Optica Applicata, Vol. LV, No. 4, 2025
DOI: 10.37190/oa/210111

Deep-learning-based spectral extraction algorithm for off-axis digital holographic signal terms

YONG KONG, KE KE LIU, HE ZHANG, GENG CHEN, HOUZANG LIAO, HUI HUI WU, XIN TANG*

School of Electric and Electrical Engineering, Shanghai University of Engineering Science, Shanghai, China

*Corresponding author: tangxin@sues.edu.cn

Digital holographic microscopy is widely applied in the study of the dynamic morphology of microscopic objects due to its characteristics of non-contact measurement, high-resolution three dimensional morphology detection, and real-time observation. This paper introduces a novel off-axis digital holographic signal spectrum extraction method based on deep learning, named DL-SEDH. The DL-SEDH algorithm is trained using approximately 1000 signal spectra from a USAF 1951 resolution test target, successfully achieving adaptive extraction of signal spectra and effective suppression of interference components. To validate its effectiveness, experiments are conducted using the USAF 1951 resolution test target and onion epidermal cells as research subjects. The experimental results demonstrate that DL-SEDH not only rapidly and accurately selects signal spectra while suppressing interference frequency components, especially coherent noise distributed near the signal spectra, but also exhibits higher accuracy, robustness, and applicability compared to traditional methods. Validation on holograms not used during training confirms the effectiveness of DL-SEDH in phase reconstruction quality. The proposed DL-SEDH method introduces innovation to the off-axis digital holography field, holding significant practical value and providing an efficient and precise solution for phase reconstruction in digital holographic microscopy.

Keywords: digital holographic microscopy, deep learning, phase reconstruction, off-axis holography, signal spectrum extraction.

1. Introduction

Digital holographic microscopy (DHM) has advantages such as high resolution [1], non-contact detection [2], and three-dimensional morphology reconstruction [3], and has been widely applied in fields such as biomedical measurements (*e.g.*, quantitative phase imaging of living cells), detection of microelectronic circuit defects, and other precision metrology applications [4, 5]. Unlike analogue holographic techniques, DHM records the complex amplitude information of objects by the interference of two laser beams onto a computer, and numerically reconstructs the amplitude and phase through established reconstruction algorithms. The hologram formed by the interfer-

ence of two laser beams generally exhibits three components in the frequency domain: the zero-order term, the signal term, and the conjugate term, where the zero-order term, also known as the DC term, is mainly determined by the intensities of the two laser beams; the signal term and the conjugate term together are referred to as the interference terms, which are mainly determined by the phase difference between the two laser beams. To reconstruct the complex amplitude information in accordance with the most widespread Fourier filtering method [6], one of the interference terms, whether it is the signal term or the conjugate term, must be extracted quickly and accurately.

In recent years, non-spatial filtering methods based on the local least squares (LLS) approach have demonstrated good performance in holographic reconstruction. The LLS method, originally proposed by LIEBLING *et al*. [7], effectively extracts signals through local fitting. Subsequently, the noise-immune off-axis SPAR algorithm was developed based on this approach, further enhancing noise robustness [8]. More recently, the method has been improved by employing confidence interval intersection techniques to optimize reconstruction of objects with sharp edges [9], which is similar to the test patterns used in our experiments. The advantage of this approach lies in its ability to maintain spatial resolution [10] while reducing signal detail loss caused by inappropriate filter window shapes in traditional spatial filtering. However, the LLS method also has drawbacks, particularly the need to tune multiple parameters, which increases algorithm complexity in practical applications. In contrast, spatial filtering methods remain advantageous due to their simplicity of implementation and high computational efficiency. Therefore, considering various filtering approaches, the improvement and optimization of spatial filtering methods continue to hold significant research value.

A primary challenge in DHM is the presence of various noise sources that degrade the quality of the reconstructed phase images. It is crucial to distinguish between these sources to effectively address them. Coherent noise is a general term for noise originating from the high coherence of the laser source [11], which primarily manifests in two forms: (i) speckle noise, a random granular pattern arising from the interference of light scattered by optically rough surfaces within the setup, and (ii) parasitic interferences, which produce structured artifacts like rings or fringes due to unwanted reflections and diffraction from optical components, dust particles, or scratches. Furthermore, background noise refers to low-frequency, slowly varying intensity fluctuations, often caused by stray light or non-uniform beam illumination. Collectively, these noise sources create significant noise interference, which superimposes random fluctuations onto the true object phase, a phenomenon known as phase noise. This phase noise directly corrupts the quantitative phase information, obscures fine structural details, and is a major limiting factor for achieving high-accuracy measurements. Recent studies continue to explore advanced denoising techniques to address these challenges.

Building upon this, various types of spatial filtering methods have been proposed in recent years for selecting the signal term spectrum in holographic images. These methods can be roughly categorized into three groups: using regular-shaped frequency domain filters, adaptive filters based on image processing and segmentation, and adap-

tive spatial frequency domain filters based on deep learning networks [12]. Each of these approaches aims to balance the trade-off between noise suppression and signal preservation, leveraging different strategies to improve reconstruction quality and computational efficiency. In the first type of filtering method, for simplicity and speed, rectangular filters (RTF), circular filters (CRF), and Butterworth filters (BWF) are generally chosen to extract the signal term spectrum from holograms [13]. However, inappropriate filter shapes often introduce interference or cause loss of signal detail frequency components. The second type of filtering methods mainly uses thresholding techniques [14], as commonly employed in various digital image processing algorithms, to determine a threshold and obtain a binary, irregular-shaped window of the same size as the hologram spectrum, consisting of only 0s and 1s. This binary mask is then multiplied by the original hologram spectrum for frequency domain filtering. Weng *et al*. [15] proposed an adaptive filtering algorithm based on the Otsu algorithm to calculate the threshold, and a method based on region recognition and iterative threshold (RRF) to generate a binary mask [14]. However, the accuracy and effectiveness of this algorithm are closely related to the initial threshold setting, and sometimes the choice of a too large threshold may result in a too small filtering window, leading to the loss of high -frequency signal components; on the other hand, choosing a threshold that is too small may result in a too large filtering window, including other interference terms. Hong *et al*. combined digital image processing and histogram analysis methods, proposing a weighted adaptive filtering method based on the signal-to-noise ratio of each pixel in the signal term spectrum [16]. However, this algorithm is not suitable for cases where the signal term spectrum is relatively small, as it may disrupt the original signal term spectrum information in the process of obtaining the weighting factor. Li *et al.* used the fuzzy C-means method to segment and cluster the hologram spectrum [17], but initializing parameters such as the fuzzy weighting index significantly affects the accuracy of clustering and image segmentation, thereby affecting the quality of amplitude and phase reconstruction. In addition, the common principle of binary adaptive frequency domain filters is to allow only the frequency components within the filtering window to pass through, while completely discarding the frequency components outside the filtering window. This can lead to edge blurring and ringing artifacts in the reconstructed amplitude and phase images. The third type of filtering method mainly utilizes deep learning networks (DNN) for selecting the signal term spectrum. Xiao *et al*. used a U-net convolutional neural network for training, with thousands of manually determined signal term spectrum regions as labels [12]. Although this algorithm can achieve automatic extraction of the signal term spectrum, both dataset preparation and network training are extremely time-consuming, and the preparation stage is complex and cumbersome. Additionally, since the labels are manually selected, their accuracy and effectiveness are difficult to guarantee. Most of the above methods focus on the adaptive selection of the signal term spectrum, but rarely consider the influence of interference term spectra on the reconstruction effect. Therefore, how to accurately extract and eliminate the influence of interference terms on holographic reconstruction becomes a key issue to be urgently addressed in adaptive spatial frequency domain filtering.

