



Fringe phase extraction with one step by using deep learning

WEIHAO CHENG,^{1,2,3} ID YUNYUN CHEN,^{1,2,3,4,*} ZHAOLOU CAO,^{1,2,3} AND WENZHUO XIE^{1,2,3}¹*State Key Laboratory Cultivation Base of Atmospheric Optoelectronic Detection and Information Fusion, Nanjing University of Information Science & Technology, Nanjing 210044, China*²*Jiangsu Engineering Research Center for Intelligent Optoelectronic Sensing Technology of Atmosphere, Nanjing University of Information Science & Technology, Nanjing 210044, China*³*Jiangsu International Joint Laboratory for Meteorological Photonics and Optoelectronic Detection, Nanjing University of Information Science & Technology, Nanjing 210044, China*⁴*Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (CICAET), Nanjing University of Information Science & Technology, Nanjing 210044, China***yunqq321@sina.cn**Received 18 April 2025; revised 5 June 2025; accepted 5 June 2025; posted 5 June 2025; published 27 June 2025*

Phase extraction from fringes is crucial for various optical measurement techniques. In this paper, we propose a deep learning-based phase extraction (DLPE) method to directly extract the true phase of optical fringes in a single step without the need for phase unwrapping, leveraging the advantages of deep learning in intelligent image perception. The proposed method's validity is confirmed by experiments with simulated fringes and application in extracting phase information under a real flow field using moiré deflectometry. The results clearly demonstrate that our method achieves excellent performance, characterized by high precision and structural similarity. This work eliminates the need for intermediate processing of the wrapped phase, thereby breaking the conventional two-step methods for optical fringe phase extraction. Consequently, it provides an important reference for the intelligent and efficient phase extraction from optical fringes. © 2025 Optica Publishing Group. All rights, including for text and data mining (TDM), Artificial Intelligence (AI) training, and similar technologies, are reserved.

<https://doi.org/10.1364/AO.565597>

1. INTRODUCTION

Optical measurement technology is a classical approach for achieving precise measurement of various objects by extracting information modulated in laser light. It has been widely applied in measuring displacement, strain, surface profiles, and refractive index. Examples include interferometry [1], digital holography [2], fringe projection profiling [3], and moiré deflectometry, among others [4,5]. The critical aspect of these techniques is the phase modulation by periodic optical fringes. Therefore, the accurate phase extraction from these fringes is a fundamental and essential step. Traditionally, extracting phase information from optical fringes involves two main steps: fringe analysis and phase unwrapping [6,7].

Conventional fringe analysis methods can generally be broadly categorized into two types based on their underlying principles: phase-shifting methods [8] and spatial phase demodulation methods [9]. Phase-shifting methods extract phase information from multiple fringe images [10,11]; they place stringent requirements on the stability of the experimental environment [12]. Consequently, it is unsuitable for technologies like moiré deflectometry which measures changing objects such as various flow fields. Spatial phase demodulation methods could extract phase information from only one fringe,

offering advantages of low cost and ease of data acquisition. These methods include Fourier analysis [13], windowed Fourier analysis [14], and wavelet analysis [15,16]. They are widely used in technologies for measuring dynamic objects, such as moiré deflectometry.

Traditional phase-unwrapping methods are generally classified into two categories based on the spatiotemporal processing dimension: temporal and spatial phase-unwrapping methods [17]. Temporal phase-unwrapping methods extract the true phase from multiple fringes with varying frequencies [18–20]. In contrast, spatial phase-unwrapping methods require only one fringe to retrieve the true phase. Therefore, spatial phase-unwrapping methods are more suitable for measuring dynamic objects than temporal phase-unwrapping methods, particularly in moiré deflectometry for complex flow field measurements [21]. Spatial phase-unwrapping methods are further categorized into path-following and path-independent approaches [9]. Compared to path-following methods, which aim to circumvent noisy or invalid points by identifying an accurate integration path but often result in defects and reduced robustness to noise and other complexities [22,23], path-independent methods achieve phase unwrapping by optimizing a global function over the entire wrapped phase distribution, which

typically offers superior noise tolerance. Despite decades of research, precisely extracting phase information from a single fringe remains a critical and unresolved challenge.

In recent years, deep learning has sparked a new wave of advancements in artificial intelligence technologies [24–26], significantly driving the intelligent development of various image processing domains [27–29]. In the field of optics, deep learning has also found extensive applications [30–33]. Notably, it has provided innovative solutions for optical fringe analysis and phase-unwrapping problems [34–36]. Recently, we proposed a deep learning-based moiré fringe analysis (DLMFA) method to predict the real and imaginary components of moiré fringe information and enable comparative analysis under different carrier frequencies [37]. In terms of phase unwrapping, Jin *et al.* [38] are pioneers in leveraging deep learning to solve ill-posed inverse problems in imaging. This breakthrough has led to the rapid emergence of numerous spatial phase-unwrapping methods based on deep learning in recent years, including regression-based and wrapped-counting methods [39–41].

However, as mentioned before, existing phase information extraction methods for optical fringe are all divided into two steps: fringe analysis and phase unwrapping. This division significantly increases computational complexity and time costs. This paper proposes a method called deep learning-based phase extraction (DLPE) to obtain the true phase in a single step without the need for phase unwrapping. This approach eliminates the requirement for intermediate processing of the wrapped phase, thereby breaking away from conventional two-step methods for optical fringes' phase extraction. First, the advantages of the DLPE over traditional methods are validated through simulated fringes. Subsequently, the optimal network structure is determined by evaluating various down-sampling steps and convolutional kernel channel configurations. Finally, the DLPE is applied to extract real flow field phase information based on moiré deflectometry. In a word, our work will provide an important reference for the intelligent and efficient phase extraction from optical fringes.

2. PRINCIPLE AND METHOD

A. Phase Extraction from Optical Fringes by Traditional Methods

Traditional phase extraction techniques involve retrieving the true phase from optical fringes through a two-step process: fringe analysis and phase unwrapping. The most commonly used fringes for modulating phase information are horizontal fringes, with a typical horizontal optical fringe being expressed as [12]

$$I(x, y) = A(x, y) + B(x, y) \cos[2\pi f y + \phi(x, y)] \quad (1)$$

where $I(x, y)$ denotes the intensity distribution of optical fringe, $A(x, y)$ represents the background intensity, $B(x, y)$ indicates the modulation amplitude, f denotes the carrier frequency, and $\phi(x, y)$ is the desired phase information, x and y represent the pixel coordinates.

