N}\right) \quad (18)$$

$$D_h(x, y) = \sum_{n=1}^N I_n^c(x, y) \cos\left(\frac{2\pi n}{N}\right) \quad (19)$$

where  $B(x, y)$  is the modulation intensity of  $I_n^c(x, y)$ ,  $Thr$  denotes a preset threshold used to distinguish the object from the low modulation background, and  $M_h(x, y)$  and  $D_h(x, y)$  are set as the numerator and denominator terms of the inverse tangent function in the wrapped phase  $\phi(x, y)$  in Eq. (2). After thresholding, the valid measurement points labeled with  $Mask(x, y)$  are further used for network training and 3D reconstruction. It is important to note that the threshold  $Thr$  should be adjusted for object surfaces with different reflectance.

### 3. Experiments and results

#### 3.1. Comparison experiments

In order to validate the actual performance of the proposed method, we built a conventional FPP system consisting of a black and white industrial camera (Basler acA640-750 $\mu$ m) and a DLP projector (LightCrafter 4500Pro). To prepare the dataset for the deep neural network, 3-step PS fringe patterns at different frequencies (including 1, 16, 24, 32, 48 and 64) were projected onto multiple object surfaces in sequence and the fringe images were acquired by the camera simultaneously. The wrapped phase maps of these six frequencies were used for training and testing. To be specific, these maps of different frequencies were trained and tested separately instead of all together. Further, the fringe order  $k_h(x, y)$  and the absolute phase of the high-frequency phase  $\Phi_h(x, y)$  are obtained to build the training data set, the validation data set and the test data set. We set the value of the threshold  $Thr$  to 8, which is applicable to most of the measurement scenarios in this work.

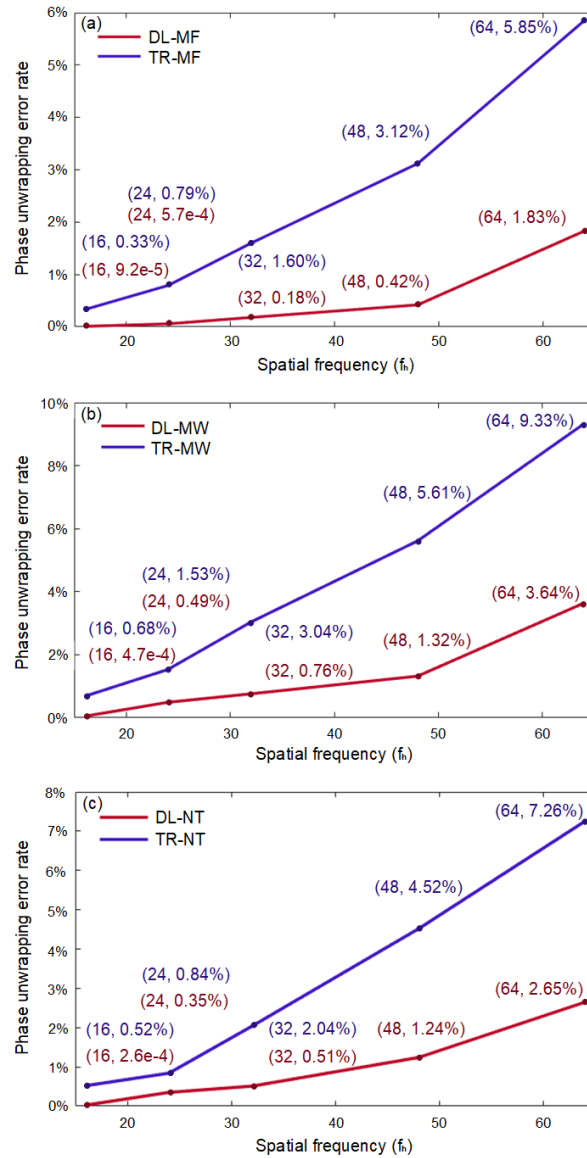
The proposed network is implemented based on the TensorFlow framework (Google) and the training, validation and testing of the network is performed on a GTX Titan graphics card (NVIDIA). In the network configuration, the loss function is set to MSE, the optimizer is *Adam*, the mini-batch size is 2, the training period is set to 200 rounds and the initial model will be learned through the above process. In order to further optimize the model, the network parameters and structure need to be adjusted according to different methods. It takes approximately 12 hours to complete model training using our proposed strategy. And our model achieves 12 fps offline prediction, which is 20% higher than U-Net. The training time for the multi-frequency method

takes 11.7 hours, and the prediction speed is at 11.6 fps. The multi-wavelength method requires 11.8 hours of training and achieves a prediction speed of 11.9 fps. The training time for the number-theoretic method is at 12 hours and the prediction speed is 12.1 fps.