The paper presents a deep-learning-based off-axis digital holographic signal term spectrum extraction algorithm (DL-SEDH). Firstly, instead of manually selecting, the method of adaptive frequency domain filtering is used to create the dataset for network training. On the other hand, interference suppression is applied during the dataset labeling process to ensure the accuracy and quality of the filtering window. Secondly, the network model is trained using approximately 1000 signal term spectra of the USAF 1951 resolution test chart. Compared to previous filtering network models, the trained model not only accurately extracts the signal term spectrum but also effectively suppresses interference terms. The trained network model enables fully adaptive selection of the signal term spectrum without any human intervention or initial parameter settings. By applying the proposed algorithm to a hologram that has not been seen during training to verify the reconstruction effect, experimental results demonstrate that the proposed algorithm exhibits adaptability, speed, robustness, and outstanding frequency domain filtering and interference suppression capabilities.

2. Method

2.1. Principle of frequency domain filtering

Off-axis digital hologram H is formed by the interference of object light wave O and reference light wave R converging at a certain angle on the hologram plane. It can be expressed as:

$$H = |O|^2 + |R|^2 + O \times R^* + O^* \times R \tag{1}$$

where $*$ denotes the complex conjugate, $|O|^2 + |R|^2$ represents the zero-order term, $O \times R^*$ and $O^* \times R$ represent the interference terms, corresponding to the signal term and the conjugate term, respectively. After performing fast Fourier transform (FFT) on the hologram, the spectrum of the hologram F can be represented as:

$$F = \mathrm{FFT}\{H\} = \mathrm{FFT}\{|O|^2 + |R|^2\} + \mathrm{FFT}\{O \times R^*\} + \mathrm{FFT}\{O \times R^*\} + \mathrm{FFT}\{O^* \times R\} \tag{2}$$

where $\mathrm{FFT}\{|O|^2 + |R|^2\}$ represents the spectrum of the zero-order term, $\mathrm{FFT}\{O \times R^*\}$ and $\mathrm{FFT}\{O^* \times R\}$ represent the spectra of the signal term and the conjugate term, respectively. These three frequency components collectively form the spectrum of the off-axis hologram, and they are separated from each other without interference.

Figure 1 shows the off-axis digital holographic microscopy (DHM) system used in this study. The light source is a helium-neon (He-Ne) laser with a wavelength of 631 nm and an output power of 50 mW, which is linearly polarized. The laser beam first passes through an adjustable attenuator, composed of a half-wave plate (HWP) and a polarizing beam splitter (PBS), to control the intensity ratio between the object and reference beams. The beam is then split by the first beam splitter (BS1). The reflected beam serves as the reference arm. It is expanded and collimated by a beam expander consisting of a microscope objective (MO1, Olympus Plan Achromat, 40×, 0.6 NA, working distance = 0.7 mm) and a collimating lens (L1, f = 150 mm). The collimated

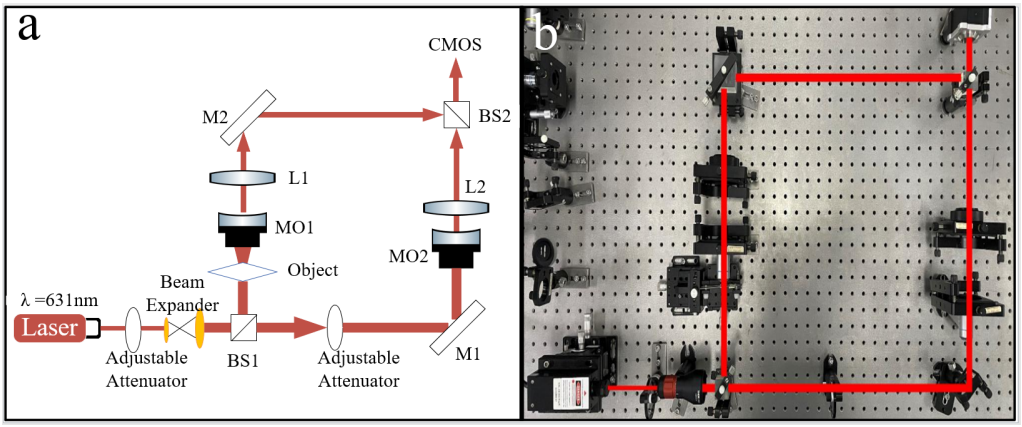

Fig. 1. Off-axis digital holographic microscopy setup. (a) Schematic diagram of the optical path. Components include: HWP, half-wave plate; PBS, polarizing beam splitter; M, mirror; BS, beam splitter; MO, microscope objective; L, lens; CMOS, image sensor. (b) Photograph of the experimental system.

plane wave is then directed by beam splitter BS3 (acting as a mirror) towards the final beam splitter BS2. The transmitted beam serves as the object arm. It is directed by mirror M1 to illuminate the sample. The light transmitted through the sample is collected and imaged by a microscope system composed of an identical objective (MO2, Olympus Plan Achromat, 40×, 0.6 NA, working distance = 0.7 mm) and a tube lens (L2, f = = 180 mm), which forms a magnified image of the sample. Finally, the object and reference beams interfere at a small angle at BS2. The resulting interference pattern, *i.e.*, the off-axis hologram, is recorded by a CMOS image sensor and transmitted to a computer for subsequent numerical processing.

To facilitate rapid and accurate extraction of signal term spectra and suppression of coherent noise, this paper proposes a framework based on deep learning for obtaining signal term spectra from holographic images, as depicted in Fig. 3. It mainly consists of four parts: holographic data preprocessing, signal term spectrum selection based on the U-Net network, frequency domain filtering, and spectral center shifting, as well as amplitude and phase map reconstruction.

2.2. Effect of the shape of the filter window on phase reconstruction

Generally, most transmitted samples are smooth. Therefore, their holographic image spectra appear as a bright cross-shaped spot, with the brightness attenuating from the center towards the periphery, and the energy of frequency components mainly concentrated at the center of the low-frequency spot, while the high-frequency region contains detailed signal information and coherent noise, as shown in Fig. 2(b). Figure 2(a) shows a holographic image captured for the fourth element of the fifth group on the USAF 1951 calibration plate, with a magnified local view inside the orange rectangle in the upper right corner, clearly revealing interference fringes. Figure 2(b) presents