According to Eq. (1), the wrapped phase $\Phi(x, y)$ could be calculated by the inverse tangent function after removing $A(x, y)$, by

$$\begin{aligned} \Phi(x, y) &= \arctan \frac{\frac{s}{2} B(x, y) \sin[2\pi f y + \phi(x, y)]}{\frac{s}{2} B(x, y) \cos[2\pi f y + \phi(x, y)]} \\ &\quad - \arctan \frac{\frac{s}{2} B(x, y) \sin(2\pi f y)}{\frac{s}{2} B(x, y) \cos(2\pi f y)}, \end{aligned} \quad (2)$$

where s is the number of needed fringes in fringe analysis (e.g., $s = 1$ in spatial phase demodulation methods, $s = N$ in N -step phase-shifting methods). Affected by the range of the inverse tangent function, the phase values of $\Phi(x, y)$ are limited to $[-\pi, \pi]$. Therefore, phase unwrapping is essential to obtain the true phase $\phi(x, y)$, as follows:

$$\phi(x, y) = UW\{\Phi(x, y)\}, \quad (3)$$

where $UW\{\cdot\}$ represents phase unwrapping through extracting the true phase $\phi(x, y)$ from the wrapped phase $\Phi(x, y)$.

Finally, the true phase information of optical fringes is extracted by fringe analysis [Eq. (2)] and phase unwrapping [Eq. (3)].

B. Datasets Production of the DLPE based on Moiré Deflectometry

To generate the dataset of the DLPE, moiré deflectometry is taken as the scene to extract the phase information from moiré fringes to get the label data of dataset. The spatial phase demodulation method exemplified by wavelet analysis and the spatial phase-unwrapping method exemplified by the quality-guide method are employed as the theoretical basis for generating datasets. The experimental setup of the moiré deflectometry system is schematically illustrated in Fig. 1. The wavelength and output power of the laser serving as the probing light are 532 nm and 400 mW. The beam expansion and collimation system consisted of an expanding lens with a focal length of 3 mm and a diameter of 4 mm, and a collimation lens with a focal length of 300 mm and a diameter of 50 mm. This system produced a collimated beam that passed through the measurement region containing the measured flow field. The key components of the moiré deflectometry system are two Ronchi gratings, each with an effective size of 50 mm × 50 mm and a grating constant of 0.02 mm. A subsequent 4-f system is employed to filter the desired diffraction order, producing a moiré fringe pattern on the screen. These fringes are captured using a CCD camera, specifically the GUPPY PRO F125C model.

The moiré fringe containing the phase information can be captured on the optical screen, and its intensity can be specified as [21]

$$I(x, y) = a \left[1 + \cos \left(\frac{4\pi}{d} y \sin \frac{\alpha}{2} + \phi(x, y) \right) \right], \quad (4)$$

where a is the background intensity of the moiré fringe, d represents the grating constant, and α is the angle between the two gratings.

Taking the Morlet wavelet $\psi(x)$ as an example, the one-dimensional wavelet transformation $WI_x(m, n)$ of Eq. (4) in the y direction can be described as

$$\psi(x) = \frac{1}{\pi^{1/4}} \left(\frac{2\pi}{\gamma} \right)^{1/2} \exp \left[-\frac{(2\pi/\gamma)^2 x^2}{2} + i 2\pi x \right], \quad (5)$$

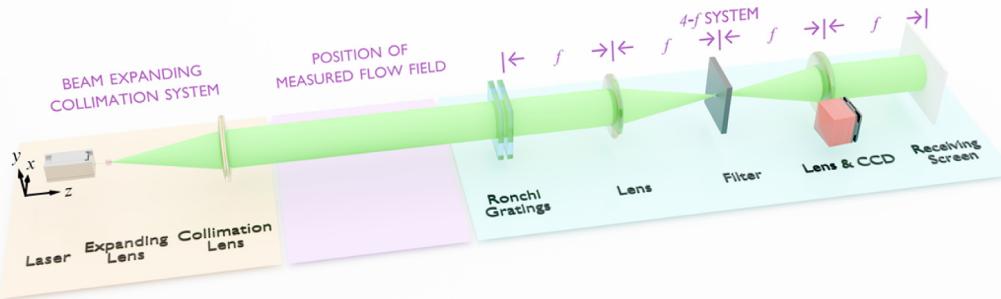


Fig. 1. Experimental setup of moiré deflectometry.

$$WI_x(m, n) = \frac{1}{\sqrt{m}} \int_{-\infty}^{\infty} I(x, y) \psi^* \left(\frac{y-n}{m} \right) dy, \quad (6)$$

where m represents the scale parameter, n denotes the position parameter, and $\psi = \pi \sqrt{2/\ln 2}$.

In further, the wrapped phase could be obtained by

$$\Phi(x, y) = \arctan \left\{ \frac{\text{Im}[WI_x(m_{r,n}, n)]}{\text{Re}[WI_x(m_{r,n}, n)]} \right\}, \quad (7)$$

where $m_{r,n}$ is the value of m relating to the ridge point at n , $\text{Re}[\cdot]$ and $\text{Im}[\cdot]$ are the real and imaginary parts, respectively.

However, due to the range limitations of the arctangent function, the obtained result from Eq. (7) is merely the wrapped phase $\Phi(x, y)$ constrained within the range of $(-\pi, \pi]$. Here, the quality-guide method within the path-following methods is selected for phase unwrapping of Eq. (3). This method unwraps the wrapped phase by following a path that depends on the reliability of the pixels [23]. At this point, the true phase information $\phi(x, y)$ we need has been extracted.

C. Phase Information Extraction by DLPE

The overall process of the DLPE method is illustrated in Fig. 2. Initially, experimental fringes are obtained using the moiré deflectometry. Following the procedure described in Section 2.B, the true phase information obtained from wavelet analysis and phase unwrapping is used as the label data, which is then used to create the datasets required for training and testing the CNN. After the training set is fed into the CNN, the training loss is backpropagated, enabling iterative updates to the CNN parameters and progressively enhancing prediction accuracy. Once the loss converges, the CNN training is considered complete, enabling the trained model to predict the phase information of new fringes.

D. CNN Architecture and Training Details

The CNN used for the DLPE is based on ResUNet [42], which integrates the architectures of U-Net [43] and residual networks [44], as shown in Fig. 3. It consists primarily of two main components: down-sampling and up-sampling, which are connected through a bridge at the bottom layer. In the down-sampling part, each layer comprises two convolutional layers, a batch normalization (BN) layer and a parametric rectified linear

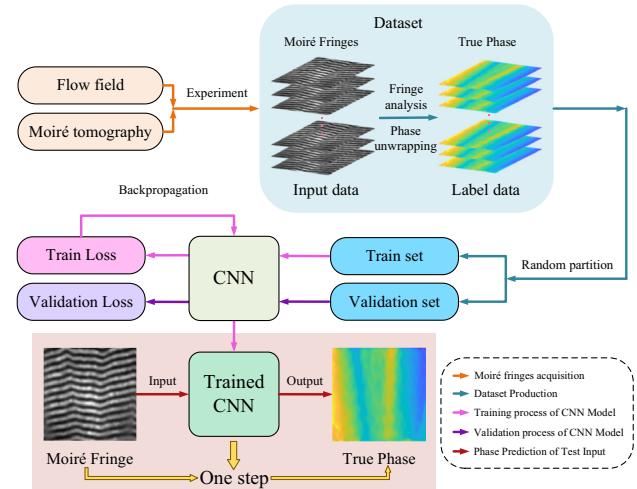


Fig. 2. Process of DLPE.

