Deep learning-assisted wavefront correction with sparse data for holographic tomography

Li-Chien Lin[1], Chung-Hsuan Huang[2], Yi-Fan Chen[1], Daping Chu[3] and Chau-Jern Cheng[2,*]

[1]Department of Communications Engineering, Feng Chia University, Taiwan
[2]Institute of Electro-Optical Engineering, National Taiwan Normal University, Taiwan
[3]Centre for Photonics Devices and Sensors, Department of Engineering, University of Cambridge, UK
**Email: cjcheng@ntnu.edu.tw*

Abstract

In this paper, a novel approach using deep learning-assisted wavefront correction in beam rotation holographic tomography to acquire three-dimensional images of native biological cell samples is described. With digitally recorded holograms, the wavefront aberration is contained in the reconstructed phases. However, there are large computation costs for compensating the phase aberration during the reconstruction. With the aid of a deep convolution network, we present an effective algorithm on the reconstructed phases with sparse data for active wavefront correction. To accomplish this, we developed a Res-Unet scheme to segment the cell region from its background aberration and a deep regression network for the representation of the aberration on Zernike orthonormal basis. Moreover, a sparse data fitting algorithm was used to predict the Zernike coefficients of whole scanning angles from the collected sparse data. As a result, the proposed algorithm is capable of accurately correcting the background aberration and much faster than the original plain algorithm.

Keywords: Holographic tomography, Deep learning, Wavefront correction

1. Introduction

Due to the rapid development of semiconductor technology, digital holographic microscopy (DHM)[1-3] has been used to acquire quantitative phase information via a modified interferometric architecture. Studies have been conducted on correction of the phase aberration for the DHM system, with signal modeling of the background aberration in sparse domain, such as sparsity in the polynomial representation[4] and Zernike polynomial[5-7]. Recently, DHM has been widely applied to the label-free analysis of living cells and tissues, to examine biophysical characteristics. Such research has extended to the fields of vascular and tumor biology, as well as monitoring of bacterial colonies, which require accurate three-dimensional (3D) reconstruction of a biological micro-sample. However, DHM cannot provide 3D information just with single-shot hologram. It has been found that tomographic imaging can be used to measure the 3D refractive index (RI) distribution inside a biological micro-sample by calculating the integral phase from different illumination directions. In a previous study[8,9], holographic tomography was constructed by Fourier diffraction theorem[10] and verified to provide high-precision measurement of cell RI with spatial resolution reaching the submicron (0.5 um) level of accuracy. This has become the essential

technique for the analysis of 3D biological micro-samples. In studies based on tomographic holography, there have been two common scanning architectures: beam rotation (BR) and sample rotation. Sample rotation requires rotation of the sample holder during the hologram recording process. Even though the sample rotation architecture can achieve 360° scanning of the sample, there are still some major problems to be resolved, such as the mechanical movement of the sample holder and the adverse effect of surrounding fluid on phase accuracy. Moreover, it is difficult to generate accurate rotation for biological samples that are close to circular in shape. In contrast, the BR scheme provides tilted illumination beam for scanning the biological sample at different angles to obtain sectional images. The essential problem with this scheme the limited range of scanning angles, which leads to low resolution in the axial direction. Moreover, it is necessary to have a high numerical aperture (NA) objective lens, limiting the depth of focus and leading to the optical aberration phase, especially for large illumination angles. Based on optical diffraction theory[10-12], missing data is core issue for reconstruction in holographic tomography. To resolve this problem digitally, a regularization term is added to the signal sparsity domain for the phase recovery to achieve better convergence, which includes the constraints of minimization of the total variation, nonnegative, and edge preservation[13,14]. In a previous study, we proposed optical tweezer to trap the biological sample rotating in both x-z and y-z planes[15,16]. In another study[17], combination of sample rotation and BR mechanisms resolved this problem.

In recent years, due to the development of the internet and increases in computing speed, the acquisition and computing of huge amounts of data have led to progress in deep learning. During the 2012 ICSVRC competition, the network architectures of AlexNet[18] and ImageNet[19] were proposed to classify more than one million images, which was a groundbreaking breakthrough (error rate dropped from 25.8% to 16.4%). These architectures deepen the hidden layers in machine learning and use convolution layers to capture image features, to carefully design the activation function of each layer and reduce the possibility of over-fitting. Even with huge amounts of data, convergence can be achieved during the training process. Deep learning is widely used in various fields, such as computer vision, image processing, sound recognition, facial recognition, and medical imaging[20]. At the same time, some new architectures have been developed to reduce the number of calculations and to achieve more accurate learning performance and rapid convergence. One of the more well-known methods is the VGG architecture proposed by an Oxford research team in 2015[21]. The feature extraction layer of VGG is stacked with multiple small 3×3 convolutional layers to replace large kernel convolutions and reduce the number of parameters in the network. Stacking multiple convolution

modules, which combine multiple convolutional layers and pooling layer, simplifies and deepens the network architecture. In that same year, the network architecture of the Inception series[22] was proposed by a Google team. In addition to deepening the depth of the network, this architecture increases the width of the convolutional cognitive layer. The split and merge of different convolution kernels (1×1, 3×3, 5×5, 7×7) allows this architecture to cover the multiscale and multilevel feature extractions in a single Inception module and greatly reduce the network parameters. Although the network depth is 22 layers, there are 12 times fewer parameters than AlexNet's 8 layers. The latest version Inception-V4[23] combines the concept of residue connection to design the Inception module. After 2015, the vanishing gradient problem of deep architecture was resolved using residue connection[24,25] and skip connection[26]. In a residue module, the input of a module is directly added to the output of the two convolutional layers. In this way, the back propagated gradient can be directly passed to the previous layer during backpropagation training. Using the same concept in U-net, the convolution output of the Encoder is directly skip connected to the decoder convolutional layer of the same feature level. In the recent literature[27], the squeeze excitation (SE) network and Convolutional Block Attention Module (CBAM)[28,29] network have given different weights to the feature maps in the convolutional layer, thereby emphasizing channel and spatial features within the probability of activation. For the application of deep convolution network in DHM[30-34], a deep convolutional network is used to quickly perform autofocusing and extended depth of focus processing[35,36], to restore the acquisition distortion for the in-line hologram[37], to reduce the speckle noise[38] from the coherent light source, to reduce the noise led by the low photon count in the light source[39], to correct optical phase aberration[40,41] , to perform the phase unwrapping for the digital hologram[42,43], to reconstruct a super-resolution hologram in DHM system[44], and to implement 3D holographic tomography[45].

In this study, we propose a BR planar scanning mechanism, named BR holographic tomography (BR-HT), to acquire 3D information using a phase-encoding computer-generated hologram (CGH) to generate tilted incidence beam scanning on the transverse plane of the optical axis. The reconstruction of accurate 3D information requires a large number (~10^2) of two-dimensional scanning holograms. Thus, aberration correction becomes a very time-consuming task. Speeding up this process is crucial to promoting the usability of this system. It's true to say that many conventional methods[46,47], for example, Shack-Hartman wavefront sensor, have been proposed to sense wavefront errors for characterizing an imaging system. However, the main goal of the proposed method with deep learning is to solve the wavefront error and aberration problems of beam rotation holographic tomography system under sample testing condition, and

once completed sparse data training with deep learning, it can be applied to speed up whole wavefront correction and imaging processes without manual interference for future automation application. To resolve this problem, we developed a deep sparse data aberration correction (Deep-SDAC) algorithm to accelerate the overall process with the aid of deep convolution network technique. Moreover, a Res-Unet scheme to segment the cell region, a deep regression network to build up the background aberration phase with its Zernike coefficients, and a sparse data fitting algorithm to predict the Zernike coefficients of whole scanning angles from its sparse data were incorporated. With these improvements, the correction with the proposed algorithm was as good as that with the plain algorithm but at a speed which is 13 times faster.