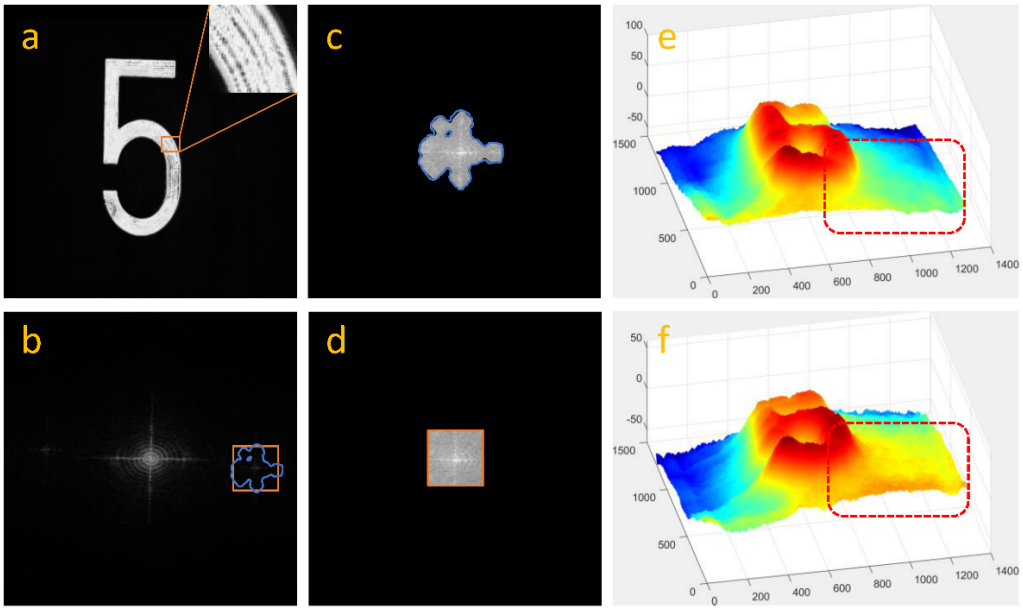

Fig. 2. Effect of different filter shapes on reconstructed phase. (a) Captured hologram; (b) Spectrum of the hologram; (c) Signal term spectrum and filtering window selected by adaptive method; (d) Signal term spectrum and rectangular filtering window selected manually; (e) 3D phase reconstruction result using adaptive filtering; (f) 3D phase reconstruction result using manual filtering.

the spectrum of the off-axis holographic image, with energy concentrated mainly in the zero-order term area, while the intensities of the signal term and conjugate term appear much weaker. To achieve high-quality phase reconstruction, the filtering window H should remove interference terms while including as many signal term spectral components as possible. Figure 2(c, d) displays the adaptively selected filtering window and manually selected rectangular filtering window for filtering the signal term spectrum and shifting it to the center of the holographic image spectrum, corresponding to the blue and orange regions in Fig. 2(b). Figures 2(e, f) show the respective results of three-dimensional phase reconstruction for adaptive and manual filtering. By comparing the results of phase reconstruction within the red dashed rectangle in Fig. 2(e, f), it is evident that the adaptive method preserves more phase details and exhibits a noticeable suppression effect on interference terms. Furthermore, manually selecting the filtering window cannot accurately identify the signal term spectrum, resulting in the filtering window containing not only the signal term spectrum but also a large amount of interference term components, thereby affecting the quality of phase reconstruction.

To observe the morphology and microstructure of samples accurately, it is necessary to extract the signal term spectrum accurately. Therefore, the design of the spatial frequency domain filtering window must filter out interference terms that cause coherent noise while retaining as many high-frequency components of the signal term

as possible, striking a balance between the quality of phase reconstruction and the suppression of interference terms.

2.3. Frequency domain filtering method based on deep learning

This paper employs the classic U-Net convolutional neural network from deep learning to suppress interference terms and extract the signal term spectrum. The principle framework is illustrated in Fig. 3. First, the holographic image directly acquired from the CMOS is subjected to data preprocessing, including rotation and translation, to adjust the target to the center of the image, as shown in Fig. 3(b). Next, the original spectrum of the holographic image is obtained using FFT, as depicted in Fig. 3(c). Then, the zero-order term and conjugate term are eliminated [18], and the original holographic image is cropped to half size, resulting in Fig. 3(d). Subsequently, the cropped spectrum is input into the trained U-Net, and the output is the binarized frequency domain filtering window, as shown in Fig. 3(e). To perform frequency domain filtering, the filtering window needs to be padded to the original size, as illustrated in Fig. 3(f). Then, the frequency domain filtering is conducted by multiplying it with the original spectrum, resulting in the signal term spectrum with interference terms removed, as

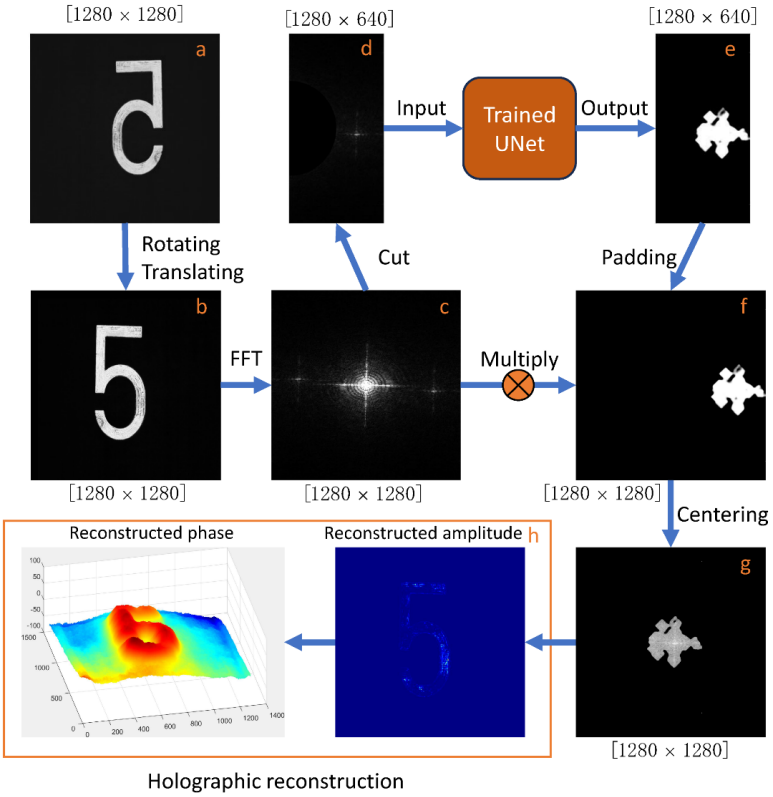

Fig. 3. Framework diagram of DL-SEDH based frequency domain filtering algorithm.

shown in Fig. 3(g). Finally, holographic reconstruction is performed on the amplitude and phase maps, as shown in Fig. 3(h).

3. Data preparation and model training

3.1. Preparation of data sets

To establish the dataset for training the deep learning network model, partial digital holographic images of the USAF 1951 resolution test chart were captured using the off-axis DHM setup, as shown in Fig. 1(a). The density and orientation of the interference fringes are adjusted by tuning the horizontal position and tilt angle of BS2, which adjusts the distribution of the signal term, conjugate term, and zero-order term in the holographic image spectrum, according to the angular spectrum theory. The actual optical setup is depicted in Fig. 1(b).

The holographic images of the third, fourth, and fifth elements of the fourth group of the calibration chart were captured. The resolution of each element was determined to be 20.16, 22.63, and 25.40 LP/mm, respectively. The corresponding limit resolutions per line pair were calculated as 49.6, 44.2, and 39.4 μm. A total of 915 holographic images were obtained by capturing 305 holographic images for each element, with 5 images taken at each position and 10 different positions captured in each direction (front, back, up, down, left, and right). During the capturing process, efforts were made to ensure that the target element remained within the field of view of the CMOS image sensor to prevent it from moving out of view. It is important to note that different elements have different frequency spectra, and noise interference may also be present. Therefore, the dataset contains reconstructed images with different resolutions and various distributions of coherent noise. The label images were generated by combining the adaptive filtering algorithm proposed in [15] with morphological processing in digital image processing. In the label image generation process, considering that interference around the signal term may affect the quality of the final phase reconstruction, a phase noise filtering step was added during frequency domain filtering to produce high-quality label images. Additionally, to facilitate better learning by the network, the zero-order term and the conjugate term were first removed, and the original spectrum was cropped before network training [12]. This approach simplifies the selection of the signal term to resemble a straightforward image segmentation problem and significantly reduces computational complexity, thereby alleviating the cache pressure on the computer.