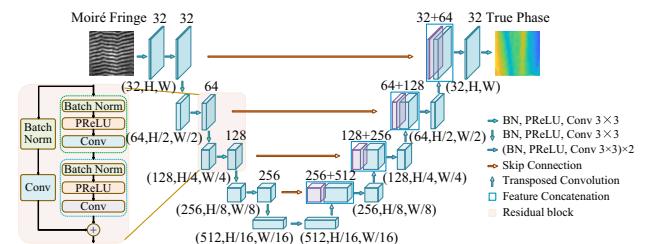


Fig. 3. Architecture diagram of CNN.

unit (PReLU) activation layer. Residual blocks are applied after these convolutional layers. The down-sampling process, which aims to extract more abstract features, is achieved through an initial convolution in the residual block with a stride of 2, doubling the number of feature channels. The up-sampling process employs transposed convolutions to restore the dimensions of the input data. To retain more feature information, feature maps from the down-sampling path of the same layer are merged via skip connections. Finally, an output convolutional layer reduces the number of feature channels to 1, enabling the prediction of the true phase information of optical fringes. In summary, this CNN architecture establishes a direct mapping between the optical fringes and its phase information.

The CNN model is trained using the PyTorch 2.0 framework on Python 3.10, executed on an RTX 4090 GPU with 24 GB

of memory. The computational platform included an Intel Xeon Platinum 8352 V CPU (2.10 GHz) and 90 GB of server memory. The training consisted of 200 epochs with a dataset of 2,400 images, each sized at 256×256 pixels. The batch size is set to 10 and the learning rate is configured to 10^{-5} . The change in the loss function value served as a key metric for evaluating the efficiency and accuracy of the training process, with the mean squared error (MSE) chosen as the loss function. Upon completing the training process, the CNN model demonstrated high efficiency, requiring approximately 700 ms to predict the true phase on a laptop equipped with a 12th-generation Intel Core i7-12700 H (2.30 GHz) CPU, a 3060 GPU, and 6 GB of memory.

3. NUMERICAL SIMULATION

To verify the feasibility and reliability of the DLPE, a simulation study is conducted in this section. The dataset is created by simulating fringes and corresponding phase information, which are subsequently used for training and testing the CNN model.

A. Training and Testing Datasets

Creating an accurate dataset is essential for ensuring the efficiency and reliability of deep learning. Unlike most datasets previously used for image segmentation, where pixel-level data represent discrete categories, both the input and label images in our study contain continuous values at each pixel. Consequently, the precision requirements for dataset generation in this work are exceptionally high at the pixel level. To generate the training dataset for the CNN model, a simulated Gaussian phase field ϕ_{simu} is constructed, which is represented as

$$\phi_{\text{simu}}(x, y) = \frac{c_1}{\sqrt{2\pi}} \exp \left[-\frac{(x + c_2)^2 + (y + c_3)^2}{b} \right] + \frac{c_4}{\sqrt{2\pi}} \exp \left[-\frac{(x + c_5)^2 + (y + c_6)^2}{b} \right], \quad (8)$$

where b is a constant, c_1, c_2, c_3, c_4, c_5 , and c_6 are the random integer.

By using Eqs. (1) and (8), the simulated fringes modulated with a random Gaussian phase field can be obtained. To enhance the CNN model's robustness to noise, random Gaussian white noise is added to the simulated fringes. From this process, 2400 fringes loaded with Gaussian phase are generated as input data, and the corresponding phases are used as label data. The dataset is divided randomly into a training set and a validation set at a ratio of 19:1, while 100 separate fringes and their phases constitute the test dataset. Both the input fringe and phase label data have dimensions of 256×256 . The input fringe and phase label data are shown in Fig. 4.

B. Feasibility and Accuracy of the DLPE

To evaluate the prediction accuracy of the DLPE, five simulated fringes not included in the training dataset are selected to predict the true phase. The results are compared with those obtained using classical Fourier analysis method, wavelet analysis method, deep learning-based fringe analysis (DLFA)

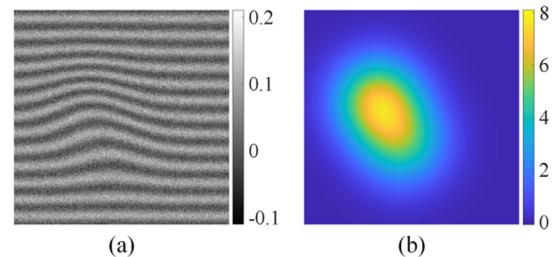


Fig. 4. Input fringe and label phase of the dataset. (a) Fringe and (b) true phase.

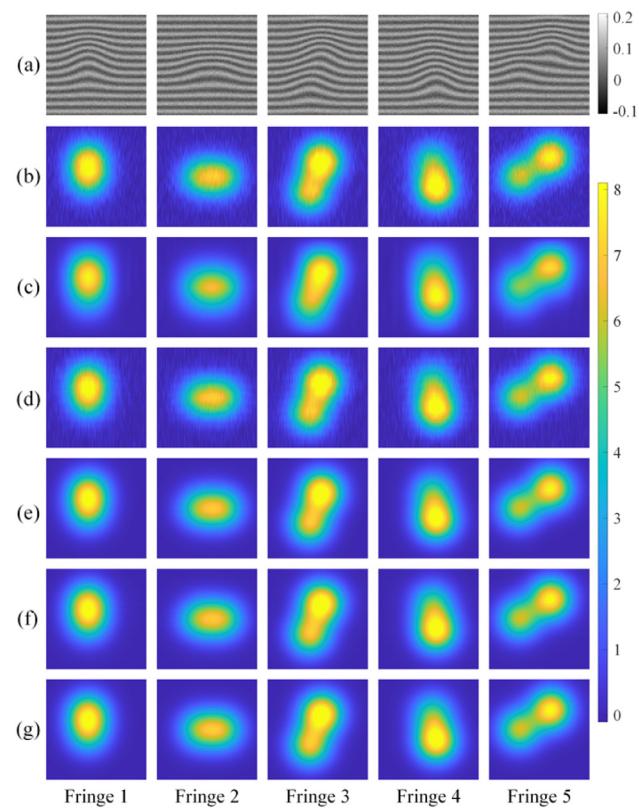


Fig. 5. True phase results of different methods on the fringes those are not in the training process. (a) Simulated fringes with noise, (b) Fourier analysis, (c) wavelet analysis, (d) DLFA method, (e) DLPU method, (f) our proposed DLPE method, and (g) label data of the true phase.

method, and deep learning-based phase unwrapping (DLPU) method, as shown in Fig. 5.

According to the comparison shown in Fig. 5, the true phase results obtained through Fourier analysis and DLFA methods are more sensitive to noise, resulting in relatively dense bar errors. In contrast, the results from wavelet analysis, DLPU, and our DLPE methods are smoother. However, wavelet analysis introduces some ripple errors in the smooth areas around the flow field. To clearly illustrate the error details, the phase error diagram is shown in Fig. 6.