2. Beam rotation holographic tomography (BR-HT) system

The experimental BR-HT system with active wavefront correction (AWC) is shown in Fig. 1. This system was based on improved Mach-Zehnder interferometer architecture with a diode-pumped solid-state (DPSS) laser of 532 nm wavelength as the light source. The laser beam passed through BS_1 before splitting into reference beam and object beam. Phase-only spatial light modulator (SLM) (pixel number: 1920×1080, pixel size: 6.4 μm) was used to display a series of design phase CGHs. The object beam illuminated the phase-only SLM and reflected it back to the objective (MO_1, 100x, NA=0.8) to illuminate the object. The wavefront information of the object was collected by 4-f imaging system (MO_2: 60x, NA=1.2, water-immersion, L_1: 180 mm). Finally, object beam interfered with reference beam as hologram and complementary metal-oxide-semiconductor (CMOS) image sensor (pixel number: 1448×1086; pixel size: 3.45 μm) was used to record the digital hologram. After that, the 3D object information of the BR-HT system was obtained by numerical reconstruction based on the optical diffraction theory[9]. Accordingly, the reconstructed complex wavefront was propagated to its back-focus plane by the digitally simulated Fresnel formula. Then, the Fourier

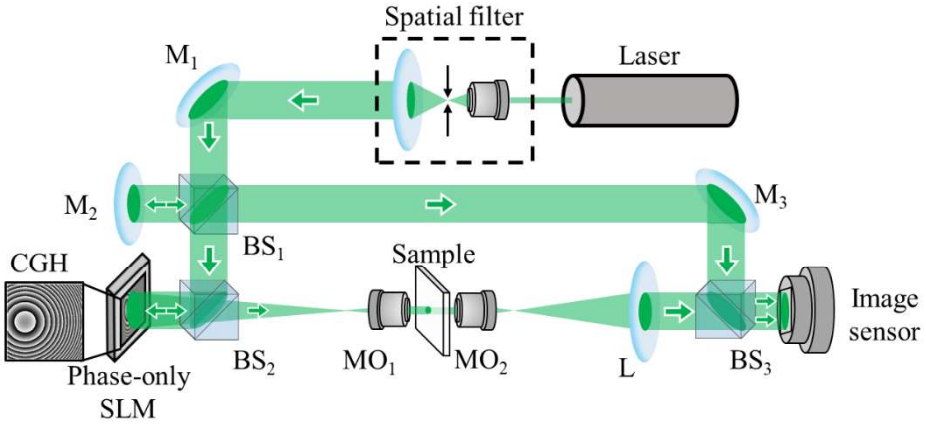

Fig. 1. Experimental setup. M: mirror, BS: beam splitter, SLM: spatial light modulator, CGH: computer-generated hologram, MO: microscope objective, L: lens.

transform of the complex wavefront for each angle was mapped as grid points on the Ewald hemisphere of the object frequency domain. Subsequently, the complex information from all scanning angles was integrated in the frequency domain and inverse 3D Fourier transform was performed to obtain the 3D RI distribution inside a biological micro-sample. Under this procedure, one can use the acquired phase from the system to estimate the background aberration phase and generate its compensated phase. This phase can then compensate the aberration wavefront during the acquisition process.

The flowchart of the AWC used in the BR-HT system is shown in Fig. 2. The processing steps were as follows: (1) Two CGHs cascaded together to encode the phase information of the blazed grating and Fresnel lens. The phase information was encoded in a phase-only SLM. The phase wavefront of blazed grating can be expressed as $\Phi_{\mathrm{BG}} = \frac{i\pi}{\Lambda}(x + y)$, where Λ is grating period and *x,y* are the spatial coordinates. This grating created incident light to scan the 3D sample at different angles and acquire multiple-view holograms. In addition, the phase wavefront of Fresnel lens was expressed as $\Phi_{\mathrm{FL}} = \frac{i\pi}{\lambda f}(x^2 + y^2)$, which led to a convergence spherical wave with focal length *f*. Here, λ was the wavelength of the light source. In this way, we built up a BR mechanism by changing the value of Λ in $\Phi_{\mathrm{BG}} + \Phi_{\mathrm{FL}}$ to generate different illumination angles. In our CGHs designed, the encoded phase contains Fresnel lens (f = 150 mm, 312 Fresnel zones) and blaze. In the CGHs designed on phase-only SLM, the encoded phase consists of a Fresnel lens (f = 150 mm) with about 312 zones and a blazed grating with variable period round from 19.3 μm to 13.5 μm, to span the full scanning range [-45°, 45°] after the angular magnification of a microscopic objective (MO_1, 100x, NA=0.8). According to phase modulation of the SLM device, the phase map is employed to the phase encoding for optimal design of the Fresnel lens and blazing grating to fully utilized SLM phase depth. Further, the fill factor of SLM pixel is over 93%, this only leads to a little diffraction loss and distortion in phase encoding for Fresnel lens and blazing grating. Also, the small nonlinear response of phase modulation over 2π in SLM device might induce some encoding noise, but it is not severely detrimental to the resultant tomographic reconstruction. (2) With the acquired digital hologram, the wavefront information was recorded in the reconstructed phase, which contains cell wavefront distribution, background aberration wavefront, and acquisition noise. For the different scanning angles, the object wavefront passes through different parts of the objective lens with a large extended pupil (NA=1.2). In this way, the aberration wavefront was generated during acquisition, and related to its scanning angles. Our goal is to estimate this aberration wavefront from the acquired phase and generate the compensated wavefront to correct this aberration. (3) For

accurate and efficient estimation of the wavefront aberration, we employed deep convolution network for the cell region detection and used the Zernike polynomial as the mathematical model for the aberration on the acquired phase. In addition, the corresponding Zernike coefficients of full-angle beam rotation were obtained from the aberration distribution with sparse illumination angles using a data fitting algorithm. The compensated aberration distribution, Φ_{AWC}, was the conjugate phase distribution from the proposed Deep-SDAC algorithm. In next section, this algorithm will be discussed in detail. (4) Thin CGH encoded conjugates phase was added to the original CGH series, the phase wavefront of which was encoded as $\Phi_{\mathrm{BG}} + \Phi_{\mathrm{FL}} + \Phi_{\mathrm{AWC}}$ and named AWC CGHs. This information was displayed in phase-only SLM. The wavefront generated from AWC CGHs actively corrected the wavefront aberration for the corresponding illumination angle during the holographic acquisition. So, (5) in the recorded hologram after step (4), the effects of aberration is removed from the numerical reconstructed phase distribution. With a series of AWC CGHs displayed on the phase-only SLM for beam rotations, the results from the recorded holograms are used for further tomographic reconstructions.

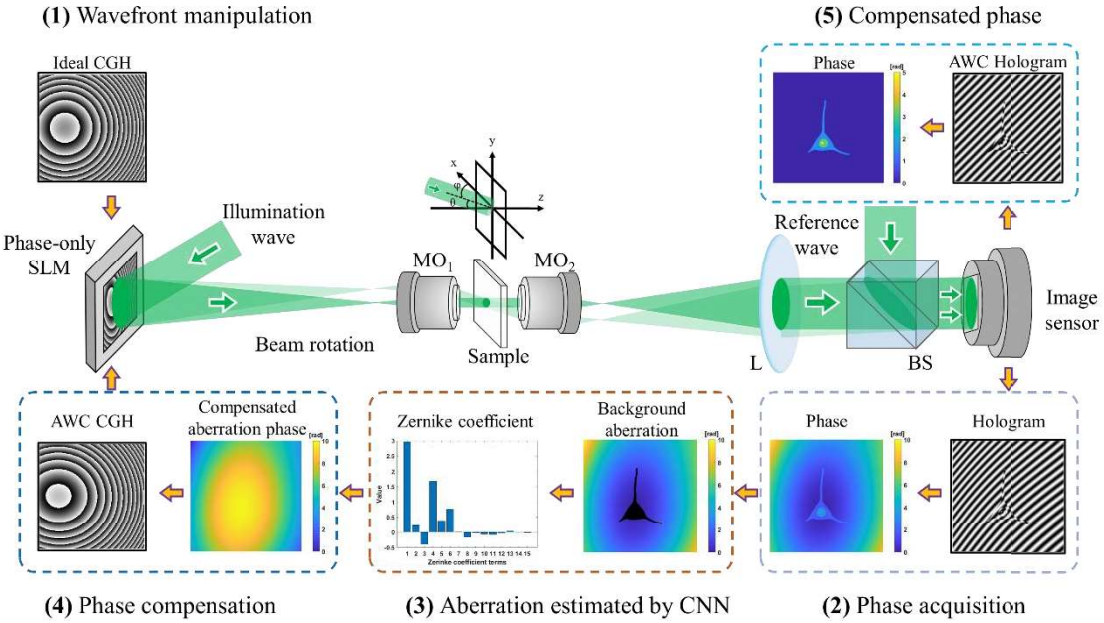

Fig. 2. Flowchart of AWC for BR-HT. CNN: convolutional neural network.