3.2. Model training

The CNN model used in this study is an improved U-Net architecture [19,20], as illustrated in Fig. 4. The network input and output have the same spatial dimensions. The input is the cropped signal term spectrum, represented as a grayscale image, while the output is the corresponding signal term filtering window, represented as a binary image. The improved U-Net model comprises multiple convolutional layers equipped

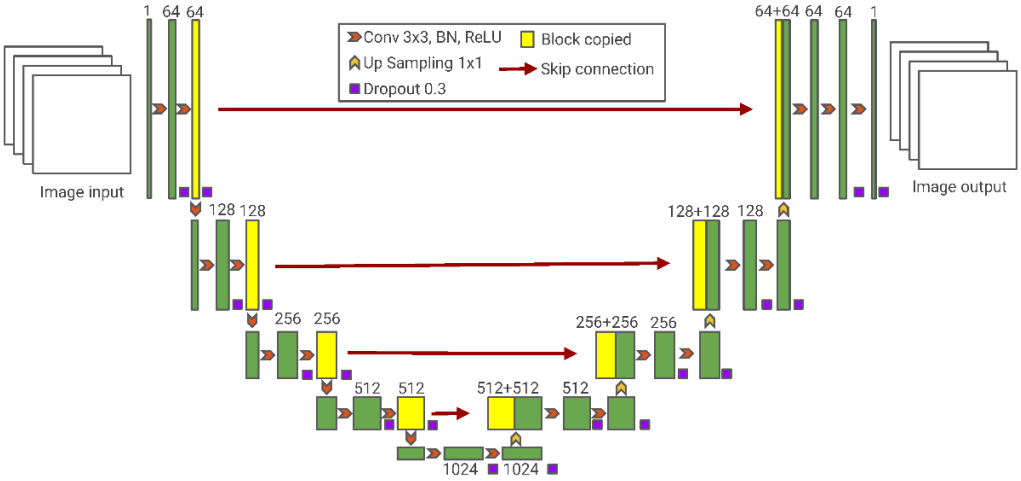

Fig. 4. Improved U-Net convolutional neural network model structure.

with ReLU nonlinear activation functions, batch normalization (BN) layers [21,22], upsampling layers, and skip connections. The network consists primarily of convolutional blocks, downsampling modules (left half of the diagram), and upsampling modules (right half), exhibiting a symmetric overall architecture.

Within each convolutional block, the convolutional kernel size is fixed at 3×3 with a stride of 1 and padding of 1; the padding mode is set to "reflect" padding to preserve edge information and enhance feature extraction capability compared to standard zero -padding. Following each convolution, batch normalization is applied to stabilize and accelerate training by normalizing the feature distributions, which also helps mitigate overfitting. Subsequently, a dropout layer with a dropout rate of 0.3 is employed to randomly deactivate 30% of neurons during each training iteration, serving as a regularization technique to further reduce overfitting and improve model generalization. Finally, the ReLU activation function introduces nonlinearity, enabling the network to learn complex mappings.

In contrast to the original U-Net that uses max-pooling for downsampling—a process which can cause loss of valuable feature information and lacks trainable parameters—this study replaces max-pooling with a 2×2 convolution operation with stride 2, padding of 1, and reflect padding mode. This convolutional downsampling layer is followed by batch normalization and ReLU activation, allowing the network to learn more meaningful feature representations while reducing spatial resolution. During downsampling, the number of input and output channels remains unchanged. For upsampling, nearest-neighbor interpolation with a scale factor of 2 is utilized to restore spatial resolution. To enhance segmentation accuracy and preserve spatial context, skip connections concatenate feature maps from shallower layers of the encoder to corresponding deeper layers in the decoder. Channel dimensionality is adjusted during upsampling by applying a 1×1 convolution with stride 1; this layer does not perform feature ex-

traction but solely reduces the number of channels to half of the input size. The skip connections ensure the consistency of input and output channel dimensions, facilitating effective fusion of multi-scale features.

The input images and their corresponding label images are fed into the network simultaneously during training, and the adaptive moment estimation (Adam) optimizer is employed for iterative parameter optimization. Training proceeds in batches, where the signal term spectrum is input into the network model, and the optimizer minimizes the discrepancy between the network output and the ground truth labels. The loss function utilized is the cross-entropy loss, which effectively quantifies the difference between predicted probabilities and true class labels in classification tasks.

To prevent overfitting and enhance the model's generalization capability, a dropout layer with a dropout rate of 0.3 is applied during training. Dropout works by randomly deactivating 30% of neurons in each training iteration, which forces the network to learn more robust and redundant representations instead of relying on specific pathways. Other training parameters are as follows: batch size is set to 1, input image size is 1280×640 pixels, the number of channels in the first convolutional layer is 64, and the deepest convolutional layer has 1024 channels. The network is trained for 20 epochs. This model is implemented using the PyTorch deep learning framework in Python and accelerated with an NVIDIA Quadro P5000 GPU.

To evaluate the trained model, randomly selected holographic images from three samples are analyzed. The differences between the filtering windows produced by the trained network and those obtained via the adaptive spatial frequency domain filtering algorithm are compared, as shown in Fig. 5. Specifically, the first column displays the grayscale holographic images of the three samples; the second column shows their corresponding spectra after removal of the zero-order and conjugate terms followed by cropping; the third column presents the filtering windows determined by the adaptive filtering method; and the last column depicts the concatenation of the filtering windows generated by the neural network (right half) and the ground truth images (left half).

The automatic signal term spectrum selection algorithm based on the improved U-Net network proposed in this study was trained on a self-constructed dataset derived from the USAF 1951 calibration target experiment. The spatial frequency domain filtering results using the trained model are presented in Fig. 5. The first column of Fig. 5 displays the grayscale holograms of the USAF 1951 calibration target, including Element 3 (Fig. 5(a)), Element 4 (Fig. 5(e)), and Element 5 (Fig. 5 (i)). Correspondingly, Figs. 5 (b), (f), and (j) show the cropped spectra after the removal of conjugate terms. It can be observed that the zero-order diffraction terms have been eliminated via a high -pass filtering (HPF) method. Additionally, a bright spot appears in the lower right corner of the signal term spectrum due to environmental interference, which is referred to as noise in this study. Notably, although all three samples originate from the fourth group of elements on the calibration target, the distributions of the signal term spectra exhibit distinct variations. This variation is clearly reflected in the shapes of the adaptive filtering windows obtained by conventional algorithms, as illustrated in Figs. 5 (c), (g), and (k). The last column of Fig. 5 compares the filtering windows predicted by the