As shown in Fig. 6, the errors in the results of Fourier analysis and DLFA method primarily stem from poor noise resistance, and the error distribution is relatively uniform. In contrast, the errors in wavelet analysis are mainly concentrated in the flow

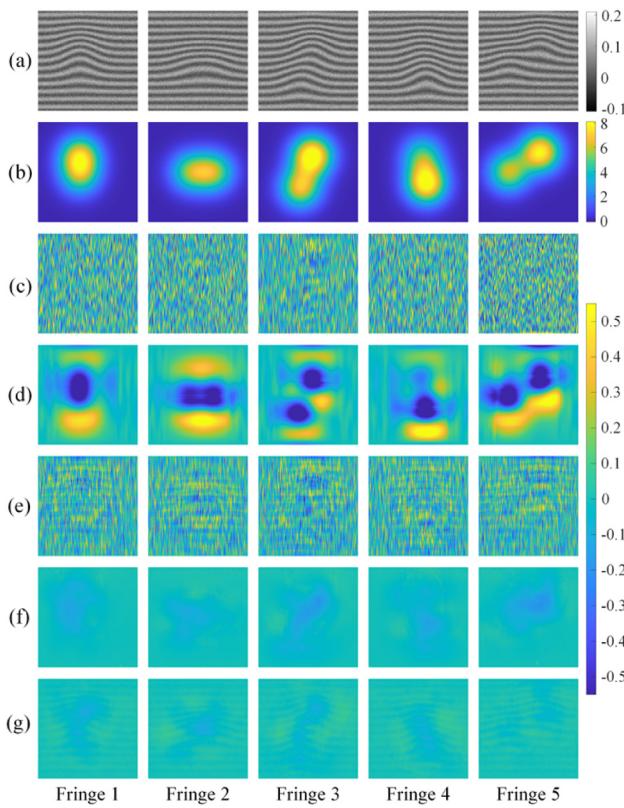


Fig. 6. Comparison of the phase error of different methods. (a) Fringe, (b) phase label data, (c) Fourier analysis, (d) wavelet analysis, (e) DLFA method, (f) DLPUs method, and (g) our proposed DLPE method.

Table 1. Prediction Evaluation of the True Phase with Different Methods

	Fringe 1	Fringe 2	Fringe 3	Fringe 4	Fringe 5
Fourier	0.2357	0.2362	0.2128	0.2394	0.2747
Wavelet	0.2031	0.2349	0.2121	0.2022	0.2651
DLFA	0.1942	0.1945	0.1990	0.1955	0.1960
DLPUs	0.0388	0.0316	0.0484	0.0368	0.0506
DLPE	0.0262	0.0270	0.0295	0.0281	0.0204

field and its surrounding areas. For the DLPUs and DLPE methods, the errors are attributed to fringe patterns that are nearly imperceptible, resulting in overall lower error magnitudes. This indicates that the prediction performance of DLPUs and DLPE is generally superior to that of Fourier analysis, wavelet analysis, and DLFA. To further quantify the accuracy of the phase results, the RMSE values for each method are shown in Table 1.

Table 1 shows that the RMSE of the phase results obtained using Fourier analysis, wavelet analysis, and DLFA is approximately 0.2, with DLFA achieving a slightly lower RMSE than the other two methods. In comparison, the RMSE of the DLPUs method is around 0.04. Notably, the DLPE method achieves an RMSE of approximately 0.02, which is nearly one-tenth that of Fourier analysis, wavelet analysis, and DLFA, and about half that of the DLPUs method. This indicates the high precision of the DLPE in predicting the true phase. To evaluate the predictive time performance of the DLPE, the phase extraction

Table 2. Prediction Time of the True Phase with Different Methods (Unit: s)

	Fringe 1	Fringe 2	Fringe 3	Fringe 4	Fringe 5
Fourier	0.804	0.823	0.820	0.844	0.854
Wavelet	1.414	1.452	1.445	1.425	1.466
DLFA	0.775	0.820	0.756	0.720	0.723
DLPUs	0.856	0.850	0.841	0.850	0.848
DLPE	0.691	0.698	0.693	0.706	0.696

time required by the DLPE, compared with Fourier analysis, wavelet analysis, DLFA method, and DLPUs method are shown in Table 2.

As shown in Table 2, Fourier analysis, DLFA method, and DLPUs method require approximately 0.8 s to extract phase information, while wavelet analysis takes around 1.4 s. In contrast, our DLPE method requires approximately 0.7 s, demonstrating a clear time advantage over other phase extraction methods.

C. Discussion of the Structure Parameter of CNN

The prediction accuracy is influenced by various structural parameters of the CNN model, such as the number of down-sampling steps n_{ds} and the channel number of convolutional kernels n_{ck} . To select the optimal parameters, one can adjust the value of a single parameter while keeping the other parameters constant, and then evaluate the performance using the loss function value. To determine the optimal selection of network depth, the CNNs with $n_{ds} = 0, 1, 2, 3, 4$ are selected separately. It should be emphasized that, compared to the CNN structure depicted in Fig. 3, the discussion below from Fig. 7 to Fig. 8 solely varies the n_{ds} (or network depth), while the n_{ck} remains consistent with that in Fig. 3 (i.e., Case 1 as described in the subsequent discussion on the n_{ck}). Additionally, down-sampling is performed in a top-down order, for example, $n_{ck} = (64, 128)$ with $n_{ds} = 1$ (i.e., network depth is 2). The performance of the training process for each selection is shown in Fig. 7.

Figure 7 shows that the loss function of the CNN model with $n_{ds} = 0$ converges to a highest value, approximately 3. The CNN model with $n_{ds} = 1$ also converges to a higher error value, whereas models with $n_{ds} \geq 2$ converge to smaller error values. Notably, the loss function of the CNN model with $n_{ds} = 4$ exhibits the least convergence value, despite some fluctuations. Additionally, the initial value of the loss function decreases as the n_{ds} increases. Overall, a higher n_{ds} (indicative of greater

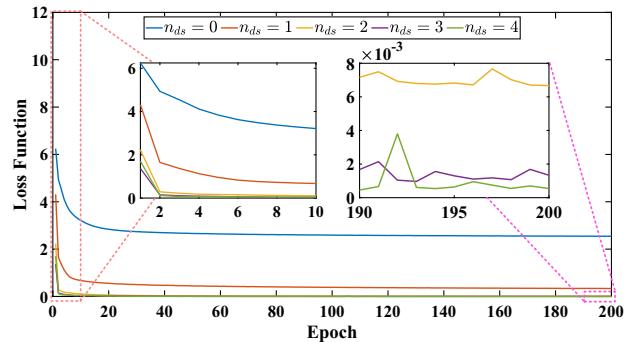


Fig. 7. Values of loss function in the training process under different numbers of down-sampling steps.