Firstly, we performed phase unwrapping for different high-frequency wraps on the test dataset to quantitatively analyze the phase unwrapping accuracy of DL-TPU and traditional TPU by calculating the average error rate of TPU, which is shown in Fig. 3. The figure shows that for any of the TPU methods, the error rate of phase unwrapping gradually increases as the high-frequency fringe frequency  $f_h$  gradually increases. For the multi-frequency method (Fig. 3(a)), the phase unwrapping error rate in the traditional method (TR-MF) increases from 0.33% to 5.85%. This result proves that due to the presence of non-negligible noise and other error sources in the actual measurement, it is challenging for the traditional method to unwrap the high-frequency phase as it increases. As seen from the graphs we have presented, the correct rate of phase unwrapping under deep learning (DL-MF) has significantly improved. For the multi-wavelength method (Fig. 3(b)), our proposed method (DL-MW) shows similarly better performance through deep learning, with the absolute phase recovered correctly for more than 95% of the valid pixels. When using the number-theoretic method (Fig. 3(c)) for phase unwrapping, the proposed method (DL-NT) can reduce the phase unwrapping error rate of the traditional method (TR-NT) by a factor of 2 to 3. Furthermore, we find that DL-MF shows the smallest error rate among these deep-learning-based methods.

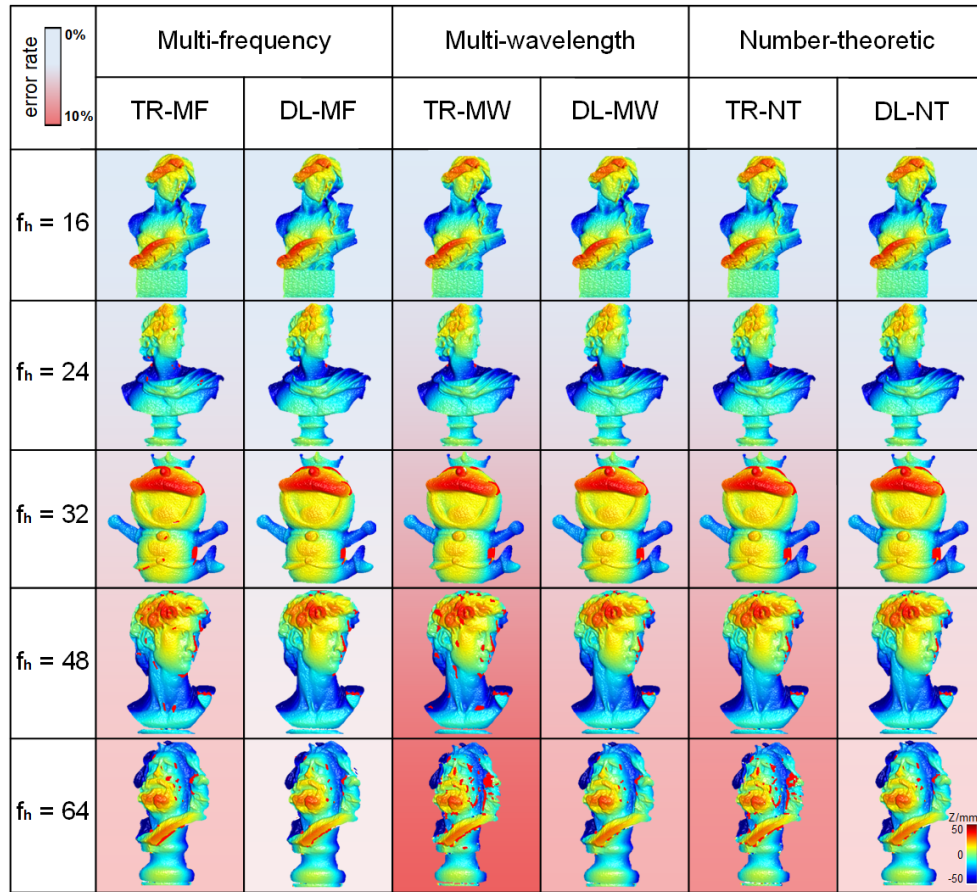
To provide a more intuitive comparison of the actual phase unwrapping performance of the traditional TPU and DL-TPU, Fig. 4 shows the results of 3D measurements for a number of complex models. The corresponding high-frequency fringe frequencies are 16, 24, 32, 48 and 64 for different low-frequency fringe frequencies. It can be seen from Fig. 4 that for these three TPU methods, the phase unwrapping error rate gradually increases with the increase of high frequency. Figure 5 shows the detail areas of the 3D measurement results for each method at a frequency of 48, including hair and skin. From the 3D measurement results, it can be seen that for all three traditional algorithms, the hair and skin of the plaster statue are deficient to different degrees. For the DL-MF, DL-MW and DL-NT, the plaster statue has more complete surface. It can be seen that for different textured surfaces, our methods have better results. Compared with traditional TPU, the quality of the 3D reconstruction with our proposed DL-TPU is significantly improved, and the proposed method can still reliably unwrap the phase even at higher frequencies. It is noteworthy that although we tested only these frequencies, this neural network can also be used for other frequencies, as long as the corresponding wrapped phase maps were seen by the neural network during the training process.