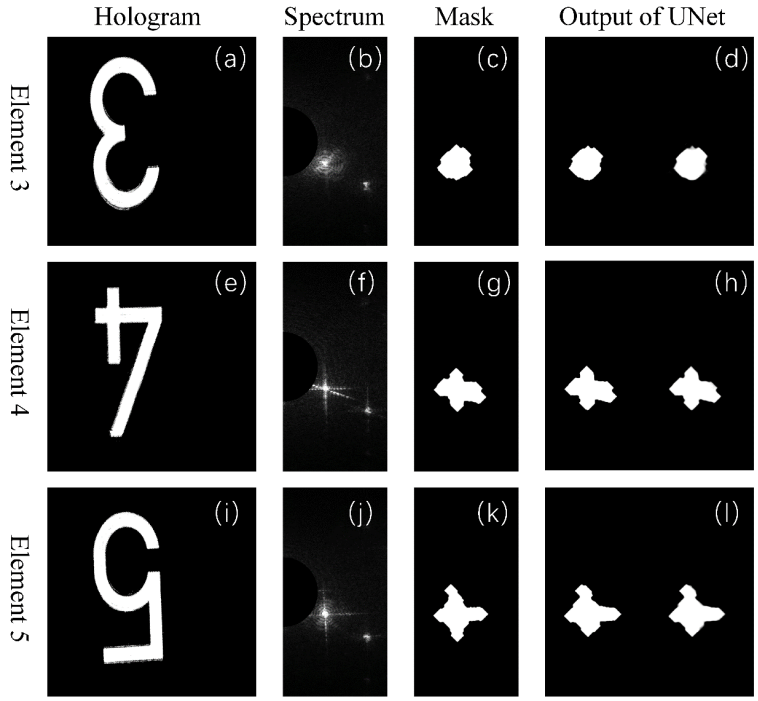

Fig. 5. Validation of the effectiveness of the network model. (a) Element 3 hologram; (b) Element 3 cropped spectrum; (c) Element 3 label (filtering window); (d) Element 3 model output; (e) Element 4 hologram; (f) Element 4 cropped spectrum; (g) Element 4 label (filtering window); (h) Element 4 model output; (i) Element 5 hologram; (j) Element 5 cropped spectrum; (k) Element 5 label (filtering window); (l) Element 5 model output.

trained U-Net model with those derived from traditional adaptive spatial frequency domain filtering methods. The results demonstrate that the proposed improved U-Net based automatic signal term spectrum selection algorithm effectively learns the intrinsic features of the signal term spectrum in off-axis holographic images.

4. Experimental verification and results

To evaluate the effectiveness and applicability of the proposed deep-learning-based signal term spectrum feature learning algorithm, a series of phase reconstruction results obtained using the automatic signal term spectrum selection algorithm based on the improved U-Net network proposed in this paper were compared in detail with those obtained using traditional phase reconstruction algorithms. This comparison includes experiments conducted with the USAF 1951 calibration board for validating the effectiveness of the proposed algorithm and experiments with onion specimen biological cells to demonstrate the algorithm's applicability.

The phase reconstruction results obtained using the commonly used circular filter (CRF) [17] and rectangular filter (RTF) from the traditional-shape-based filtering

methods, as well as the segmentation-based adaptive filtering (SAF) [13] and the deep -learning-based filtering algorithm DL-SEDH proposed in this paper, were compared. All the spatial frequency domain filtering methods mentioned above were implemented in MATLAB 2021a environment, using an Intel® Xeon® Silver 4114 CPU with a clock frequency of 2.20 GHz and 64 GB of memory.

4.1. USAF 1951 calibration board experiment

Figure 6(a) shows the comparison of the three-dimensional phase reconstruction results obtained using the proposed algorithm and three other commonly used frequency domain filtering algorithms. These methods were applied to process a hologram captured from a portion of the USAF 1951 calibration target, as depicted in Fig. 1. An enlarged view of the region outlined by the orange dashed rectangle is shown in Fig. 6(b). Figure 6(c) displays the off-axis spectrum corresponding to the hologram, revealing three distinct frequency components that are separated from each other. The orange rectangle in the spectrum corresponds to the –1 frequency component, which represents the signal frequency component. An enlarged view of this region is presented in Fig. 6(d). It is pertinent to clarify the selection of these specific target elements from the USAF 1951 test chart. The primary objective of this experiment was not to determine the ultimate spatial resolution of our imaging system, which is fundamentally limited by the numerical aperture of the objective and the pixel pitch of the CMOS sensor. Instead, our goal was to rigorously validate the performance of the proposed DL-SEDH algorithm in terms of its phase reconstruction accuracy and interference suppression capabilities. By selecting larger elements (group 4, elements 3–5) that are clearly resolved by our DHM system, we ensure that the recorded holograms provide a high-contrast, high-fidelity ground truth, free from artifacts introduced by the system's modulation transfer function (MTF) at its resolution limit. This approach allows for an unambiguous and isolated assessment of the algorithm's ability to accurately extract the signal spectrum while rejecting noise, which is the central focus of this study.

Figures 6(e) and (f) depict the circular ring filter (CRF) and rectangular tapered filter (RTF), respectively, which belong to manually selected filters (MSFs) among regular-shaped filters. The advantage of MSFs lies in their precise selection of position and size, which can greatly ensure the quality of the reconstructed phase. However, due to the manual intervention required in the phase reconstruction process, MSFs suffer from low time efficiency and lack systematic real-time requirements. Comparing CRF to RTF, under the same filtering radius condition, CRF contains fewer frequency components. Although this reduces background noise, some details are also lost. RTF, on the other hand, retains more detail information but is also affected by background noise. Figure 6(g) presents the result of the segmentation-based adaptive filter (SAF), which belongs to irregular-shaped filters.

It is evident that the shape of the filtering window of SAF is basically consistent with the actual shape of the signal frequency spectrum. Therefore, SAF can contain more signal frequency components, especially high-frequency components corresponding

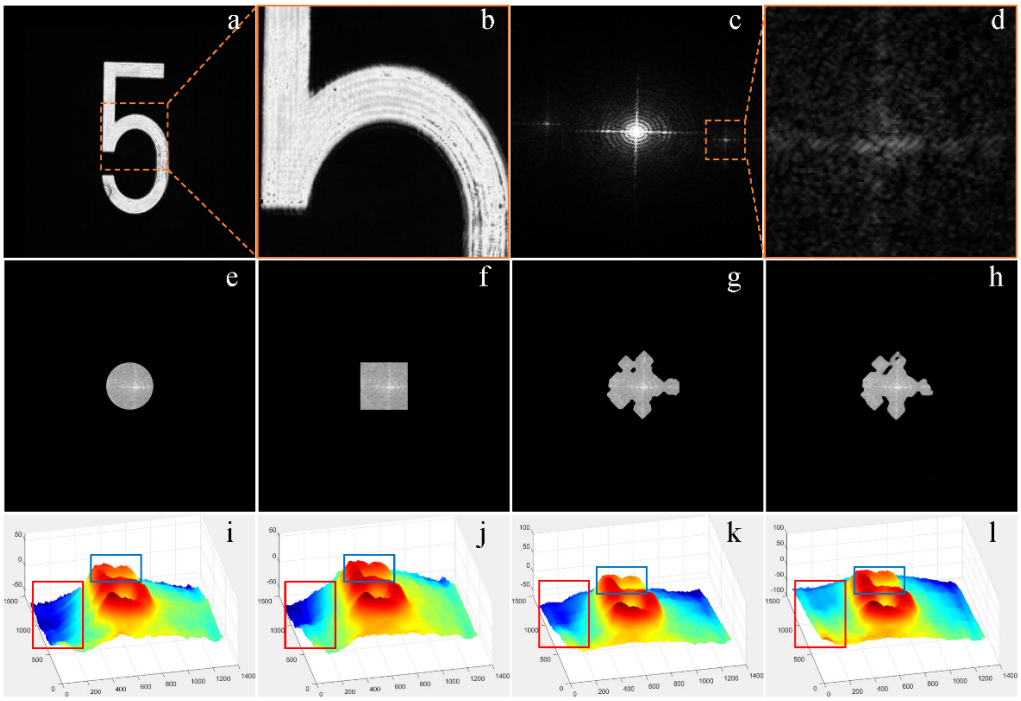

Fig. 6. Comparison of phase reconstruction effects of DL-SEDH and three other filtering algorithms. (a) Hologram; (b) Enlarged view of the region outlined by the orange dashed rectangle in subfigure (a); (c) Off-axis spectrum corresponding to the hologram; (d) Enlarged view of the –1 frequency component outlined by the orange rectangle in subfigure (c); (e) Signal term spectrum extraction and shifting using CRF; (f) Signal term spectrum extraction and shifting using RTF; (g) Signal term spectrum extraction and shifting using SAF; (h) Signal term spectrum extraction and shifting using DL-SEDH; (i) Phase reconstruction using CRF; (j) Phase reconstruction using RTF; (k) Phase reconstruction using SAF; (l) Phase reconstruction using DL-SEDH.