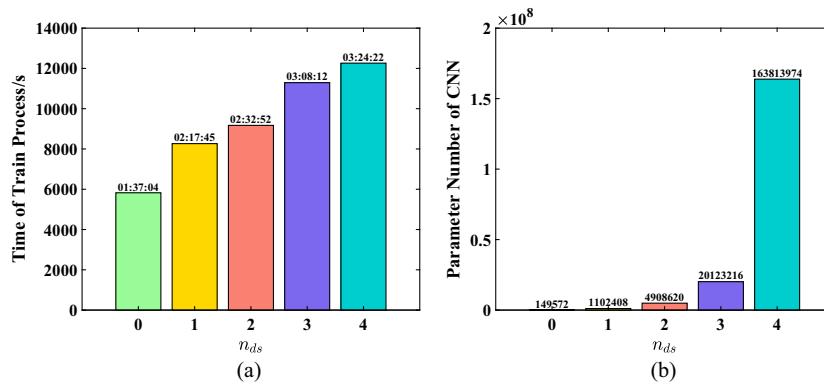


Fig. 8. Time and space costs of the training process of CNN models with different numbers of down-sampling steps. (a) training time and (b) number of parameters.

Table 3. Evaluation of Predictions with Different Numbers of Down-Sampling Steps

n_{ds}	Fringe 1		Fringe 2		Fringe 3		Fringe 4		Fringe 5	
	RMSE	SSIM								
0	1.6210	0.0106	1.5701	0.0078	1.9933	0.0116	1.7381	0.0115	1.7768	0.0109
1	0.5902	0.1637	0.7107	0.1407	0.8509	0.2005	0.7163	0.1724	0.8027	0.1727
2	0.0836	0.8249	0.0966	0.8275	0.1175	0.8645	0.1107	0.8367	0.1152	0.8453
3	0.0372	0.8750	0.0427	0.8986	0.0592	0.9254	0.0412	0.9005	0.0528	0.9054
4	0.0262	0.9164	0.0270	0.9570	0.0295	0.9611	0.0281	0.9455	0.0204	0.9436

network depth) improves the training of CNN models in DLPE methods.

Meanwhile, the loss function values on the test set indicate the final prediction accuracy of the deep learning model, providing insights into the rationale behind the settings of the structural parameters. Here, the quality of prediction results for n_{ds} is evaluated by root mean square error (RMSE) and structural similarity index (SSIM), as shown in Table 3.

Based on the evaluation index values of the prediction results for the five test fringes in Table 3, the RMSE of the CNN with $n_{ds} = 4$ is the lowest, approximately 0.02, while the SSIM is the highest, around 0.95. As n_{ds} increases, the RMSE decreases, and the SSIM increases. However, CNNs with excessively low n_{ds} exhibit poor prediction performance. Specifically, the RMSE of prediction results for CNNs with $n_{ds} = 0$ and $n_{ds} = 1$ rapidly increases to approximately 0.7 and 1.6, respectively, while the SSIM rapidly decreases to about 0.2 and 0.01. This indicates that increased n_{ds} positively impact the prediction effectiveness of CNNs, both in terms of error reduction and structural similarity improvement. Additionally, CNNs with $n_{ds} = 2$ and $n_{ds} = 3$ can also be considered when exceptionally high prediction accuracy is not required. To illustrate the time and space costs of CNN models with different n_{ds} , the training time and number of parameters of these CNNs are shown in Fig. 8.

According to Fig. 8, it is evident that both the training time and the number of parameters of the CNN increase with the n_{ds} , which is logical. Specifically, the training time for a CNN with $n_{ds} = 0$ is approximately one and a half hours, while it exceeds two h for a CNN with $n_{ds} \geq 1$, with the longest training time exceeding three h for a CNN with $n_{ds} = 4$. Regarding the number of parameters, the CNN parameters for $n_{ds} = 4$ are notably higher compared to other values, while the difference in

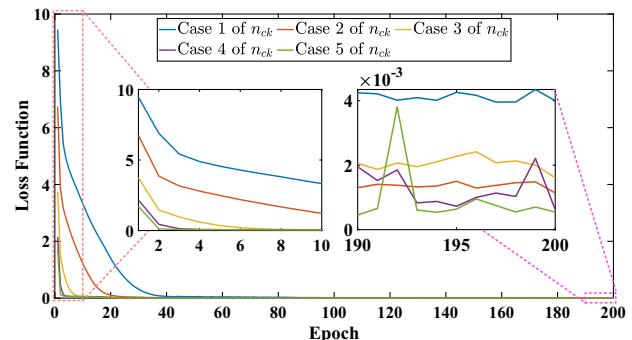


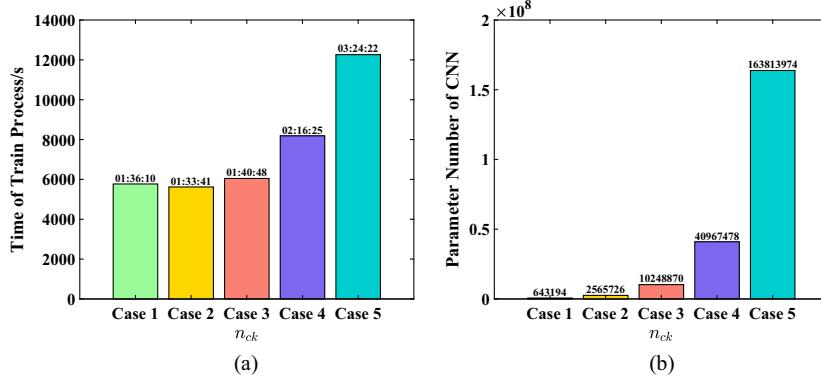
Fig. 9. Values of loss function in the training process with different numbers of convolutional kernels.

the number of parameters when $n_{ds} < 4$ is relatively smaller. In summary, although there is a significant increase in the number of CNN parameters with $n_{ds} = 4$, the corresponding increase in required training time is not substantial, indicating a relatively high cost-performance ratio for CNNs with $n_{ds} = 4$.

Similarly, to determine the optimal selection of n_{ck} , five cases of the CNN with different n_{ck} are considered. Due to the convention in many classical CNN architectures where n_{ck} is typically in the form of 2^n [45], the five cases of n_{ck} per layer are defined as follows: (4, 8, 16, 32, 64) (Case 1), (8, 16, 32, 64, and 128) (Case 2), (16, 32, 64, 128, and 256) (Case 3), (32, 64, 128, 256, and 512) (Case 4), and (64, 128, 256, 512, and 1024) (Case 5), and the number of parameters varies from few to many across Cases 1 to 5. It is important to note that such a definition is fixed with $n_{ds} = 4$, meaning the network depth is 5, thus each case comprises 5 numbers. The performance of the training process for each configuration is shown in Fig. 9.