In order to further verify the generalizability of our proposed method, we compared the 3D reconstruction results for different material measurement objects. Figure 6 shows the measurement results of three TPU methods with deep learning for the plaster statue, plastic model (plastic toy) and metal (lithium battery metal base) at the frequency of 24. The first row shows the original images of the measurement objects of three different materials. For the plaster statue, all three methods can recover the majority of the effective pixels, and complex regions of the plaster statue (such as eyes, nose and hair) have been reconstructed correctly. For the plastic model, all three TPU methods also have good results with fewer error points. For metal surfaces, the captured grating images usually have highlights. The phase information at the corresponding position is also often difficult to obtain correctly due to the effects of the highlights. Therefore, it also generate a large number of error points if we perform phase unwrapping directly. To solve this problem, we used the multiple exposure techniques [52] when capturing metal surfaces. We can obtain the correct wrapped phase without interference from highlights by this technique. Then we used the neural network described in this paper for phase unwrapping. This experiment demonstrates that our proposed method is effective for different materials of measurement objects. We measured standard ceramic plates to quantify the 3D measurement accuracy of the proposed

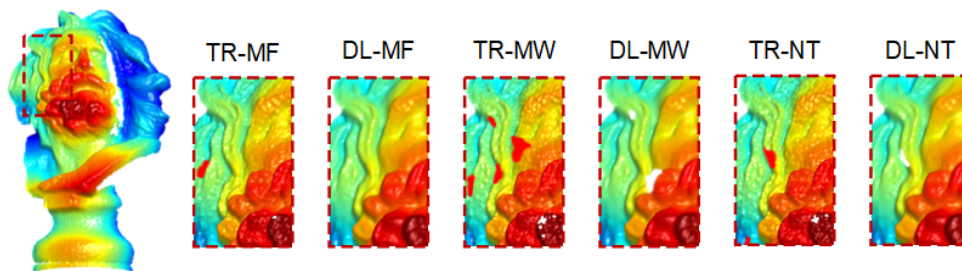


**Fig. 3.** Comparisons between traditional temporal phase unwrapping and deep learning-based temporal phase unwrapping on the test dataset for the average error rate of phase unwrapping for wrapped phases at different high frequencies (e.g.  $f_h = 16, 24, 32, 48$  and  $64$ ). (a) The task of multi-frequency phase unwrapping; (b) The task of multi-wavelength phase unwrapping; (c) The task of number-theoretic phase unwrapping.

method. In our experiments, we projected 16, 24, 32, 48 and 64 period fringe patterns onto a standard ceramic plate sequentially and used the multi-frequency method, multi-wavelength method and number-theoretic method for phase unwrapping. Table 1 shows the RMS values of the standard ceramic plate at different frequencies for traditional and proposed methods. It can be seen from Table 1 that for the TR-MF, the RMS value decreases from  $98.2\mu\text{m}$  to  $53.1\mu\text{m}$  as the frequency increases from 16 to 64. The corresponding RMS value in DL-MF can be reduced from  $81.9\mu\text{m}$  to  $44.2\mu\text{m}$ . For the multi-wavelength method, the RMS value changes from

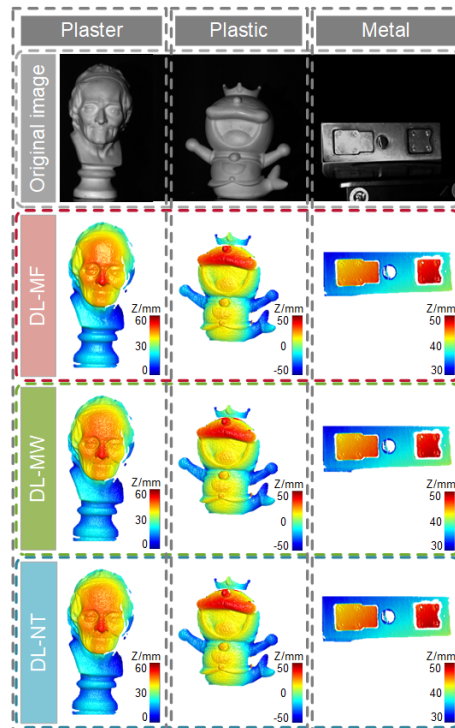


**Fig. 4.** Comparison of 3D reconstruction results for complex models with traditional TPU and the proposed deep learning based TPU, where high-frequency fringe frequencies are 16, 24, 32, 48 and 64 respectively.