to the details of the reconstructed phase, while suppressing a certain degree of interference frequency components corresponding to background noise in the reconstructed phase. Figure 6(h) shows the filtering window generated by the DL-SEDH algorithm proposed in this paper after training the network. The shape and size of the window are basically consistent with the distribution of the signal frequency spectrum, similar to SAF. However, unlike SAF, DL-SEDH handles the high-frequency part of the signal frequency spectrum more accurately. This is because the deep learning network is trained and learned from a large number of different holographic spectrum images, enabling it to comprehensively and accurately learn the commonalities and details in the signal frequency spectrum. Therefore, the filtering window generated by DL-SEDH fits the original signal frequency spectrum more accurately in terms of shape, position, and size.

Figures 6(i)–(l) respectively show the three-dimensional phase reconstruction results of the hologram in Fig. 6(a) using CRF, RTF, SAF, and the DL-SEDH algorithm proposed in this paper. The red rectangle in the images represents the inherent back-

ground noise in the signal, corresponding to the low-frequency components in the signal frequency spectrum. It can be observed that CRF has the largest background noise interference, which even affects the details of the signal itself at the edges.

RTF and SAF have significantly improved background noise compared to CRF, especially SAF, which shows good improvements in both amplitude and affected areas. The blue rectangle in the images represents the details within the signal, corresponding to the high-frequency components in the signal frequency spectrum. Clearly, CRF and RTF can retain sufficient signal details because their filtering window radius is large enough to preserve more frequency components. However, SAF loses some high-frequency components in order to match the distribution of the signal frequency spectrum, resulting in severe distortion of the internal details of the reconstructed phase. The DL-SEDH algorithm proposed in this paper effectively addresses this issue. Like SAF, DL-SEDH also generates a filtering window that matches the distribution of the signal frequency spectrum. However, its essence lies in the deep convolutional neural network learning from a large number of holographic spectrum images. Therefore, DL-SEDH can learn high-frequency signal details that SAF may overlook. In summary, the DL-SEDH algorithm proposed in this paper can achieve high-quality reconstruction of three-dimensional phase holograms.

4.2. Onion epidermal cell biology experiment

Figure 7(a) shows an off-axis hologram of onion epidermal cells captured using a CMOS image sensor, with dimensions of 2135×2135 pixels. Figure 7(b) presents a zoomed-in view of the hologram within the red rectangle, clearly showing the interference fringes. Figure 7(c) displays the frequency spectrum distribution of the onion epidermal cell hologram on a logarithmic scale, with the magnified view of the signal frequency spectrum shown in Fig. 7(d). It can be observed that the signal frequency spectrum is scattered with a large number of interference components, affecting the quality of the reconstructed phase. The second column of Fig. 7 presents the filtering windows determined by the DL-SEDH algorithm proposed in this paper and three other filtering algorithms. Correspondingly, the third column of Fig. 7 shows the three-dimensional phase reconstructions of the onion epidermal cells obtained using each algorithm. Figures 7(e)–(h) show the filtering windows generated by CRF, RTF, SAF, and the trained DL-SEDH algorithm for the onion epidermal cell hologram in Fig. 7(a). The reconstructed phases corresponding to Fig. 7(i)–(l) are presented accordingly.

Due to the use of in-line holography technique during the capture of onion epidermal cells [23,24], the basic morphology of the onion epidermal cells can be clearly observed in Fig. 7(a). In-line holography refers to the condition where the imaging plane of the sample coincides with the receiving surface of the CMOS sensor during hologram capture, allowing for the acquisition of holograms with clear sample contours. According to the angular spectrum theory [25,26], adjustments in the tilt direction and density of interference fringes in Fig. 7(b) enable control over the distribution of the three-fre-