3. Sparse-data aberration correction by deep learning (Deep-SDAC) algorithm

To obtain accurate 3D information of the measured object, hundreds of scanning angle 2D holograms are required. However, correcting the phase aberration holograms for all scanning angles is computationally intensive. In this paper, we propose a sparse-data aberration correction algorithm by Deep-SDAC algorithm to accelerate the overall process, as indicated in Fig. 3. In the algorithm, the acquired multi-scanning angle holograms were obtained by the previously described BR-HT system. The phase

information was firstly unwrapped and an initial tilted parameter was estimated. Since the scanning angle was given, it was easy to obtain this tilt parameter. Next, a deep convolution network Res-Unet was used to segment cell region of the phase image, incorporating residue connection and U-net architecture for the segmentation. Our goal was to estimate the background aberration image. Thus, the phase corresponding to the cell object was removed. Once the cell region was detected, the background phase of the cell region could be replaced by the interpolated data from the non-cell (background) region. Accordingly, the phase aberration image could be obtained from its background with some acquisition noise. After that, we employed orthonormal Zernike basis as the theoretical model for the background phase aberration. Based on orthogonality, there exists unique representation for the phase aberration and this representation is robust to the random noise. In the next step, we developed a deep regression network, which cascaded the Inception-ResV2 and full-connection network, to establish the relationship between the phase aberration image and its Zernike coefficients. Once the Zernike coefficients were obtained, the aberration compensation phase was generated, which was the conjugated phase image generated by the resultant Zernike coefficients. To this end, the aberration compensation phase was added to AWC CGH hologram and encoded in the phase-only SLM. Then, the object wave after the AWC CGH constructed an active illumination wavefront to remove the system aberration during the holographic recording process. Even though the overall process, as shown in Fig. 3, is efficiency. However, the process must be implemented in all 244 holograms with different scanning angles, and thus it required a large amount of computation. In this study, we constructed a sparse data fitting method which uses Zernike coefficients of the sparse angles in 15-degree scanning steps, to predict all the scanning angles, which were in 1.5-degree scanning steps for the whole scanning range [-45°, 45°]. Only the reconstruction phase of the sparse angles went through the overall process in Fig. 3, while the other Zernike coefficients were evaluated only by the sparse data fitting method. In this way, computation cost of the algorithm was greatly reduced. In the following sections, we introduce two deep convolution networks for cell region segmentation and Zernike coefficient estimation, as well as sparse data fitting technique for the Zernike coefficients.

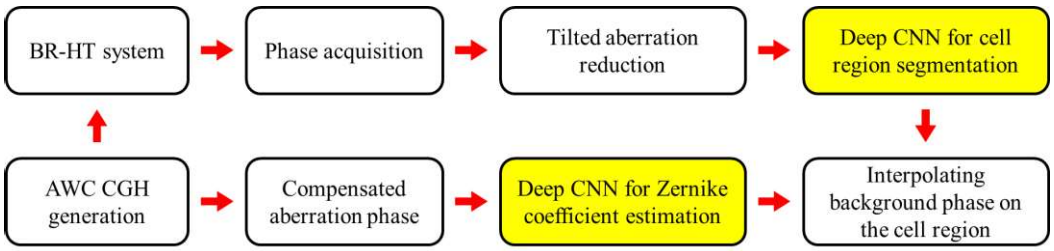

Fig. 3. The proposed Deep-SDAC algorithm. Notes that the color-marked blocks are implemented by deep learning architectures and the others are processing by digital programs.

3.1 Deep learning Res-Unet for cell region segmentation

The acquired phase information of the BR-HT system consisted of three parts: cell wavefront distribution, background aberration, and system noise. Our purpose was to remove the influence of the optical aberration by BR scheme. For this purpose, a fast process was required to separate the cell region from the background in the phase image. We utilized Res-Unet deep convolution network, which combines two network architectures: deep U-net and residue connection. The deep U-net architecture is a mature technique for the segmentation of 2D images[26]. With this technique, there is a question of whether the neural network can extract the feature maps of the cell wavefront distribution. Deep learning stacks up large numbers of convolutional layers to obtain multi-scale features, from fine-grid to coarse-grid. The fine-grid features corresponded to the low-level features, such as pixel color, line, edge, and corner, whereas the high-level integration of these low-level features was used to construct complex, abstract, and close to visually recognizable features. In contrast to digital image, the cell phase or the phase aberration possesses less sharp edge information and more mid-level (two or three upper levels) features. Thus, four or five convolutional blocks are suitable for extracting the cell wavefront distribution edge information. In addition, the residue connection[2,24] was adopted to avoid the problem of vanishing gradient during backpropagation training.

As indicated in Fig. 4, the Res-Unet architecture was established for cell region segmentation, which is an image-in image-out structure. There are two strategies to deal with the input data. First, to overcome the overfitting problem, we employed data argumentation to increase the number of training images. Next, multiple inputs, the original phase image and its gradient, were fed into the network. In this way, the size

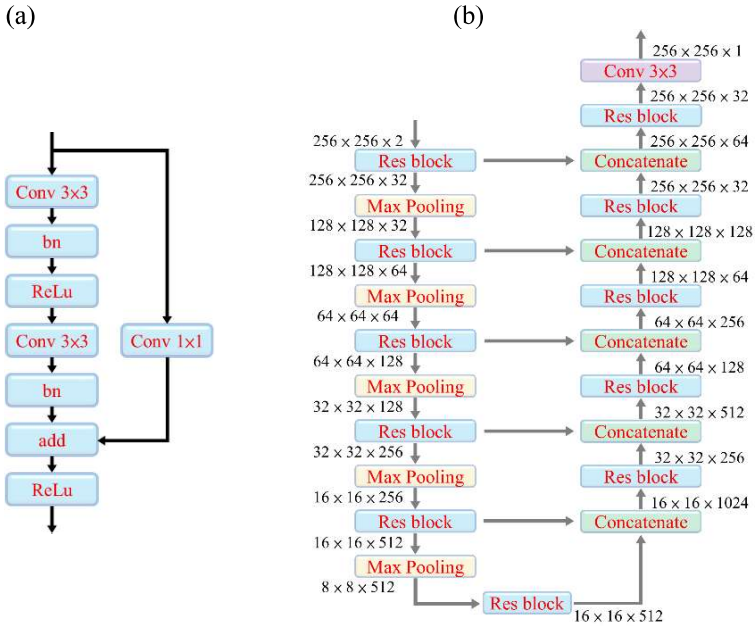

Fig. 4. (a) Residue block (Res-Bk) and (b) The Res-Unet architecture.