**Fig. 5.** Detail results of 3D measurements.

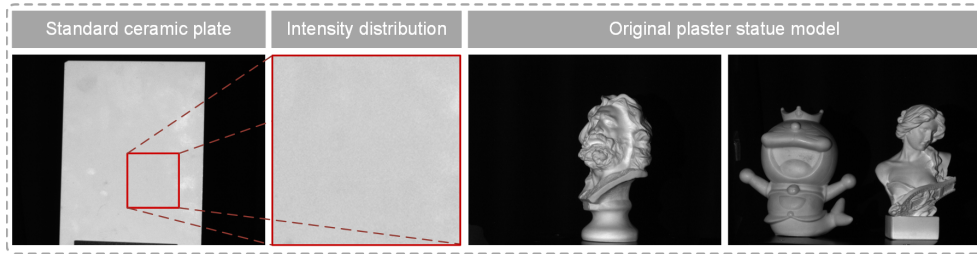
106.6 $\mu\text{m}$  to 63.7 $\mu\text{m}$  with increasing frequency under the traditional method. By comparison, the RMS value of DL-MW was reduced from 97.7 $\mu\text{m}$  to 58.9 $\mu\text{m}$ . For TR-NT, the RMS value changes from 101.2 $\mu\text{m}$  to 59.4 $\mu\text{m}$  after the frequency becomes higher. And the RMS value of the standard ceramic plate under deep learning changes from 91.4 $\mu\text{m}$  to 51.4 $\mu\text{m}$ . This experiment demonstrates that for different TPU methods, the accuracy of our proposed method is better than that of the traditional method.



**Fig. 6.** 3D measurement results for different materials.

**Table 1.** The RMS values of standard ceramic plates at different frequencies for traditional and proposed methods.

| RMS/ $\mu\text{m}$ | TR-MF | TR-MW | TR-NT | DL-MF | DL-MW | DL-NT |
|--------------------|-------|-------|-------|-------|-------|-------|
| $f = 16$           | 98.2  | 106.6 | 101.2 | 81.9  | 97.7  | 91.4  |
| $f = 24$           | 86.4  | 94.2  | 88.2  | 67.2  | 86.1  | 77.8  |
| $f = 32$           | 73.4  | 82.4  | 78.1  | 56.7  | 74.4  | 66.3  |
| $f = 48$           | 62.2  | 72.1  | 67.8  | 49.1  | 65.1  | 57.2  |
| $f = 64$           | 53.1  | 63.7  | 59.4  | 44.2  | 58.9  | 51.4  |



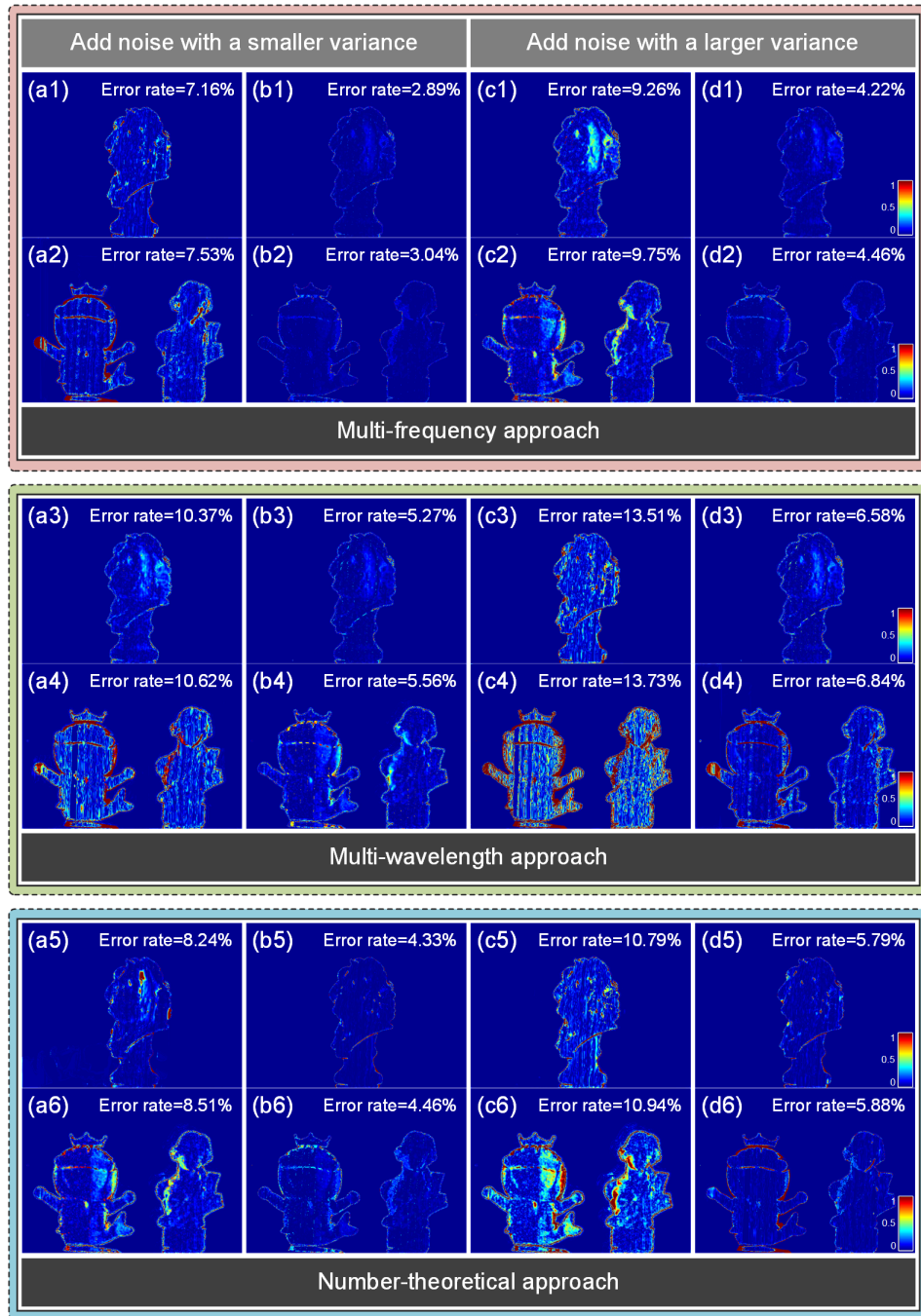
**Fig. 7.** Captured standard ceramic plate image and plaster statue models by camera.