Table 4. Evaluation of Predictions with Different Numbers of Convolutional Kernels

Case	Fringe 1		Fringe 2		Fringe 3		Fringe 4		Fringe 5	
	RMSE	SSIM								
1	0.0531	0.7367	0.0499	0.7467	0.0725	0.8078	0.0497	0.7732	0.0575	0.7871
2	0.0440	0.9096	0.0292	0.9169	0.0595	0.9349	0.0529	0.9276	0.0522	0.9326
3	0.0362	0.8019	0.0277	0.8021	0.0374	0.8557	0.0353	0.8575	0.0359	0.8566
4	0.0465	0.9310	0.0394	0.9298	0.0556	0.9430	0.0475	0.9317	0.0442	0.9436
5	0.0262	0.9164	0.0270	0.9570	0.0295	0.9611	0.0281	0.9455	0.0204	0.9436

**Fig. 10.** Time and space costs of the training process of CNN models with different numbers of convolutional kernels. (a) Training time and (b) number of parameters.

As shown in Fig. 9, the training process of the CNN in all five cases of n_{ck} gradually converges to a small loss function value. In addition, the initial value of the loss function decreases as n_{ck} increases. Overall, the difference in the final training of the CNN across the five cases is not particularly significant. Similarly, the final performance of the CNN depends on its prediction performance on the test set. To quantify the prediction performance of the CNN in detail for the five cases, the prediction results are evaluated using RMSE and SSIM, as shown in Table 4.

Table 4 shows that the difference in prediction performance between CNNs with different n_{ck} narrows compared to different n_{ds} . The RMSE and SSIM of Case 5, which has the best performance, reach approximately 0.02 and 0.94, respectively. However, it is worth noting that the RMSE of Case 3 is lower than those of Cases 2 and 4, but its SSIM is also lower than those of Cases 2 and 4. Generally, without considering cost, using as many n_{ck} as possible can better improve the prediction performance of the CNN. Besides, the training time and the number of parameters of the CNN model with different n_{ck} are shown in Fig. 10.

Similar to Fig. 8, the training time and number of parameters of the CNN generally increase with n_{ck} , although the training time of Case 1 is slightly longer than that of Case 2. Specifically, the training times of Cases 1 to 3 are approximately one and a half hours, and they are very close in duration. The training times of Cases 4 and 5 exceed two h, with the longest training time of Case 5 being more than three h. In terms of the number of parameters, Case 5 has significantly more parameters than the other cases, reaching the order of 1.6×10^8 , while the differences in the number of parameters among the other cases are slightly smaller and roughly within the order of

$10^5 \sim 10^7$. To sum up, although the number of parameters in Case 5 increased significantly, the required training time did not increase proportionately, making the cost performance of CNN in Case 5 relatively high.

Therefore, to summarize, in terms of the accuracy and structural similarity of prediction results, CNNs with $n_{ds} = 4$ and n_{ck} as defined in Case 5 provide the best performance. Although this configuration is not the lowest in terms of training time and space costs, it is the most cost-effective when considering both accuracy and structural similarity of the predicted results. Consequently, in the DLPE method applied in Section 4, this CNN configuration is selected to extract the phase information of the real flow field. However, in environments with high-cost constraints, other CNN configurations can also be considered.

D. Discussion of the Noise Level

To evaluate the robustness of the DLPE method under different noise levels, a test fringe not involved in the training process is randomly selected, and random noise with signal-to-noise ratio (SNR) of 5, 6, 7, 8, 9, and 10 dB are added. The phase results and error of the DLPE method are shown in Fig. 11.

By observing Fig. 11, under different noise levels, the phase prediction results of the DLPE method closely resemble the phase label data. Although the error slightly increases as the SNR of the fringe decreases, the phase prediction errors of the DLPE across all noise levels remain within a small range. To further quantify the impact of noise levels on prediction accuracy of the DLPE method, RMSE and peak signal-to-noise ratio (PSNR) are selected to evaluate its performance at various noise levels. The specific performance metrics are shown in Table 5.

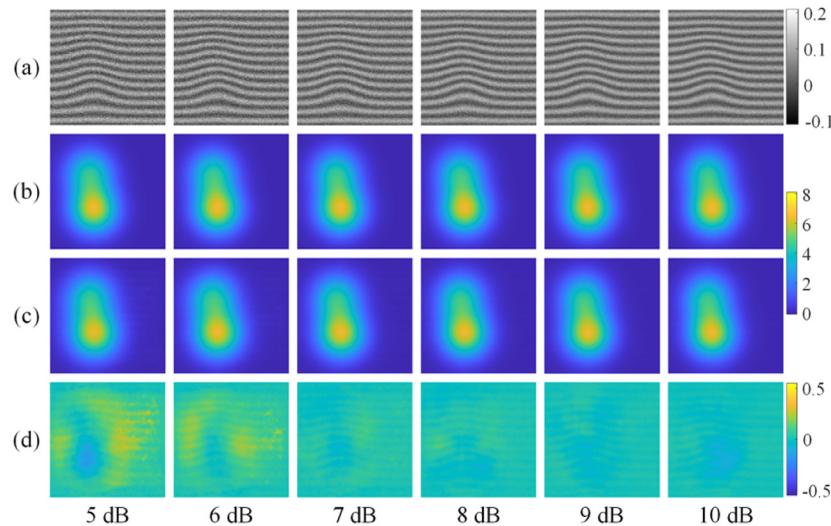


Fig. 11. True phase results and error of our proposed DLPE method under different SNRs. (a) Fringes, (b) phase label data, (c) phase results, and (d) error of phase results.

Table 5. Evaluation of Predictions with Different Noise Levels

SNR	5 dB	6 dB	7 dB	8 dB	9 dB	10 dB
RMSE	0.1059	0.0915	0.0355	0.0303	0.0210	0.0256
PSNR	35.6975	37.1287	45.3388	46.7096	49.8247	48.1025

As shown in the specific values in Table 5, the RMSE remains generally below 0.1, except for the 5 dB case, which is slightly higher than 0.1. Notably, when the SNR exceeds 7 dB, the RMSE stays below 0.04. Meanwhile, the PSNR remains above 35, and when the SNR is greater than 7 dB, the PSNR exceeds 45. These metrics indicate that the prediction results of the DLPE method are minimally affected by varying noise levels, demonstrating that our proposed DLPE method has strong robustness to noise.

4. PHASE INFORMATION EXTRACTION OF REAL FLOW FIELD BASED ON THE DLPE

To assess the performance of DLPE in extracting phase information from real flow fields and applicability in different scenarios, three types of flow fields are selected for experimentation: air heat flow field (electric soldering iron heat flow field), non-premixed combustion flow field (candle combustion flow field),

and premixed combustion flow field (propane combustion flow field), as shown in Fig. 12.

Utilizing the moiré deflectometry device shown in Fig. 1, 2,400 moiré fringes are acquired for each flow field to extract phase information using the theoretical process outlined in Section 2.B. Therefore, the moiré fringes are used as the input data, and the corresponding phase information obtained through wavelet analysis and phase unwrapping is used as the label data for the dataset. Similar to the simulation dataset described in Section 3, the 2,400 images are randomly divided in a 19:1 ratio. Each group assigns 2,280 moiré fringes to the training set, while the remaining 120 images form the validation set, ensuring the representativeness of the fringe in the validation set, which are not involved in the training process. The performance of DLPE under the three real flow fields during the training process is illustrated in Fig. 13.