of the inputs was 256×256×2, which is resize from its the original recoded hologram (1448×1086 pixels) for the limited computer memory in network architecture. We found that the curvature of the cell was valuable for properly recognizing the cell boundary. Gradient image was the most convenient way to access this information. The network architecture consisted of three modules: encoding, intermediate, and decoding. The encoding module extracted the feature maps from fine-grid to coarse-grid by stacking a series of convolution layers. The decoding part recovered the pixel size of the feature map to that of the original input. Convolution block (CB) consisted of two 3×3 convolutional layers with strike of 2. The basis unit of Res-Unet was a residue block (Res-Bk). Here, we combined CB block with residue connection, letting x, y denote the input and output of the block. Then we performed the residue connection by $y = x + CB(x)$, where $CB(x)$ means the convolutional result after CB block with input x, as indicated in Fig. 4(a). In the encoding module, we employed the depth-wise separable convolution to reduce the parameters of the network. Moreover, the transposed convolution[48] restored the size of the feature maps in the decoding pass. As shown in Fig. 4(b), the encoding pass was comprised of a series of layers with structure 5×(Res-Bk – Max pooling layer) modules. Similarly, the decoding module encompassed the 5×(concatenate layer – Res-Bk) modules. The intermediate module served as the interconnection between the encoding and decoding modules, which connected one Res-Bk block. In this architecture, the skip connection was used to incorporate the feature maps in the encoding module into the corresponding feature level in the decoding module. In this way, the gradient flow was properly backpropagated through the encoding path for fast convergence. For the output, the Soft-max activation function was used for segmentation, with $\{f_i, S_i\}$ denoting the input and the desired segmented image (binary). In the training process, the input image was fed into the deep network. Then, the resultant output of the network $O_i = Net(f_i, W)$ was considered the probability map of the image pixels belonging to the cell region, where W are the weighted parameters of the network. As a result, a single threshold segmented the output to a binary image and the resultant binary output T_i was compared with the desired segmented image S_i for defining the error function. In the experiments, we found that, other than the sample object and the background area, adding edge as the third region can improve the segmentation. Thus, we utilized the multiclass segmentation, which included three objects: the sample, the background, and the edge. In advance, under normal circumstances, the edge area was a very small, normally less than 5 %, which led to the imbalanced multiclass segmentation problem. In this case, the dice loss (DL) can achieve most accurate segmentation result, and defined as

$$\mathrm{DL} = 1 - \frac{2\times(S_i \cap T_i)}{|S_i| + |T_i|} \tag{1}$$

where $S_i \cap T_i$ denotes the intersection of these two binary sets and $|\bullet|$ is the area of the set. During training, the gradient values of the error function with respect to the W for each layer were evaluated and backpropagated from the output layer to the input layer successively. The gradient values were then used to adjust the W of the network to minimize the loss function.

3.2 Deep regression network for Zernike coefficient estimation

In this section, Zernike polynomial was employed to construct an orthonormal basis[49] for the aberration phase. The orthonormal basis provided a simple way to transform the aberration phase image and its coefficients using integration operation. To increase the efficiency of the algorithm, we developed a deep regression network to establish the relationship between aberration phase image and its Zernike coefficient. In BR-HT system, the tilted scanning beam passed through different apertures of the objective lens with large extended pupil (NA=1.2, water-immersion) at different angles. Thus, phase aberration depended on scanning angle. The unflattened surface of the SLM element or the varying RI of the microslide generated the aberration phase. In the experiments, 15 terms of Zernike coefficients were sufficient for modeling the phase aberration of the proposed BR-HT system, which included tilted, astigmatism, coma, trefoil, and spherical aberrations. The phase aberration model can be expressed as

$$\mathrm{P(x,y)} = \sum_{k=1}^{15} a_k \, \varphi_k(x,y) \qquad (2)$$

where $\varphi_k(x,y)$ is the orthonormal Zernike basis and a_k are the corresponding Zernike coefficients. Since the Zernike basis was orthonormal, the phase aberration was uniquely represented by its Zernike coefficients. This relationship was established by a deep regression neural network. Given the input phase aberration image, there is corresponding output of Zernike coefficient. For the accurate estimation of Zernike

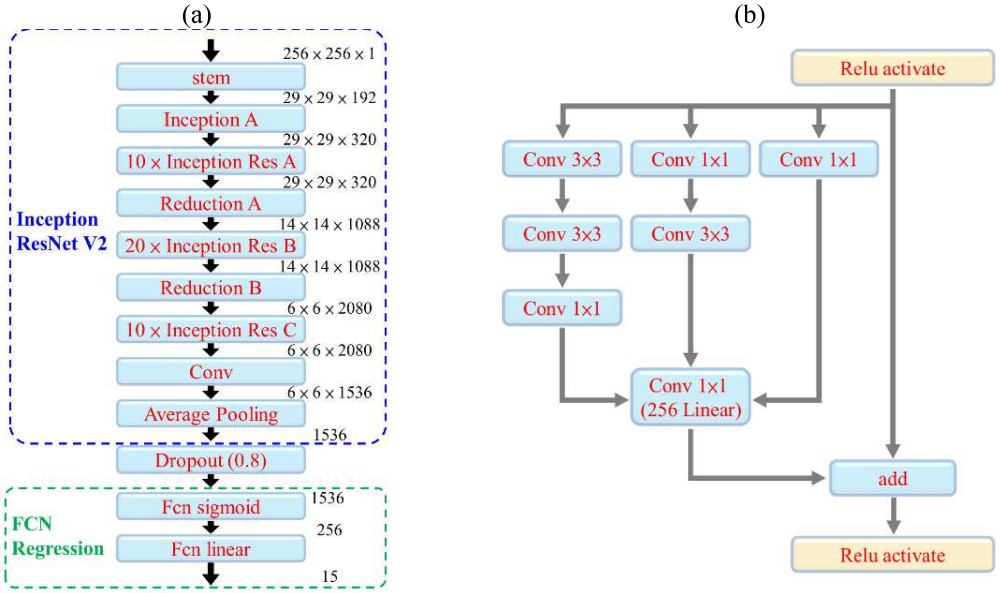

Fig. 5. (a) The deep regression network architecture and (b) Inception Res A module.

coefficients, the deep regression network should be able to extract the surface feature maps from different aberrations, such as tilted, astigmatism, coma, trefoil, and spherical. Recently, deep convolutional networks have matured enough to extract multiscale feature maps, which include tilted, spherical, or low-order polynomial surfaces. Thus, we combined these features for the estimation of the Zernike coefficients. Here, we employed the deep convolution network, Inception ResNet-V2, which has performed the best in terms of feature extraction in recent years. The basic setup of the deep regression network combined Inception ResNet-V2 for feature extraction and a series of full-connection networks for coefficient estimation. To obtain exact Zernike coefficients, large numbers of training images were generated to train the neuron parameters. The overall scheme of the proposed deep regression network is shown in Fig. 5. As indicated, we adopted an Inception ResNet-V2 network to extract the surface feature maps of the aberration phase, with flattened feature maps imported into the next deep regression module, which was composed of a series of full-connection layers.

Inception ResNet-V2[23] incorporates two main techniques in deep convolutional network: Inception block and residue connection. Inception block, as developed by Google-Net Group, combines multi-scale convolution of different sizes, 1×1, 3×3, 5×5, by split, transform, and merge, to deepen and widen the architecture. In this way, spatial information of different scales (at fine and coarse grid levels) can be captured in a single Inception block. As an example, in Fig. 5(b) of the scheme for 35×35 grid (Inception Res A) module, the input was split into three convolutional branches to obtain multi-scale feature maps and then merged into the output. In this scheme, residue connection was also performed, with the input directly added to the merged output of the convolution. Meanwhile, with the Inception block there was emphasis on learning efficiency, with 1×1 convolution before a large kernel size convolution. Due to merging of the feature maps in the output, some irrelevant features were ignored. In this way, there was effective training and fast convergence. In recent years, residue connection has inherently contributed to the training of a very deep network architecture. Since Inception ResNet-V2 tends to be very deep, it was natural to incorporate the filter cascade of the Inception block with residue connection. As shown in Fig. 5(a), different modules were combined, including Stem, Inception Res A, Reduction A, Inception Res B, Reduction B, and Inception Res A, and there were serial cascades multiple times. For example, the notation 10x Inception Res A means that 10 convolution modules (Inception Res A) are connected in series. The Reduction modules employed the convolution with strike of 2 to reduce the spatial grid to its lower coarse grid. The details of these Inception modules can be found in ref. 23

Given an input aberration phase $P(x, y)$, the output of network can be represented as

$$a_k = Net(P(x, y), W),\ k = 1,2, \dots ,15 \qquad (3)$$

During training, the corresponding desired Zernike coefficients z_k were employed to evaluate the root mean square error for each training step, which was $\mathrm{RMSE} = \sqrt{\sum_{k=1}^{N} \|a_k - z_k\|^2}$, where N=15. The training process was used to minimize this error in each step.