### 3.2. Performance analysis under the influence of different noises

In this work, the fringe distribution captured by the camera can be represented by Eq. (1). In fact, the images are usually captured with sensor noise. We assume a Gaussian distributed additive noise in captured fringe patterns. This assumption is valid for image sensors in which thermal or shot noise is the main noise type. In the next experiments, we will further verify the noise immunity performance of the three TPU methods with deep learning.

In order to obtain the noise variance of the camera in the experiment, we measured the single-vision system used in the experiment. We took a standard ceramic plate as the measurement object, and the projector projected a fixed light source onto the plate. During the experiment, we only modified the camera's exposure time without changing the brightness of the light source and acquired 200 sets of plate images as shown in Fig. 7. After the acquired images were averaged and the average image was subtracted from each image separately, the resulting noise was averaged again to obtain a final noise variance of  $\sigma^2 = 2.4629$ . To further validate the performance of the proposed method for different levels of noise, we added twice and three times the noise of the original to the original fringe pattern,  $2\sigma$  and  $3\sigma$ , respectively.  $\phi_h(x, y)$  and  $\phi_l(x, y)$  were calculated by the 3-step PS algorithm, and the fringe order  $k_h(x, y)$  was obtained by our proposed three TPU methods (multi-frequency method, multi-wavelength method, and number-theoretic method) were obtained separately. In order to compare the reliability of the phase unwrapping, a 6-step PS algorithm and a 3-frequency phase-based multi-frequency method (frequencies are 1, 8 and 64) were used to obtain high-quality fringe order as reference orders. By comparing the fringe order and the reference order based on the different methods, the corresponding error rate of the phase unwrapping can be accurately calculated.

Firstly, we chose two static scenes never seen in our network before, including separate plaster statue objects and a combined plaster statue model as shown in Fig. 7. These scenes involved a single object with a continuous complex surface shape and a combination of multiple objects with isolated surfaces. After unwrapping the phase for the selected scene, we fed them directly into the trained neural network to predict the high-frequency fringe order corresponding to the input wrapped phases, and the predictions were compared with groundtruth to obtain the fringe order error plot shown in Fig. 8. Figure 8(a1)-(a6) show the phase error of the traditional method with  $2\sigma$  noise added, and Fig. 8(b1)-(b6) show the phase error of the proposed method with the corresponding noise. For the multi-frequency method, the use of a deep neural network can reduce the error rate of phase unwrapping for a single object from 7.16% to 2.89%. Using the multi-wavelength method reduces the error rate for a single object from 10.37% to 5.27%. We can see that the deep learning approach is successful in eliminating phase errors and that the error rate for a single object can be reduced from 8.24% to 4.33% under the number-theoretic method when applying our approach. It still shows a good performance in the multi-object combination scenario shown in the second row of each method. Figure 8(c1)-(c6) show the phase error of the traditional method with  $3\sigma$  noise added, and Fig. 8(d1)-(d6) show the phase error of our method with the corresponding noise. The error rate for individual objects in the phase expansion based on the multi-frequency method can be reduced from 9.26% to 4.22%, even under the influence



**Fig. 8.** Comparison of the phase error between the traditional phase unwrapping method and the proposed method at the high frequency of 48. (a1)-(a6), (c1)-(c6) Phase errors calculated by the traditional method; (b1)-(b6), (d1)-(d6) Phase errors calculated by the proposed method.

of  $3\sigma$  noise. In addition, for the multi-wavelength and number-theoretic methods under the same noise conditions, the error rate of phase unwrapping for a single object is reduced from

13.51% and 10.79% to 6.58% and 5.79% respectively. Even in more complex scenes, more than 90% of the pixels can be recovered. It can be seen from the experimental results that the fringe order error increases as the noise increases, but our method always yields smaller error results compared to traditional methods.