It is observed that CNN training of all three flow fields converges to a small loss at the end, with some oscillation in the loss function observed during the training process for the electric soldering iron flow field. Additionally, regarding the initial loss function values, the candle combustion field exhibits the highest initial loss, followed by the electric soldering iron heat flow field, with the propane combustion field having the smallest initial loss. This phenomenon may be attributed to the relatively stable

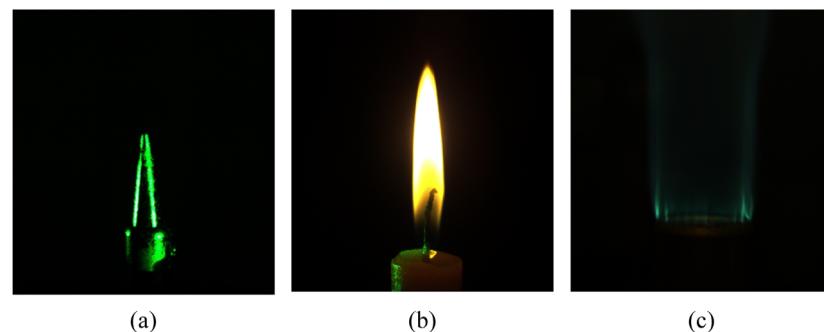


Fig. 12. Measured real flow fields. (a) Electric soldering iron, (b) candle, and (c) propane.

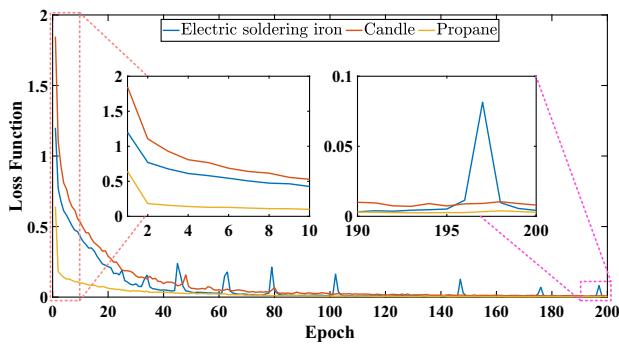


Fig. 13. Values of the loss function in the training process of three real flow fields.

jet region of propane being measured, whereas candles and soldering irons do not produce jets, resulting in a more stable structure for the propane-air flame flow field. Furthermore, the higher burning temperature of the candle compared to the soldering iron leads to more intense phase disturbances, which may further contribute to this observed difference.

After completing the training process, two test fringes per flow field are independently selected from the test sets to demonstrate the model's performance in the DLPE under real flow field conditions. The prediction results of the true phase based on DLPE are shown in Fig. 14.

As shown in Fig. 14, the prediction results of our DLPE method under the three real flow fields closely match the corresponding label data in the test sets. The flow field structure is accurately depicted, and the errors are minimal. This demonstrates that our DLPE method performs well in extracting phase information from real flow fields, further highlighting the practicality and versatility of the DLPE method. Consequently, the DLPE method shows great potential in the diagnostic technology of real complex flow fields. Most importantly, the DLPE method achieves one-step extraction of moiré fringe phase information, significantly enhancing the efficiency of

moiré deflectometry in measuring real complex flow fields. Meanwhile, considering the optimal physical conditions for phase extraction from fringes, the CNN in DLPE method is trained on datasets generated under specific physical conditions, thereby fixing the fringe characteristics during prediction. To extract phase information from fringes acquired under different physical conditions—such as variations in fringe type, period, orientation, distortion type, or contrast—the CNN model must be retrained using datasets corresponding to those specific conditions.

5. CONCLUSION

In this paper, we proposed a one-step true phase information extraction method for optical fringes based on deep learning (DLPE), eliminating the need for an intermediate wrapped phase. Comparing DLPE with Fourier analysis, wavelet analysis, DLFA method, and DLPD method demonstrated superior accuracy from simulated fringes. To design the optimal CNN structure, various configurations are discussed and compared by examining different down-sampling steps and convolutional kernel channels, with a focus on changes in the loss function, training time, and space cost. The results indicate that when time and space cost constraints are not particularly stringent, our CNN structure with four down-sampling steps and the Case 5 configuration of convolutional kernel channels offer the best cost-performance ratio in terms of prediction accuracy and structural similarity. Finally, applying DLPE to extract phase information in real flow fields using moiré deflectometry showed minimal errors and accurately depicted flow field structures. Our work lays a crucial technical foundation for enhancing optical measurement technologies requiring phase extraction and for the intelligent diagnosis of complex real flow fields using moiré deflectometry.

Funding. National Natural Science Foundation of China (61975083).

Disclosures. The authors declare no conflicts of interest.

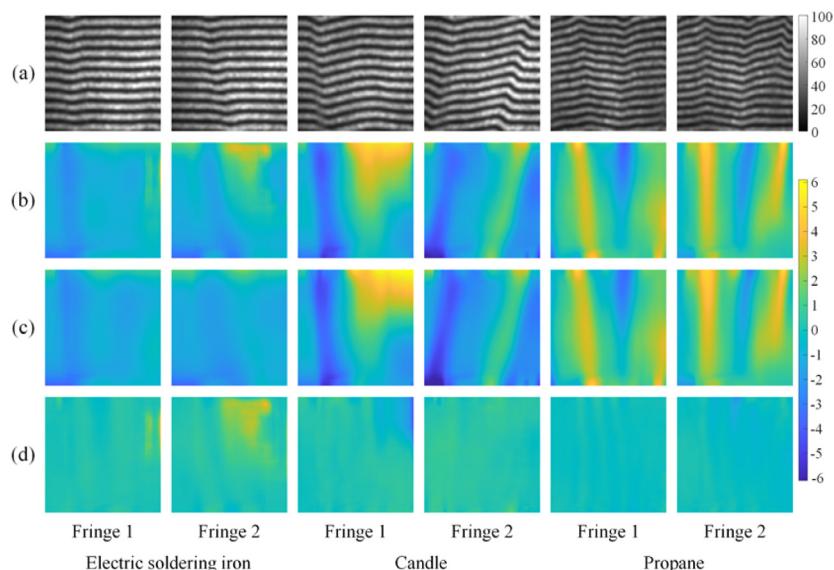


Fig. 14. Prediction results of the true phase based on the DLPE. (a) Fringes, (b) prediction results of the DLPE, (c) label data of test sets, and (d) prediction error of the DLPE.

Data availability. No data were generated or analyzed in the presented research.