3.3 Sparse data fitting for the Zernike coefficients at different angles

In order to obtain complete 3D information of the living cell, we acquired 244 2D holograms at different scanning angles from the system. Optical diffraction theory was utilized to restore 3D living cell data. Even though the overall aberration correction process in Fig. 3 was computationally effective, it had to be implemented in all acquired holograms. Correcting such a large number of holograms is still costly in terms of reconstruction. For sparse data fitting, we used Zernike coefficients from the sparse scanning angles, for example 15° steps from the whole scanning range [-45°, 45°], to predict the coefficients of the other scanning angles. For this problem, a mathematical model was required for the relationships of the Zernike coefficients at different angles. By assuming that the optical axis of the BR-HT system was the z-axis and with the tilted illumination beam scanned in the x-direction at the angle of θ, we derived the mathematical expression which represents the relationship of the 2nd Zernike coefficient at different angles. From the polar representation of the 2nd Zernike basis, the aberration led by the tilted scanning angle could be written as

$$c_1 \sin(\theta) = a_2(\theta) \cdot x \tag{4}$$

where $a_2(\theta)$ denotes the 2nd Zernike coefficient depending on the scanning angle and $c_1$is a constant. Here, the right expression $\mathrm{x} = \cos(\theta)$ of this equation is the Zernike basis. As a result, the relationship of the 2nd Zernike coefficient with respect to the tilted scanning angle can be considered $a_2(\theta) = c_2 \tan(\theta)$, where $c_2$is a constant. In practice, a small perturbation can be added for the tilted scanning angle. Accordingly, we altered the expression as $a_2(\theta) = c_2 \tan(\theta + \Delta\theta)$. Taylor expansion led to the first order approximation of this expression with first two terms

$$a_2(\theta) = k_1 \tan(\theta) + k_2 \, sec^2(\theta) + \cdots \tag{5}$$

of a prediction model for the 2nd Zernike coefficient at different tilted scanning angles. With this model, we developed a data fitting algorithm for accurately predicting this coefficient at different angles using only sparse angle data. Predictive models for other coefficients can be derived in a similar way to utilize the polar coordinates of the Zernike basis in ref. 45

4. Experimental results

In this study, we determined the aberration correction performance of the proposed Deep-SDAC algorithm using living neuroblastomas (SY-SH5Y) as our biological

samples in BR-HT system to obtain the 3D RI of these cells. The neuroblastomas, which were comprised of cell body and synaptic structures, were cultured in phosphate buffered saline (PBS, RI: 1.33). Here, we propose a BR planar scanning mechanism to acquire 3D information of the biological samples, using a phase-encoding CGH to generate tilted incidence beam scanning on the transverse plane of the optical axis, named BR holographic tomography. In this scheme, the hologram is recoded 1.5-degree steps of the optical axis (z-axis). The maximum of illumination angle is ± 45°. The beam rotation directions are the angles of 0°, 45°, 90°, and 135° along x-axis on x-y plane. The total number of recorded holograms during beam rotation is 244 images. Computation was carried out with AMD Rysen5 2600X CPU, GTX 1060 3GB graphic processing unit, Windows 10 operating system, MATLAB program 2019 version, and deep convolutional network with Tensorflow 2.0 & Keras.

Figure 6 shows the resultant aberration correction process of the reconstructed unwrapped phase with single biological cell for the BR-HT system. In Fig. 6, the encoded phase of the ideal CGH (1920×1080 pixels) displayed on phase-only SLM was indicated, as previous stated this element was used to generate a dynamic focal spot by a convergence spherical wave and frequency conversion grating to record as multiple view holograms. With a known tilted angle, it is easy to initially estimate the tilted parameters and remove the tilted aberration. In additions, the phase recovery and unwrapping algorithm can achieve the reconstructed phase (1448×1086 pixels), as shown in Fig. 6(b). Our purpose was to correct the background aberration. Thus, cell information was removed during the correction process. The proposed deep convolution network Res-Unet was used to segment the cell region, with the resultant

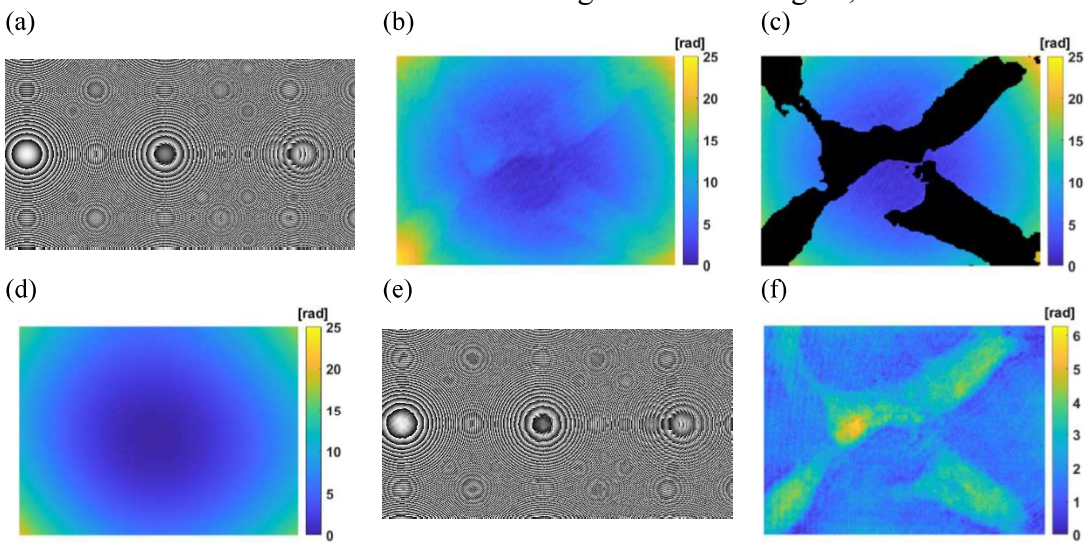

Fig. 6. Resultant images of the proposed Deep-SDAC algorithm. (a) Ideal CGH displayed on a phase-only SLM, (b) Reconstructed phase image includes object information and background aberration, (c) Cell segmentation result of (b) by Res-Unet, (d) Aberration phase from Zernike coefficient estimation by deep regression network, (e) AWC CGH by adding the conjugate of aberration phase to (a), and (f) Active compensated phase result.

segmented cell region (black region) overlapped to the background information (1448×1086 pixels), as shown in Fig. 6(c). As indicated, other than cell information, the input phase of the Res-Unet network also contained aberration and acquisition noise. In order to collect a large number of these phase images, we acquired training phases directly from the BR-HT system with scanning at different angles. Given the input phase of the Res-Unet network in Fig. 6(b), there was good cell segmentation, as shown in Fig. 6(c). As a result, a proper trained network can distinguish smoothness on the cell wavefront distribution from that of the aberration surface and noise. It achieves good segmentation result, even if the cell region is not clear to the human eye. To remove the cell information, the background phase (1448×1086 pixels) contained only the aberration with some noise, as shown in Fig. 6(d). The background phase was then input into a deep regression network for the estimation of the 15 terms of Zernike coefficient. Finally, the estimated Zernike coefficients were used to obtain the aberration compensated phase. The conjugation of this aberration phase was added to the encoded phase of the ideal CGH in Fig. 6(a) and led the resulting phase in AWC CGH (1920×1080 pixels), as shown in Fig. 6(e). Given the active compensated wavefront generated from AWC CGH, the hologram acquired in BR-HT system led the compensated phase result (1448×1086 pixels), as indicated in Fig. 6(f).