In order to analyze the method quantitatively, 150 completely different sets of scenes were selected for validation with the addition of noise. We calculated the error rate of phase unwrapping in each of these scenes with different levels of noise added, and the average error rate is shown in Table 2. The extensive validation results show that the error rate of phase unwrapping using the traditional multi-frequency method is 7.28% and 9.54% for the case of adding less and more noise respectively to the fringe patterns we have acquired. For our method, the error rate of phase unwrapping was reduced to 3.32% and 4.46%, respectively. For the traditional multi-wavelength method (TR-MW), the error rate of phase unwrapping was reduced from 10.66% and 13.79% to 5.42% and 6.99% respectively after our network processing. For the number-theoretic method, the error rate was reduced from 8.71% and 11.02% to 4.47% and 5.91%, respectively. Compared with the traditional phase unwrapping method, our method can reduce the error rate by nearly half. The deep learning-based phase unwrapping method has better performance under all three different phase unwrapping methods (multi-frequency method, multi-wavelength method, and number-theoretic method). We can see that noise has different effects on these three TPU at a frequency of 48 and the phase unwrapping error rate increases as the noise increases. In contrast, the multi-frequency method (DL-MF) has the lowest influence by noise. For the traditional TPU, the low-quality fringe order leads to severe phase errors and therefore a larger phase unwrapping error rate. For our proposed DL-TPU, the phase unwrapping eliminates phase errors close to  $2\pi$  between adjacent pixels, allowing for a complete and unambiguous phase. In the case of increased noise, our method also overcomes the noise influence better. The reason why the proposed work can suppress noise may lie in the fact that a similar process like image denoising is embedded in the training process. People usually use filters to remove the effect of noise in image processing. The image filtering is essentially a convolutional operation. Here, the structure of our proposed neural network is based on convolutional neural networks. Therefore, the process of noise removal tends to be embedded during training when a convolutional layers extract information from training data (convolve with input data).

**Table 2. The error rate of phase unwrapping with different levels of noise added for the traditional and proposed methods.**

| Error rate | Multi-frequency |           | Multi-wavelength |           | Number-theoretic |           |
|------------|-----------------|-----------|------------------|-----------|------------------|-----------|
|            | $2\sigma$       | $3\sigma$ | $2\sigma$        | $3\sigma$ | $2\sigma$        | $3\sigma$ |
| DL-TPU     | 3.32%           | 4.46%     | 5.42%            | 6.99%     | 4.47%            | 5.91%     |
| TR-TPU     | 7.28%           | 9.54%     | 10.66%           | 13.79%    | 8.71%            | 11.02%    |

#### 4. Conclusions

This paper focuses on three mainstream temporal phase unwrapping methods, including multi-frequency method, multi-wavelength method, and number-theoretic method. Conventionally, these temporal phase unwrapping methods are seldom considered in a unified framework. In this work, we have proposed a generic temporal phase unwrapping neural network framework to enhance the performance of traditional temporal phase unwrapping methods. Experiments demonstrate that trained deep neural networks can significantly improve the accuracy of phase unwrapping. This deep learning-based method, using different frequencies of wrapped phases as input to predict accurate fringe orders, retrieves the complex or discontinuous surface of objects. Furthermore, we have tested the performance of these methods under different noise conditions to verify the phase unwrapping capability of the proposed DL-TPU method under different cases. Experimental results have shown that compared with the traditional TPU method,

the proposed method can effectively suppress the effect of noise on phase unwrapping results. After we have increased the noise to twice and three times the original noise, our method has better performance in both cases. We believe this method will demonstrate significant potential for robust and accurate phase unwrapping and 3D measurements.

**Funding.** National Key Research and Development Program of China (2022YFB2804600, 2022YFB2804603, 2022YFB2804605); National Natural Science Foundation of China (62005121, 62075096, U21B2033); “333 Engineering” Research Project of Jiangsu Province (BRA2016407); Leading Technology of Jiangsu Basic Research Plan (BK20192003); National Major Scientific Instrument Development Project (62227818); Jiangsu Provincial “One belt and one road” innovation cooperation project (BZ2020007); Fundamental Research Funds for the Central Universities (30919011222, 30920032101, 30921011208); Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX22\_0412); Open Research Fund of Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense (JSGP202105, JSGP202201).