REFERENCES

1. M. Servin, J. A. Quiroga, and M. Padilla, *Fringe Pattern Analysis for Optical Metrology: Theory, Algorithms, and Applications* (John Wiley & Sons, 2014).
2. U. Schnars, C. Falldorf, J. Watson, et al., *Digital Holography and Wavefront Sensing: Principles, Techniques and Applications* (Springer, 2015).
3. S. S. Gorthi and P. Rastogi, "Fringe projection techniques: whither we are?" *Opt. Lasers Eng.* **48**, 133–140 (2010).
4. Y. Chen, Y. Song, F. Gu, et al., "High-temperature flow field's electron number density measurement by two-wavelength moiré tomography," *Opt. Lett.* **41**, 1640–1643 (2016).
5. W. Cheng, Y. Chen, Q. Zhang, et al., "Spatial and temporal distributions of atmospheric refractive-index structure parameter measured by moiré deflectometry," *Opt. Commun.* **550**, 129966 (2024).
6. Y. Chen, M. Xu, W. Cheng, et al., "Phase information extraction for moiré fringes based on multiresolution analysis," *Appl. Opt.* **59**, 3543–3550 (2020).
7. M. Wang, "Fourier transform moiré tomography for high-sensitivity mapping asymmetric 3-D temperature field," *Opt. Laser Technol.* **34**, 679–685 (2002).
8. C. Zuo, S. Feng, L. Huang, et al., "Phase shifting algorithms for fringe projection profilometry: a review," *Opt. Lasers Eng.* **109**, 23–59 (2018).
9. L. Huang, Q. Kema, B. Pan, et al., "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**, 141–148 (2010).
10. H. Canabal, J. A. Quiroga, and E. Bernabeu, "Improved phase-shifting method for automatic processing of moire deflectograms," *Appl. Opt.* **37**, 6227–6233 (1998).
11. S. K. Debnath and Y. Park, "Real-time quantitative phase imaging with a spatial phase-shifting algorithm," *Opt. Lett.* **36**, 4677–4679 (2011).
12. W. Yin, Y. Che, X. Li, et al., "Physics-informed deep learning for fringe pattern analysis," *Opto-Electron. Adv.* **7**, 230034 (2024).
13. M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," *J. Opt. Soc. Am.* **72**, 156–160 (1982).
14. K. Qian, "Two-dimensional windowed Fourier transform for fringe pattern analysis: principles, applications and implementations," *Opt. Lasers Eng.* **45**, 304–317 (2007).
15. L. R. Watkins, S. M. Tan, and T. H. Barnes, "Determination of interferometer phase distributions by use of wavelets," *Opt. Lett.* **24**, 905–907 (1999).
16. K. Pokorski and K. Patorski, "Separation of complex fringe patterns using two-dimensional continuous wavelet transform," *Appl. Opt.* **51**, 8433–8439 (2012).
17. C. Zuo, J. Qian, S. Feng, et al., "Deep learning in optical metrology: a review," *Light Sci. Appl.* **11**, 39 (2022).
18. G. Sansoni, M. Carocci, and R. Rodella, "Three-dimensional vision based on a combination of gray-code and phase-shift light projection: analysis and compensation of the systematic errors," *Appl. Opt.* **38**, 6565–6573 (1999).
19. J. M. Huntley and H. Saldner, "Temporal phase-unwrapping algorithm for automated interferogram analysis," *Appl. Opt.* **32**, 3047–3052 (1993).
20. J. C. Wyant, "Testing aspherics using two-wavelength holography," *Appl. Opt.* **10**, 2113–2118 (1971).
21. Y. Song, Y. Chen, A. He, et al., "Theoretical analysis for moire deflectometry from diffraction theory," *J. Opt. Soc. Am. A* **26**, 882–889 (2009).
22. M. Zhao, L. Huang, Q. Zhang, et al., "Quality-guided phase unwrapping technique: comparison of quality maps and guiding strategies," *Appl. Opt.* **50**, 6214–6224 (2011).
23. M. A. Herraez, D. R. Burton, M. J. Lalor, et al., "Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path," *Appl. Opt.* **41**, 7437–7444 (2002).
24. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
25. G. Menghani, "Efficient deep learning: a survey on making deep learning models smaller, faster, and better," *ACM Comput. Surv.* **55**, 259 (2023).
26. K. Sharifani and M. Amini, "Machine learning and deep learning: a review of methods and applications," *World Inf. Technol. Eng. J.* **10**, 3897–3904 (2023).
27. A. Dosovitskiy, J. T. Springenberg, M. Tatarchenko, et al., "Learning to generate chairs, tables and cars with convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 692–705 (2017).
28. I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative adversarial networks," *Commun. ACM* **63**, 139–144 (2020).
29. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM* **60**, 84–90 (2017).
30. A. Sinha, J. Lee, S. Li, et al., "Lensless computational imaging through deep learning," *Opt.* **4**, 1117–1125 (2017).
31. Y. Nishizaki, M. Valdivia, R. Horisaki, et al., "Deep learning waveform sensing," *Opt. Express* **27**, 240–251 (2019).
32. R. Kulkarni and P. Rastogi, "Fringe denoising algorithms: a review," *Opt. Lasers Eng.* **135**, 106190 (2020).
33. B. Tao, Y. Wang, X. Qian, et al., "Photoelastic stress field recovery using deep convolutional neural network," *Front Bioeng. Biotechnol.* **10**, 818112 (2022).
34. J. Shi, X. Zhu, H. Wang, et al., "Label enhanced and patch based deep learning for phase retrieval from single frame fringe pattern in fringe projection 3D measurement," *Opt. Express* **27**, 28929–28943 (2019).
35. S. Feng, Q. Chen, G. Gu, et al., "Fringe pattern analysis using deep learning," *Adv. Photonics* **1**, 025001 (2019).
36. J. Qian, S. Feng, Y. Li, et al., "Single-shot absolute 3D shape measurement with deep-learning-based color fringe projection profilometry," *Opt. Lett.* **45**, 1842–1845 (2020).
37. Y. Chen and W. Cheng, "Moiré fringe analysis across diverse carrier frequencies by deep learning," *Optics Laser Technol.* **179**, 111384 (2024).
38. K. H. Jin, M. T. McCann, E. Froustey, et al., "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.* **26**, 4509–4522 (2017).
39. J. Zhang and Q. Li, "EESANet: edge-enhanced self-attention network for two-dimensional phase unwrapping," *Opt Express* **30**, 10470–10490 (2022).
40. K. Wang, Q. Kema, J. Di, et al., "Deep learning spatial phase unwrapping: a comparative review," *Adv Photonics Nexus* **1**, 014001 (2022).
41. G. 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.* **29**, 4862–4872 (2020).
42. F. I. Diakogiannis, F. Waldner, P. Caccetta, et al., "ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data," *ISPRS J. Photogramm. Remote Sens.* **162**, 94–114 (2020).
43. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, part III* 18 (Springer, 2015), pp. 234–241.
44. K. He, X. Zhang, S. Ren, et al., "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2016), pp. 770–778.
45. Y. Jin, W. Zhang, Y. Song, et al., "Three-dimensional rapid flame chemiluminescence tomography via deep learning," *Opt. Express* **27**, 27308–27334 (2019).