Next, the training and testing performances of two deep convolution networks are illustrated. In general, deep convolution network requires a large number of neuron parameters and the training process is based on stochastic optimization techniques, with different sources of randomness, optimization parameters, and regularization procedures, which impact on the results. In our experiments, AdaGrad optimization was employed, in which the learning rate of each neuron parameter was multiplied by the reciprocal of a weighted value, which was the square sum of its historical squared gradient values. In this way, fast changing parameters rapidly decreased the learning rate. On the contrary, there was a small decrease in learning rate for the small update parameters. The optimization parameters included batch size, learning rate, and dropout ratio. We employed a number of experimental trials to make reasonable optimization choices.

The training data for Res-Unet was acquired from the real data of BR-HT system for the experimental nerve cells. We generated about 488 phase images from the BR-HT system and utilized data argumentation to the generated phase images, such as rotation, shear, and horizontal flipping to obtain about 2000 large datasets for training. We separated the datasets into 1800 images for training data and 200 images for testing data. During training, the learning rate was initially set at 0.0001, initial accumulator value=0.1, with batch size of 12. Thus, there were 150 training steps for each epoch. In this network, there were 3.8 million neuron parameters to be trained. The training and

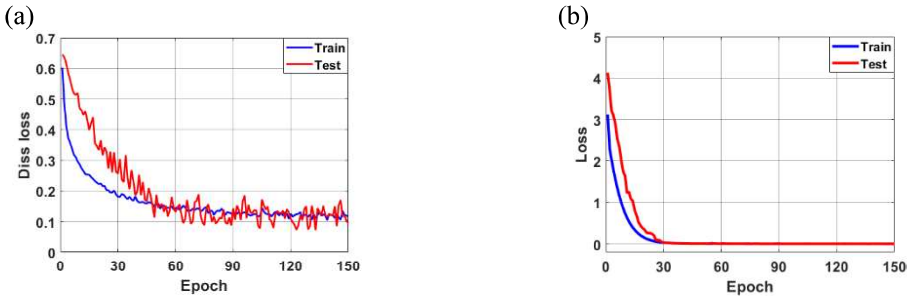

Fig. 7. Training and testing loss of the (a) Res-Unet for cell segmentation and (b) Deep regression network for Zernike coefficient estimation.

testing losses are indicated in Fig. 7(a) for 140 epochs where the loss function is defined as the dice loss error function in Eq. (1). The loss value decreased quickly from about 0.65 to 0.09 and the testing loss demonstrated randomly turbulent decline during the first 50 epochs. After that, the loss value slightly decreased until stable. Due to a small number of testing datasets, there were still some small perturbations in the testing loss during stable status. Finally, the testing loss reached its minimum value of 0.0697.

We also examined the performance of the deep regression network as described in previous section. With a background aberration image, this model accurately estimated its 15 Zernike coefficients. The background aberration was highly dependent on the optical parameters of the BR-HT system, such as the wavelength, focal distance, RI of the microslide, and pixel resolution of the image sensor. Once the system was set up, these parameters could either be fixed or were variable within a small dynamic range. Then, a few hundred background (cell free) phase images at different angles were obtained and the Zernike coefficients were determined. As a result, we attained the dynamic range of all the Zernike coefficients. Accordingly, the simulated aberration phase database was generated by randomly selecting Zernike coefficients inside this range. In this way, we created 20000 background aberration images for training this model, 16000 for training and 4000 for testing. The input of this network was 256×256 background aberration image and the output was 15 Zernike coefficients. The training process used 3200 training steps for each epoch, with batch size of 5 and learning rate of 0.0001. In this network, there were 2.1 million neuron parameters to be trained. We used root mean square error (RMSE) as the loss function of this model, as defined in previous section. The training and testing losses with respect to the number of epochs are illustrated in Fig. 7(b). From this figure, loss function convergence was fast and became stable during the first 50 epochs. After 50 epochs, either training loss or testing loss became stable. As a result, the deep regression network could accurately estimate the Zernike coefficients, with RMSE between the estimated and the desired coefficients very close to zero (0.0013).

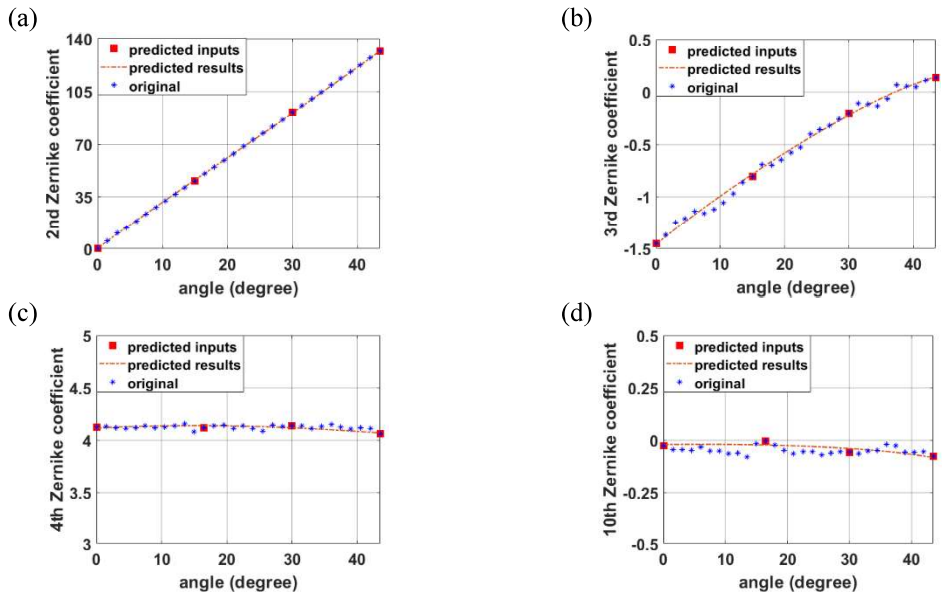

Fig. 8. Performance of the sparse data fitting: Zernike coefficients with respect to its scanning angle. (a) 2nd order. RMSE=0.37 in range [0, 132], (b) 3rd order. RMSE=0.041 in range [-1.5, 0.2], (c) 4th order. RMSE=0.023 in range [4.06, 4.14], and (d) 10th order. RMSE=0.026 in range [-0.082, -0.044].

The performance of the sparse data fitting algorithm, introduced in previous section, was evaluated. In practice, the difference in the dynamic range of these coefficients is very large. The contribution of the aberration phase is mostly from the low-order Zernike coefficient, with radial degree less than or equal to 2 and some of the high order coefficients very close to zero. For comparison, we used a MATLAB program to implement the whole correction process, as shown in Fig. 3, and calculated the Zernike coefficients for every scanning angle of the aberration phase. The result was the ground-true Zernike coefficient. Performance was evaluated using the RMSE between the predicted value and its ground-true value. The sparse data fitting algorithm, based on the derived model in previous section, was employed to predict the coefficients of all the scanning angles using only the sparse angle data; 15 degrees of step over whole scanning range. The prediction results were shown in Fig. 8. Here, we only present the prediction results of 4 coefficients in positive angles, the 2^{nd} (tilted-x, RMSE=0.37 in range [0, 132]), 3^{rd} (tilted-y, RMSE=0.0041 in range [-1.5, 0.2]), 4^{th} (defocus, RMSE=0.023 in range [4.06, 4.14]), and 10^{th} (Trefoil-x, RMSE=0.0026 in range [-0.082 -0.044]) terms. For example, in the 2nd coefficient, the sparse data was from the coefficients of 0°, 15°, 30°, and 43.5°scanning angles. Instead of 45°, we used 43.5° as the sparse input, since $\tan(45°)$ in the prediction model is undefined. From this result, it is obvious that, given a suitable model for the relationship of coefficient with respectto its scanning angle, one can accurately predict the Zernike coefficient from its sparse data and save computation costs in the overall correction process.