**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. G. M. Frank Chen and S. Mumin, “Overview of three-dimensional shape measurement using optical methods,” *Opt. Eng.* **39**(1), 10–21 (2000).
2. F. Blais, “Review of 20 years of range sensor development,” *J. Electron. Imaging* **13**(1), 231–243 (2004).
3. F. Deng, C. Liu, W. Sze, J. Deng, K. S. Fung, W. Leung, and E. Y. Lam, “Illumination-invariant phase-shifting algorithm for three-dimensional profilometry of a moving object,” *Opt. Eng.* **51**(9), 097001 (2012).
4. F. Deng, W. F. Sze, J. Deng, K. S. Fung, W. Leung, and E. Y. Lam, “Regularized multiframe phase-shifting algorithm for three-dimensional profilometry,” *Appl. Opt.* **51**(1), 33–42 (2012).
5. Y. Shu, R. Chung, Z. Tan, J. Cheng, E. Y. Lam, K. S. Fung, and F. Wang, “Projection optics design for tilted projection of fringe patterns,” *Opt. Eng.* **47**(5), 053002 (2008).
6. C. Zuo, S. Feng, L. Huang, T. Tao, W. Yin, and Q. Chen, “Phase shifting algorithms for fringe projection profilometry: A review,” *Opt. Lasers Eng.* **109**, 23–59 (2018).
7. B. Pan, Q. Kemao, L. Huang, and A. Asundi, “Phase error analysis and compensation for nonsinusoidal waveforms in phase-shifting digital fringe projection profilometry,” *Opt. Lett.* **34**(4), 416–418 (2009).
8. M. Takeda and K. Mutoh, “Fourier transform profilometry for the automatic measurement of 3-d object shapes,” *Appl. Opt.* **22**(24), 3977–3982 (1983).
9. Q. Kemao, “Two-dimensional windowed fourier transform for fringe pattern analysis: principles, applications and implementations,” *Opt. Lasers Eng.* **45**(2), 304–317 (2007).
10. L. Huang, Q. Kemao, B. Pan, and A. K. Asundi, “Comparison of fourier transform, windowed fourier transform, and wavelet transform methods for phase extraction from a single fringe pattern in fringe projection profilometry,” *Opt. Lasers Eng.* **48**(2), 141–148 (2010).
11. X. Su and W. Chen, “Reliability-guided phase unwrapping algorithm: a review,” *Opt. Lasers Eng.* **42**(3), 245–261 (2004).
12. T. J. Flynn, “Two-dimensional phase unwrapping with minimum weighted discontinuity,” *J. Opt. Soc. Am. A* **14**(10), 2692–2701 (1997).
13. C. Zuo, L. Huang, M. Zhang, Q. Chen, and A. Asundi, “Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review,” *Opt. Lasers Eng.* **85**, 84–103 (2016).
14. M. D. Pritt, “Phase unwrapping by means of multigrid techniques for interferometric sar,” *IEEE Trans. Geosci. Remote Sens.* **34**(3), 728–738 (1996).
15. K. Creath, “Step height measurement using two-wavelength phase-shifting interferometry,” *Appl. Opt.* **26**(14), 2810–2816 (1987).
16. H. O. Saldner and J. M. Huntley, “Temporal phase unwrapping: application to surface profiling of discontinuous objects,” *Appl. Opt.* **36**(13), 2770–2775 (1997).
17. L. Huang and A. K. Asundi, “Phase invalidity identification framework with the temporal phase unwrapping method,” *Meas. Sci. Technol.* **22**(3), 035304 (2011).
18. J. M. Huntley and H. Saldner, “Temporal phase-unwrapping algorithm for automated interferogram analysis,” *Appl. Opt.* **32**(17), 3047–3052 (1993).
19. R. A. Martinez-Celorio, A. Davila, G. H. Kaufmann, and G. Mendiola, “Extension of the displacement measurement range for electronic speckle-shearing pattern interferometry using carrier fringes and a temporal-phase-unwrapping method,” *Opt. Eng.* **39**(3), 751–757 (2000).
20. J. Tian, X. Peng, and X. Zhao, “A generalized temporal phase unwrapping algorithm for three-dimensional profilometry,” *Opt. Lasers Eng.* **46**(4), 336–342 (2008).
21. G. Pedrini, I. Alexeenko, W. Osten, and H. J. Tiziani, “Temporal phase unwrapping of digital hologram sequences,” *Appl. Opt.* **42**(29), 5846–5854 (2003).