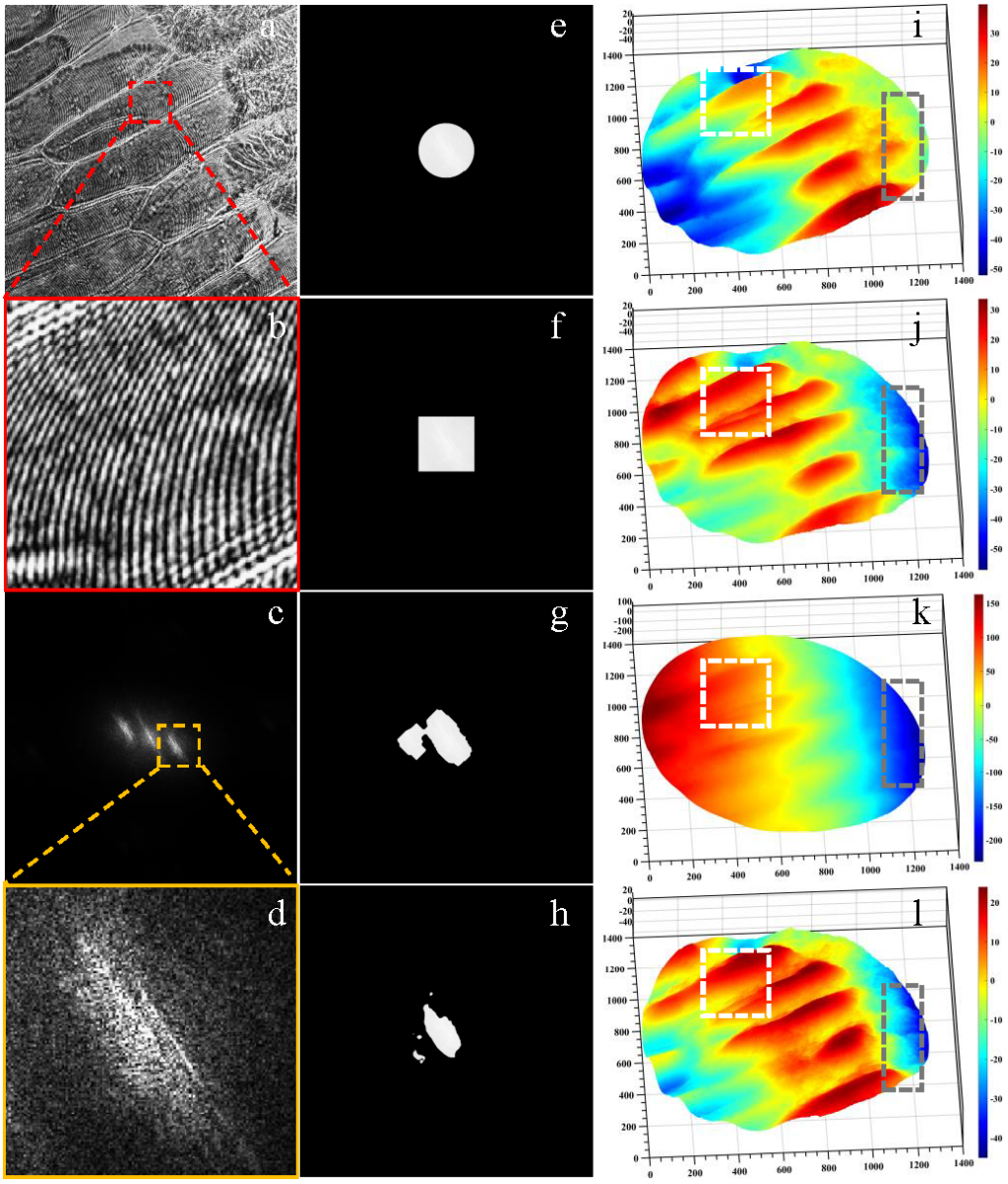

Fig. 7. Three-dimensional phase reconstruction of onion epidermal cells. (a) Hologram of onion epidermal cells; (b) Localized zoom of hologram; (c) Spectrogram; (d) Localized zoom of spectrogram; (e) CRF filtered shape; (f) RTF filtered shape; (g) SAF filtered shape;(h) DL-SEDH filtered shape; (i) CRF reconstructed phase; (j) RTF reconstructed phase; (k) SAF reconstructed phase; (l) DL-SEDH reconstructed phase.

quency spectra in Fig. 7(c), thereby achieving mutual separation of the three-frequency spectra.

The determination of the filtering windows in Figs. 7(e) and(f) requires manual intervention, where the size, position, and shape of the filtering windows are manually determined based on the observation and analysis of the distribution of signal frequency spectra. In contrast, Figs. 7(g) and (h) demonstrate automatic selection of the signal frequency spectra without the need for manual adjustment of filtering parameters. It can be observed that SAF and DL-SEDH not only accurately extract the low-frequency components of the signal frequency spectra, but also select the high-frequency components to varying degrees, ensuring the quality of phase reconstruction.

From the phase reconstruction details within the white rectangular boxes in Figs. 7(i) and (j), it can be seen that CRF suffers from severe distortion of phase details due to the loss of high-frequency components during filtering. A similar issue is also observed in SAF, as shown in Fig. 7(k), attributed to SAF's inability to accurately identify signal frequency spectra and interference, despite its filtering windows being closely aligned with the signal frequency spectra. In contrast, DL-SEDH, proposed in this paper, can rapidly and accurately distinguish between signal frequency spectra and interference after learning from a large amount of signal frequency spectrum data, thereby achieving both high-quality phase reconstruction and interference suppression. Additionally, for the edge details selected within the gray rectangular boxes on the right side of Figs. 7(i)–(l), DL-SEDH achieves results consistent with RTF, while the internal details within the white rectangular boxes are more abundant. CRF and SAF, on the other hand, result in blurred edge details due to the loss of high-frequency components, affecting the quality of phase reconstruction.

In summary, the proposed DL-SEDH algorithm based on deep learning for off-axis digital holographic signal spectrum extraction not only addresses the challenge of manual intervention and determination of filtering windows required by traditional holographic filtering algorithms but also, due to its learning from a large amount of data samples, can acquire deep features of the signal spectrum. Consequently, it accurately identifies interference and signal spectrum, thereby improving the quality of phase reconstruction.

To provide a more objective and quantitative evaluation that complements the visual assessments, we calculated the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) for the reconstructed phase images of the USAF 1951 test chart. For this analysis, the ground truth phase was meticulously generated by applying the SAF method followed by careful manual refinement of the spectrum to ensure maximum signal preservation. The results, summarized in Table 1, show a comparison between the proposed DL-SEDH and the three other filtering methods (CRF, RTF, SAF) [27-31]. As indicated in Table 1, DL-SEDH consistently achieves the highest PSNR and SSIM scores. The superior PSNR value highlights its exceptional capability in suppressing background and phase noise, while the higher SSIM score confirms that it

T a b l e 1. Quantitative comparison of phase reconstruction results using different filtering methods.

Method	PSNR [dB]	SSIM
CRF	25.13	0.82
RTF	27.45	0.86
SAF	31.58	0.91
DL-SEDH (ours)	35.72	0.96

better preserves the fine structural details of the original object. These quantitative metrics strongly support our qualitative observations and demonstrate the superior performance of the proposed method.

5. Conclusions

The present study introduces a deep-learning-based off-axis digital holographic signal spectrum extraction method, termed DL-SEDH. By training the deep learning network model on approximately 1000 signal spectra of the USAF 1951 calibration board, DL-SEDH successfully achieves adaptive extraction of the signal spectrum and effective suppression of interference components. Compared to traditional methods, this algorithm requires no manual intervention or initialization of input parameters; it only needs one round of network training to select the signal spectrum rapidly and stably. In the experimental validation section, using the USAF 1951 calibration board and onion epidermal cells as the research subjects, DL-SEDH demonstrates adaptability, speed, robustness, and outstanding frequency domain filtering and interference suppression capabilities compared to traditional phase reconstruction algorithms. Experimental results indicate that the trained DL-SEDH network can generate spatial frequency domain filtering windows, accurately selecting the signal spectrum while suppressing interference frequency components, especially coherent noise near the signal spectrum. Validation on holograms not used in training confirms the robustness and effectiveness of the DL-SEDH method in phase reconstruction quality.

In conclusion, DL-SEDH, as a deep-learning-based off-axis digital holographic signal spectrum extraction method, demonstrates significant application value in the field of off-axis digital holography. It provides an efficient and accurate solution for phase reconstruction in digital holographic microscopy. Through validation experiments in different scenarios, the universality and reliability of the DL-SEDH method have been demonstrated, laying a solid foundation for advancing the application of digital holographic microscopy in the study of dynamic morphology of small objects.

Acknowledgements
This work was supported in part by the Natural Science Foundation of China under Grant 62201338 and Natural Science Foundation of Shanghai under Grant 21ZR1424600.

References

[1] MA F., LI Y., WANG X., DU Y., GONG Q., CHENG J., QIN L., SU J., HU Y., *Investigation of the effective aperture: towards high-resolution Fresnel incoherent correlation holography*, Optics Express **29**(20), 2021: 31549-31560. https://doi.org/10.1364/OE.439449

[2] KREIS T., *Application of digital holography for nondestructive testing and metrology: a review*, IEEE Transactions on Industrial Informatics **12**(1), 2016: 240-247. https://doi.org/10.1109/TII.2015.2482900

[3] JEON P., LEE H., KIM J., LIU C., KIM D., *Analysis of three-dimensional mapping problems in incoherent digital holography*, Optics Express **28**(4), 2020: 4501-4515. https://doi.org/10.1364/OE.384477

[4] PARK Y.K., DIEZ-SILVA M., POPESCU G., LYKOTRAFITIS G., CHOI W., FELD M.S., SURESH S., *Refractive index maps and membrane dynamics of human red blood cells parasitized by* Plasmodium falciparum, Proceedings of the National Academy of Sciences of the United States of America **105**(37), 2008: 13730-13735. https://doi.org/10.1073/pnas.0806100105