The original phase images contained cell, aberration, and noise information. The proposed algorithm removed the aberration. The corrected background phase was still

not flat, as it contained acquisition noise. In this case, ground-true data for the performance evaluation of the proposed Deep-SDAC algorithm was required. With MATLAB program, the plain algorithm was used to implement the overall aberration correction process, as shown in Fig. 3, and replaced the deep convolution algorithms by commonly used image algorithms, where the watershed[50] image segmentation was used to segment the cell region and the numerical integration was employed for the Zernike coefficient estimation. Since, the Zernike basis was orthonormal with unique transformation between the Zernike coefficient and its background aberration phase. Therefore, the Zernike coefficient estimated from the process shown in Fig. 3 was an exact representation and it is robust to random noise. The aberration correction result of the proposed deep-SDAC algorithm was examined in comparison with that of the plain algorithm, which was considered ground-true data.

The aberration corrected results of these two algorithms was shown in Fig. 9. Figure 9(b) and Figure 9(c) present the results of the plain algorithm and Deep-SADC algorithm, respectively, where the original phase image without any aberration compensation was shown in Fig. 9(a). To evaluate the smoothness of the background surface, we selected a background portion (enclosed by square mark) of these results, as shown in Fig. 9(d), (e) and (f), respectively. Here, standard deviation (STD) of the surface on the selected background region was used to measure the smoothness of the background phase. The STD were 2.06, 0.28 and 0.23 in radian for original phase, the plain algorithm and the Deep-SDAC algorithm respectively. From this result, we observed that the proposed Deep-SDAC algorithm was capable of accurately correcting the background aberration in comparing with the original plain algorithm.

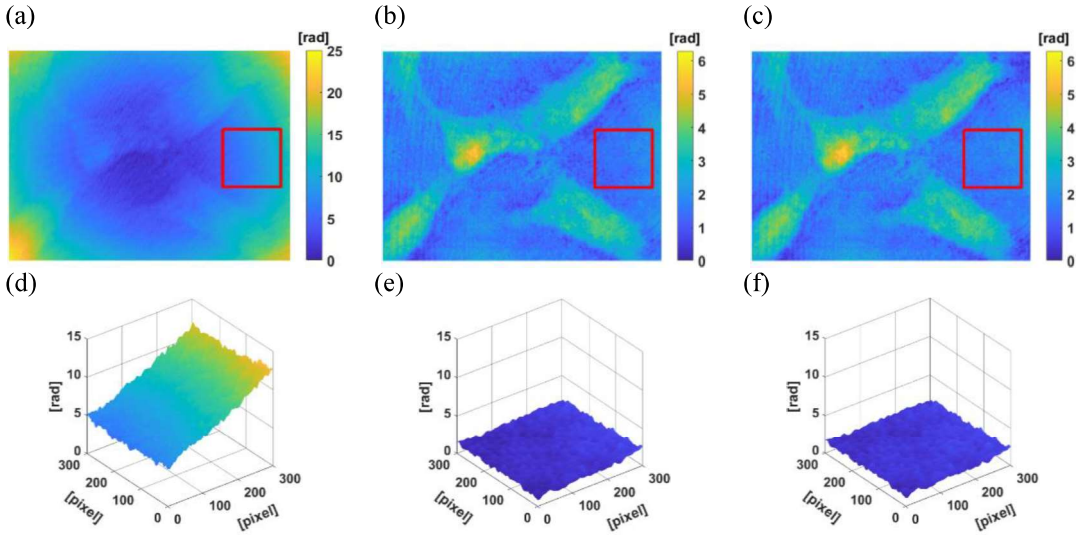

Fig. 9. Comparison of aberration compensation results. (a) Phase image without aberration compensation, (b) Aberration compensation result by the Plain algorithm, (c) Aberration compensation result by the Deep-SDAC, (d), (e), and (f) are the 3D images of red block regions in (a), (b), and (c), respectively. STD: (d): 2.06 rad, (e): 0.28 rad, and (f): 0.23 rad.

To verify the efficiency, all the computation times are shown in Table 1 for the methods used to accelerate the aberration correction algorithms presented in the previous section. As indicated, both cell segmentation and calculation of Zernike coefficients were expensive in terms of computation costs. On the contrary, deep learning algorithm sped up the processing time. In addition, the sparse data fitting allowed us to evaluate the Zernike coefficients only for the sparse angles and use them to predict the coefficients of the other scanning angles. As a result, the computation was about 13 times faster than that of the plain algorithm for correcting the aberration of the whole 244 holograms, demonstrating the efficiency of the proposed algorithm.

Table 1. Effectiveness of the proposed algorithm: Shorthand CRS: cell region segmentation, ZCE: Zernike coefficient estimation, SDF: sparse data fitting.

	Computation time in second (244 holograms)			
Processing step	CRS	ZCE	SDF	Overall process
Plain algorithm (ground-true)	369.7	6.24	none	561
Proposed Deep-SDAC algorithm	21.7	4.6	8.4	42.9

The three-dimensional tomographic reconstruction result of a single living neuroblastomas cell based on Deep-SDAC algorithm was shown in Fig. 10. According to the similar compensated phase of the reconstructed images from different scanning angles, the 3D refractive index distribution of the cell sample from tomographic reconstruction can be estimated with a reasonable accuracy of approximately 0.006 for both algorithms. As a result, the proposed method could be used as a tool for generating high speed and high quality of 3D tomographic image of label-free living cell.

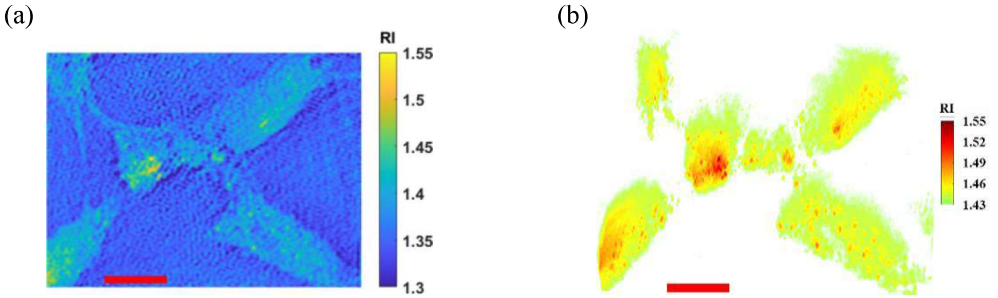

Fig. 10. Reconstructed BR-HT images based on the Deep-SDAC algorithm. (a) Slice image in x-y direction and (b) 3D view image. Scale bar: 10 μm

5. Conclusion

The proposed BR-HT imaging system was successfully used to acquire 3D RI distribution of a label-free living neuroblastoma cell, utilizing a planar BR scanning mechanism. The computational cost of the conventional aberration correction in the reconstruction was too high for practical use. To accelerate the aberration correction in

the BR-HT system, deep convolution network was introduced to segment the cell region and estimate the Zernike coefficients of the background aberration. Finally, a sparse data fitting algorithm was used to obtain the Zernike coefficients of the whole scanning range from its sparse data. The computation speed was up to 13 times faster than that of the original plain algorithm with the same correction performance. This greatly improves the practicality of the optical system.

Author's contribution

CJ and LC conceive the idea and supervised the project. CH and YF conduct the experiments and data analysis. D revise the manuscript. All the authors discuss the results and equally contributed to the manuscript.

Author statement

Chau-Jern Cheng and Li-Chien Lin conceive the idea and supervised the project.

Chung-Hsuan Huang and Yi-Fan Chen conducted the experiments and data analysis.

Daping Chu discuss the results and revise the manuscript.

Declaration of competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is financially supported by the Ministry of Science and Technology, Taiwan: MOST 109-2221-E-035-074, MOST 108-2221-E-003-019-MY3, MOST 107-2923-E-003-001-MY3, and MOST 110-2218-E-011009-MBK.