22. H. Zhao, W. Chen, and Y. Tan, "Phase-unwrapping algorithm for the measurement of three-dimensional object shapes," *Appl. Opt.* **33**(20), 4497–4500 (1994).
23. L. Kinell and M. Sjö Dahl, "Robustness of reduced temporal phase unwrapping in the measurement of shape," *Appl. Opt.* **40**(14), 2297–2303 (2001).
24. X. Peng, Z. Yang, and H. Niu, "Multi-resolution reconstruction of 3-d image with modified temporal unwrapping algorithm," *Opt. Commun.* **224**(1-3), 35–44 (2003).
25. Y.-Y. Cheng and J. C. Wyant, "Two-wavelength phase shifting interferometry," *Appl. Opt.* **23**(24), 4539–4543 (1984).
26. J. Wyant, "Testing aspherics using two-wavelength holography," *Appl. Opt.* **10**(9), 2113–2118 (1971).
27. A. Alcock and S. Ramsden, "Two wavelength interferometry of a laser-induced spark in air," *Appl. Phys. Lett.* **8**(8), 187–188 (1966).
28. C. Polhemus, "Two-wavelength interferometry," *Appl. Opt.* **12**(9), 2071–2074 (1973).
29. R. Dändliker, R. Thalmann, and D. Prongué, "Two-wavelength laser interferometry using superheterodyne detection," *Opt. Lett.* **13**(5), 339–341 (1988).
30. V. Gushov and Y. N. Solodkin, "Automatic processing of fringe patterns in integer interferometers," *Opt. Lasers Eng.* **14**(4-5), 311–324 (1991).
31. J. Burke, T. Bothe, W. Osten, and C. F. Hess, "Reverse engineering by fringe projection," in *Interferometry XI: Applications*, vol. 4778 (SPIE, 2002), pp. 312–324.
32. M. Takeda, Q. Gu, M. Kinoshita, H. Takai, and Y. Takahashi, "Frequency-multiplex fourier-transform profilometry: a single-shot three-dimensional shape measurement of objects with large height discontinuities and/or surface isolations," *Appl. Opt.* **36**(22), 5347–5354 (1997).
33. J. Zhong and M. Wang, "Phase unwrapping by lookup table method: application to phase map with singular points," *Opt. Eng.* **38**(12), 2075–2080 (1999).
34. C. E. Towers, D. P. Towers, and J. D. Jones, "Time efficient chinese remainder theorem algorithm for full-field fringe phase analysis in multi-wavelength interferometry," *Opt. Express* **12**(6), 1136–1143 (2004).
35. S. Feng, Q. Chen, G. Gu, T. Tao, L. Zhang, Y. Hu, W. Yin, and C. Zuo, "Fringe pattern analysis using deep learning," *Adv. Photonics* **1**(02), 1 (2019).
36. J. Qian, S. Feng, T. Tao, Y. Hu, Y. Li, Q. Chen, and C. Zuo, "Deep-learning-enabled geometric constraints and phase unwrapping for single-shot absolute 3d shape measurement," *APL Photonics* **5**(4), 046105 (2020).
37. K. Wang, Q. Kemao, J. Di, and J. Zhao, "Deep learning spatial phase unwrapping: a comparative review," *Adv. Photonics Nexus* **1**(01), 014001 (2022).
38. Z. Ren, Z. Xu, and E. Y. Lam, "Learning-based nonparametric autofocusing for digital holography," *Optica* **5**(4), 337–344 (2018).
39. W. Yin, Q. Chen, S. Feng, T. Tao, L. Huang, M. Trusiak, A. Asundi, and C. Zuo, "Temporal phase unwrapping using deep learning," *Sci. Rep.* **9**(1), 20175 (2019).
40. C. Zuo, J. Qian, S. Feng, W. Yin, Y. Li, P. Fan, J. Han, K. Qian, and Q. Chen, "Deep learning in optical metrology: a review," *Light: Sci. Appl.* **11**(1), 1–54 (2022).
41. Y. Li, J. Qian, S. Feng, Q. Chen, and C. Zuo, "Deep-learning-enabled dual-frequency composite fringe projection profilometry for single-shot absolute 3d shape measurement," *Opto-Electron. Adv.* **5**(5), 210021 (2022).
42. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**(8), 921–943 (2019).
43. S. Feng, C. Zuo, Y. Hu, Y. Li, and Q. Chen, "Deep-learning-based fringe-pattern analysis with uncertainty estimation," *Optica* **8**(12), 1507–1510 (2021).
44. G. Dardikman, N. A. Turko, and N. T. Shaked, "Deep learning approaches for unwrapping phase images with steep spatial gradients: a simulation," in *2018 IEEE International Conference on the Science of Electrical Engineering in Israel (ICSEE)*, (IEEE, 2018), pp. 1–4.
45. J. J. He, C. Sandino, D. Zeng, S. Vasanawala, and J. Cheng, "Deep spatiotemporal phase unwrapping of phase-contrast mri data," in *Proceedings of the 27th ISMRM Annual Meeting & Exhibition, Montréal, QC, Canada*, (2019), pp. 11–16.
46. M. V. Perera and A. De Silva, "A joint convolutional and spatial quad-directional lstm network for phase unwrapping," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (IEEE, 2021), pp. 4055–4059.
47. K. Wang, Y. Li, Q. Kemao, J. Di, and J. Zhao, "One-step robust deep learning phase unwrapping," *Opt. Express* **27**(10), 15100–15115 (2019).
48. G. Spoorthi, S. Gorthi, and R. K. S. S. Gorthi, "Phasenet: A deep convolutional neural network for two-dimensional phase unwrapping," *IEEE Signal Process. Lett.* **26**(1), 54–58 (2019).
49. V. Srinivasan, H.-C. Liu, and M. Halioua, "Automated phase-measuring profilometry of 3-d diffuse objects," *Appl. Opt.* **23**(18), 3105–3108 (1984).
50. X.-Y. Su, G. Von Bally, and D. Vukicevic, "Phase-stepping grating profilometry: utilization of intensity modulation analysis in complex objects evaluation," *Opt. Commun.* **98**(1-3), 141–150 (1993).
51. N. Ibtehaz and M. S. Rahman, "Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation," *Neural networks* **121**, 74–87 (2020).
52. S. Feng, L. Zhang, C. Zuo, T. Tao, Q. Chen, and G. Gu, "High dynamic range 3d measurements with fringe projection profilometry: a review," *Meas. Sci. Technol.* **29**(12), 122001 (2018).