[5] CAO R., XIAO W., WU X., SUN L., PAN F., *Quantitative observations on cytoskeleton changes of osteocytes at different cell parts using digital holographic microscopy*, Biomedical Optics Express **9**(1), 2018: 72-85. https://doi.org/10.1364/BOE.9.000072

[6] TAKEDA M., INA H., KOBAYASHI S., *Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry*, Journal of the Optical Society of America **72**(1), 1982: 156-160. https://doi.org/10.1364/JOSA.72.000156

[7] LIEBLING M., BLU T., UNSER M., *Complex-wave retrieval from a single off-axis hologram*, Journal of the Optical Society of America A **21**(3), 2004: 367-377. https://doi.org/10.1364/JOSAA.21.000367

[8] KATKOVNIK V., SHEVKUNOV I.A., PETROV N.V., EGIAZARIAN K., *Wavefront reconstruction in digital off-axis holography via sparse coding of amplitude and absolute phase*, Optics Letters **40**(10), 2015: 2417-2420. https://doi.org/10.1364/OL.40.002417

[9] KATKOVNIK V., SHEVKUNOV I., PETROV N.V., EGIAZARIAN K., *High-accuracy off-axis wavefront reconstruction from noisy data: local least square with multiple adaptive windows*, Optics Express **24**(22), 2016: 25068-25083. https://doi.org/10.1364/OE.24.025068

[10] KATKOVNIK V., SHEVKUNOV I.A., PETROV N.V., EGIAZARIAN K., *Sparse approximations of phase and amplitude for wave field reconstruction from noisy data*, Proceedings of the SPIE, Vol. 9508, Holography: Advances and Modern Trends IV, 2015: 950802. https://doi.org/10.1117/12.2177657

[11] XIONG W., WANG S., GAO S., GAO Z., TANG H., KOCH A.W., WU S., SUN W., *Complex amplitude domain filtering for phase measurement in speckle interferometric optics*, Optics Communications **534**, 2023: 129328. https://doi.org/10.1016/j.optcom.2023.129328

[12] XIAO W., WANG Q., PAN F., CAO R., WU X., SUN L., *Adaptive frequency filtering based on convolutional neural networks in off-axis digital holographic microscopy*, Biomedical Optics Express **10**(4), 2019: 1613-1626. https://doi.org/10.1364/BOE.10.001613

[13] ZHANG J., HUANG L., CHEN B., YAN L., *Accurate extraction of the +1 term spectrum with spurious spectrum elimination in off-axis digital holography*, Optics Express **30**(15), 2022: 28142-28157. https://doi.org/10.1364/OE.464491

[14] HE X., NGUYEN C.V., PRATAP M., ZHENG Y., WANG Y., NISBET D.R., WILLIAMS R.J., RUG M., MAIER A.G., LEE W.M., *Automated Fourier space region-recognition filtering for off-axis digital holographic microscopy*, Biomedical Optics Express **7**(8), 2016: 3111-3123. https://doi.org/10.1364/BOE.7.003111

[15] WENG J., ZHONG J., HU C., *Automatic spatial filtering to obtain the virtual image term in digital holographic microscopy*, Applied Optics **49**(2), 2010: 189-195. https://doi.org/10.1364/AO.49.000189

[16] HONG Y., SHI T., WANG X., ZHANG Y., CHEN K., LIAO G., *Weighted adaptive spatial filtering in digital holographic microscopy*, Optics Communications **382**, 2017: 624-631. https://doi.org/10.1016/j.optcom.2016.08.056

[17] Li J., Dang C., Chen Y., Luo Q., Zhao P., Zhao J., Wang C., *Adaptive spatial filtering based on fuzzy C-means and phase in off-axis digital holographic microscopy*, Optical Engineering **60**(5), 2021: 051207. https://doi.org/10.1117/1.OE.60.5.051207

[18] Dong Z., Wang H., Wang X., *Automatic filtering for zero-order and twin-image elimination in off-axis digital holography*, Optical Engineering **58**(2), 2019: 023112. https://doi.org/10.1117/1.OE.58.2.023112

[19] Bui V., Chang L.C., *Deep learning architectures for hard character classification*, [In] *Proceedings on the International Conference on Artificial Intelligence (ICAI)*, The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2016: 108.

[20] Ronneberger O., Fischer P., Brox T., *U-Net: Convolutional networks for biomedical image segmentation*, [In] Navab N., Hornegger J., Wells W., Frangi A. [Eds.], *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lecture Notes in Computer Science, Vol. 9351, Springer, Cham, 2015: 234-241. https://doi.org/10.1007/978-3-319-24574-4_28

[21] Maas A.L., Hannun A.Y., Ng A.Y., *Rectifier nonlinearities improve neural network acoustic models*, Proceedings of the 30th International Conference on Machine Learning, Atlanta, June 16-21, 2013: 3.

[22] Ioffe S., Szegedy C., *Batch normalization: Accelerating deep network training by reducing internal covariate shift*, [In] Bach F., Blei D. [Eds.], *Proceedings of the 32nd International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 37, 2015: 448-456.

[23] Petrov N.V., Putilin S.E., Chipegin A.A., *Time-resolved image plane off-axis digital holography*, Applied Physics Letters **110**(16), 2017: 161107. https://doi.org/10.1063/1.4981899

[24] Weng J., Li H., Zhang Z., Zhong J., *Design of adaptive spatial filter at uniform standard for automatic analysis of digital holographic microscopy*, Optik **125**(11), 2014: 2633-2637. https://doi.org/10.1016/j.ijleo.2013.11.035

[25] Ali R., Dahl J., *Angular spectrum method for curvilinear arrays: Theory and application to Fourier beamforming*, JASA Express Letters **2**(5), 2022: 052001. https://doi.org/10.1121/10.0010536

[26] Zhang W., Zhang H., Matsushima K., Jin G., *Shifted band-extended angular spectrum method for off-axis diffraction calculation*, Optics Express **29**(7), 2021: 10089-10103. https://doi.org/10.1364/OE.419096

[27] Jaferzadeh K., Fevens T., *HoloPhaseNet: fully automated deep-learning-based hologram reconstruction using a conditional generative adversarial model*, Biomedical Optics Express **13**(7), 2022: 4032-4046. https://doi.org/10.1364/BOE.452645

[28] Liu W., Delikoyun K., Chen Q., Yildiz A., Myo S.K., Kuan W.S., Soong J.T.Y., Cove M.E., Hayden O., Lee H.K., *OAH-Net: a deep neural network for efficient and robust hologram reconstruction for off-axis digital holographic microscopy*, Biomedical Optics Express **16**(3), 2025: 894-909. https://doi.org/10.1364/BOE.547292

[29] Ma S., Liu Q., Yu Y., Luo Y., Wang S., *Quantitative phase imaging in digital holographic microscopy based on image inpainting using a two-stage generative adversarial network*, Optics Express **29**(16), 2021: 24928-24946. https://doi.org/10.1364/OE.430524

[30] Nguyen T., Bui V., Lam V., Raub C.B., Chang L.-C., Nehmetallah G., *Automatic phase aberration compensation for digital holographic microscopy based on deep learning background detection*, Optics Express **25**(13), 2017: 15043-15057. https://doi.org/10.1364/OE.25.015043

[31] Zhang G., Guan T., Shen Z., Wang X., Hu T., Wang D., He Y., Xie N., *Fast phase retrieval in off-axis digital holographic microscopy through deep learning*, Optics Express **26**(15), 2018: 19388-19405. https://doi.org/10.1364/OE.26.019388

Received March 22, 2025
in revised form August 24, 2025