References

1. Mann CJ, Yu L, Kim MK. Movies of cellular and sub-cellular motion by digital holographic microscopy. Biomed. Eng. Online 2006;5:21.
2. Charrière F, Pavillon N, Colomb T, Depeursinge C, Heger TJ, Mitchell EAD, Marquet P, Rappaz B. Living specimen tomography by digital holography: morphometry of testate amoeba. Opt. Express 2006;14:7005-13.
3. Kim MK. Digital Holographic Microscopy. (Springer, 2011).
4. Ren Z, Zhao J, Lam EY. Automatic compensation of phase aberration in digital holographic microscopy based on sparse optimization. APL photon. 2019;4:110808.
5. Liu S, Lian Q, Qing Y, Xu Z. Automatic phase aberration compensation for digital holographic microscopy based on phase variation minimization. Opt. Lett. 2018;43:1870–3.
6. Liu S, Zhu W, Xu Z, Gao M. Automatic and robust phase aberration compensation for digital holographic microscopy based on minimizing total standard variation. Opt. Lasers Eng. 2020;134:106276.
7. Huang L., Yan L, Chen B, Zhou Y, Yang T. Phase aberration compensation of digital holographic microscopy with curve fitting preprocessing and automatic background segmentation for microstructure testing. Opt. Comm. 2020;462:125311.
8. Choi W, Fang-Yen C, Badizadegan K, Oh S, Lue N, Dasari RR, Feld MS. Tomographic phase microscopy. Nat. Methods 2007;4:717-9.

9. Balasubramani V, Kuś A, Tu HY, Cheng CJ, Baczewska M, Krauze W, Kujawińska M. Holographic tomography: techniques and biomedical applications [Invited]. Appl. Opt. 2021;10: 65-80.
10. Kak AC, Slaney M. Principles of Computerized Tomographic Imaging. (IEEE, 1988).
11. Balasubramani V, Tu HY, Lai XJ, Cheng CJ. Adaptive wavefront correction structured illumination holographic tomography. Sci. Rep 2019;9:10489.
12. Deng Y, Huang CH, Vinoth B, Chu D, Lai XJ, Cheng CJ. A compact synthetic aperture digital holographic microscope with mechanical movement-free beam scanning and optimized active aberration compensation for isotropic resolution enhancement. Opt. Lasers Eng. 2020;134:106251.
13. Sung Y, Choi W, Lue N, Dasari RR, Yaqoob Z. Stain-free quantification of chromosomes in live cells using regularized tomographic phase microscopy. PLoS ONE 2012;7:49502.
14. Lim JW, Lee KR, Jin KH, Shin S, Lee SE, Park YK, Ye JC. Comparative study of iterative reconstruction algorithms for missing cone problems in optical diffraction tomography. Opt. Express 2015;23:16933-48.
15. Balasubramani V, Montresor S, Tu HY, Huang CH, Picart P, Cheng CJ. Influence of noise-reduction techniques in sparse-data sample rotation tomographic imaging. Appl. Opt. 2021;10:81-7.
16. Lin YC, Chen HC, Tu HY, Liu CY, Cheng CJ. Optically driven full-angle sample rotation for tomographic imaging in digital holographic microscopy. Opt. Lett. 2017;42:1321–4.
17. Balasubramani V, Lai XJ, Lin YC, Tu HY, Cheng CJ. Integrated dual-tomography for refractive index analysis of free-floating single living cell with isotropic superresolution. Sci. Rep. 2018;8:5943.
18. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. Adv. Neural Inf. Process. Syst. 2012;1–9.
19. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2009;248-55.
20. Liu W, Wang Z, Liu X, Zeng N, Liu Y, Alsaadi FE. A survey of deep neural network architectures and their applications. Neurocomputing 2017;234:11–26.
21. Simonyan K, Zisserman A. Very deep convolution networks for large-scale image recognition. ICLR 2015.
22. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going Deeper with Convolutions. IEEE CVPR 2015;1-9.
23. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. 2016. arXiv:1602.07261.
24. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2015. arXiv:1512.03385.
25. Xie S, Girshick R, Dollar P, Tu Z, He K. Aggregated Residual Transformations for Deep Neural Networks. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017;1492-500.
26. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015. arXiv:1505.04597.
27. Hu J, Shen L, Sun G. Squeeze-and-Excitation Networks. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2018;7132-41.
28. Khan A, Sohail A, Ali A. A New Channel Boosted Convolutional Neural Network using Transfer Learning. 2018. arXiv:1804.08528.
29. Woo S, Park J. Lee JY, Kweon IS. CBAM: Convolutional Block Attention Module. 2018. arXiv:1807.06521.
30. Rivenson Y, Wu Y, Ozcan A. Deep learning in holography and coherent imaging. Light Sci. Appl. 2019;8:85.
31. Barbastathis G, Ozcan A, Situ G. On the use of deep learning for computational imaging. Optica 2019;6:921-41.
32. Jo YJ, Cho H, Lee SY, Choi G, Kim G, Min H, Park YK. Quantitative Phase Imaging and Artificial Intelligence: A: Review. IEEE Journal of selected topics in Quan. Elect. 2019;25:6800914.
33. Di J, Wu J, Wang K, Tang J, Li Y, Zhao J. Quantitative Phase Imaging Using Deep Learning-Based Holographic Microscope. Front. Phys. 2021;22:651313.
34. Rivenson Y, Zhang Y, Günaydın H, Teng D, Ozcan A. Phase recovery and holographic image reconstruction using deep learning in neural networks. Light Sci. Appl. 2018;7:17141.
35. Wu Y, Rivenson Y, Zhang Y, Wei Z, Günaydin H, Lin X, Ozcan A. Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery. Optica 2018;5:704–10.
36. Ren Z, Xu Z, Lam EY. Learning-based nonparametric autofocusing for digital holography. Optica

2018;5:337–44.
37. Sinha A, Lee J, Li S, Barbastathis G. Lensless computational imaging through deep learning. Optica 2017;4:1117-25.
38. Tahon M, Montresor S, and Picart P. Towards Reduced CNNs for De-Noising Phase Images Corrupted with Speckle Noise. Photonics 2021;8:255.
39. Goy A, Arthur K, Li S, Barbastathis G. Low photon count phase retrieval using deep learning. Phys. Rev. Lett. 2018;121:243902.
40. Nguyen T, Bui V, Lam V, Raub CB, Chang LC, Nehmetallah G. Automatic phase aberration compensation for digital holographic microscopy based on deep learning background detection. Opt. Express 2017;25:15043–57.
41. Paine SW, Fienup JR. Machine learning for improved image-based wavefront sensing. Opt. Lett. 2018;43:1235–8.
42. Dardikman G, Shaked NT. Phase unwrapping using residual neural networks. In Imaging and Applied Optics 2018.
43. Spoorthi GE, Gorthi S, Gorthi RKSS. PhaseNet: a deep convolutional neural network for two-dimensional phase unwrapping. IEEE Signal Process. Lett. 2019;26:54–8.
44. Liu T, Haan KD, Rivenson Y, Wei Z, Zeng X, Zhang Y, Ozcan A. Deep learning-based super-resolution in coherent imaging systems. Sci. Rep. 2019;9:3926.
45. Nguyen TC, Bui V, Nehmetallah G. Computational optical tomography using 3-D deep convolutional neural networks. Opt. Eng. 2018;57:043111.
46. Akondi V, Dubra A. Multi-layer Shack-Hartmann wavefront sensing in the point source regime. Biomed. Opt. Express 2021;12:409-32.
47. Dubey N, Kumar R, Rosen J. COACH-based Shack-Hartmann wavefront sensor with an array of phase coded masks. Opt. Express 2021;29:31859-74.
48. Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning. 2018. arXiv:1603.07285.
49. Mahajan NV, Dai G. Orthonormal polynomials in wavefront analysis: analytical solution. J. Opt. Soc. Am. A 2007;24:2994-3016.
50. Lou S, Pagani L, Zeng W, Jiang X, Scott PJ. Watershed segmentation of topographical features on freeform surfaces and its application to additively manufactured surfaces. Precis. Eng. 2020;63:177-86.