



Deep learning wavefront sensing for fine phasing of segmented mirrors

YIRUI WANG,^{1,2} FENGYI JIANG,^{1,2} GUOHAO JU,^{1,3} BOQIAN XU,¹
QICHANG AN,¹ CHUNYUE ZHANG,¹ SHUAIHUI WANG,¹ AND SHUYAN XU^{1,4}

¹CAS Key Laboratory of On-orbit Manufacturing and integration for Space Optics System, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin 130033, China

²University of Chinese Academy of Sciences, Beijing 100049, China

³juguhao@ciomp.ac.cn

⁴xusy@ciomp.ac.cn

Abstract: Segmented primary mirror provides many crucial important advantages for the construction of extra-large space telescopes. The imaging quality of this class of telescope is susceptible to phasing error between primary mirror segments. Deep learning has been widely applied in the field of optical imaging and wavefront sensing, including phasing segmented mirrors. Compared to other image-based phasing techniques, such as phase retrieval and phase diversity, deep learning has the advantage of high efficiency and free of stagnation problem. However, at present deep learning methods are mainly applied to coarse phasing and used to estimate piston error between segments. In this paper, deep Bi-GRU neural work is introduced to fine phasing of segmented mirrors, which not only has a much simpler structure than CNN or LSTM network, but also can effectively solve the gradient vanishing problem in training due to long term dependencies. By incorporating phasing errors (piston and tip-tilt errors), some low-order aberrations as well as other practical considerations, Bi-GRU neural work can effectively be used for fine phasing of segmented mirrors. Simulations and real experiments are used to demonstrate the accuracy and effectiveness of the proposed methods.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Segmented mirror space telescopes represent the development trend of future extra-large-aperture astronomical telescopes [1–2]. Segmented primary mirrors are efficient solutions to the problems with large monolithic primary mirror manufacture and testing, transportation and launch. However, the imaging quality of segmented-mirror telescopes rely heavily on the system alignment state, especially the phasing error between of primary mirror segments. To achieve an acceptable imaging quality, the RMS phasing error between segments should be less than $\lambda/40$. Therefore, segment co-phasing is a crucial technique for the construction of segmented mirror space telescopes.

Image-based co-phasing error sensing method are most popular, mainly including: phase retrieval (PR) [3–5], phase diversity (PD) [6,7] and curvature sensing [8]. Both of them use the diffraction model of the light field, through an iterative algorithm infinitely close to the real exit pupil plane light field distribution to obtain the true phase. These methods have no special hardware requirements, and have a high resolution accuracy. However, the iterative optimization process is usually time-consuming, and it is easy to fall into a local minimum, so that the global optimal solution cannot be obtained.

In recent years, deep learning has been widely applied for many optical imaging problems, including Fourier ptychography [9–11], phase unwrapping [12–14], microscopic image enhancement [15,16], single-pixel imaging [17,18], optical interferometry [19,20], phase retrieval

[21–26], and wavefront sensing [27–36]. Most of them use supervised deep learning to establish a non-linear mapping relationship between phase distribution and intensity image information. It is shown that the low-order continuous aberrations (mainly including defocus, astigmatism and coma, et al.) in monolithic mirror systems can be estimated to an acceptable accuracy using convolutional neural network (CNN) or long short-term memory (LSTM) network. Compared to traditional PR or PD methods, the neural networks have advantages such as a high efficiency (no iterative process) and robustness (no local optimal problem).

Deep learning has also been introduced to phasing of segmented telescopes. Li et al. proposed a CNN-based piston sensing technology (only a rough solution) [37,38]. In their research, five CNNs are trained to determine the translation error range of each sub-mirror in one wavelength. Subsequently, Ma et al. have shown that the piston error can be directly estimated from the intensity image through a sufficiently deep CNN (DCNN) network [39,40]. They discussed that the piston prediction accuracy is affected by the number and size of the aperture. Hui et al. also used deep CNN to diagnose piston error with feature images [41]. Their purpose is to let piston diagnosing approaches get rid of the imaging target of the optical system. In summary, it can be seen that deep learning has a not bad effect on co-phasing.

However, the current deep learning wavefront sensing methods applied to segmented mirrors are mainly for coarse phasing, not applicable for fine phasing. Specifically, they only concern piston error, and other aberrations (tip/tilt and defocus, etc) have not been comprehensively analyzed. Although deep CNNs are widely used, its unique convolution structure is complicated, which consumes a lot of computer storage and computing power. In most of these references, the average RMSE between the real piston value and the predicted piston value of CNNs is approximately 0.04 waves. When the aperture number or diameter increases, the piston accuracy become worse. Therefore, the accuracy of co-phasing only meets the requirement of coarse phasing process, and it does not meet the requirement of fine phasing process. Besides, there are no experiments to prove its feasibility in the actual environment.

In this paper, we establish a Bilateral-Gated Recurrent Unit (Bi-GRU) network to realize the sub-aperture fine phasing problem by learning two phase-diversity intensity images to achieve end-to-end wavefront sensing. We first decompose the pair of point spread function (PSF) images into a sequence of image blocks. The output of Bi-GRU network is still a sequence. Then we connect a fully-connected layer behind the Bi-GRU network. This sequence is used as input and finally we can obtain a set of aberration coefficients. Because there is no complicated operation of gradual convolution in Bi-GRU, its operation is faster than CNN, it takes up less computer resources and is easier to implement. Bi-GRU has long-term memory and can recognize and utilize the inherent relationship between image intensity. Simulations and experiments verify the effectiveness and accuracy of the method, ensuring that it can be successfully applied to the co-phasing recovery of segmented telescopes.

This paper is structured as follows: the second section introduces the source and implementation of Bi-GRU. The third section introduces the image degradation process and how to use Bi-GRU to achieve the co-phasing of segmented telescopes. The fourth section gives the simulation results and compares the co-phasing accuracy of different networks. The fifth section verifies the effectiveness of the method through experiments. Section 6 summarizes the conclusions.

2. Introduction of GRU neural network

GRU network is a variant of recurrent neural network (RNN). The so-called simple RNN has a recurrent hidden state $h^{(t)}$ (n-dimensional), which is a non-linear transformation of input $x^{(t)}$ (m-dimensional), previous hidden state $h^{(t-1)}$ and bias b_h through a activation function, as shown in Eq. (1). Output $y^{(t)}$ is dependent on the current hidden state $h^{(t)}$ and bias b_y . Specifically, in this case, W_{hx} is an $n \times m$ vector, W_{hh} is an $n \times n$ vector, and b_h is an $n \times 1$ vector. Then, the total

parameters in the RNN are equal to n^2+nm+n .

$$\begin{aligned} h^{(t)} &= \psi \left(W_{hx}x^{(t)} + W_{hh}h^{(t-1)} + b_h \right), \\ y^{(t)} &= W_{hy}h^{(t)} + b_y, \end{aligned} \tag{1}$$

As shown in Fig. 1, the special feature of RNN is that the hidden layer is a loop structure, which means that RNN is performing the same operation every time, but input data is different. This structure enables RNN calculations at the current moment to combine historical input information, which helps to process sequence-related tasks. However, basic RNN can only memorize short-term historical input information, and cannot effectively solve the problem of long-term memory.

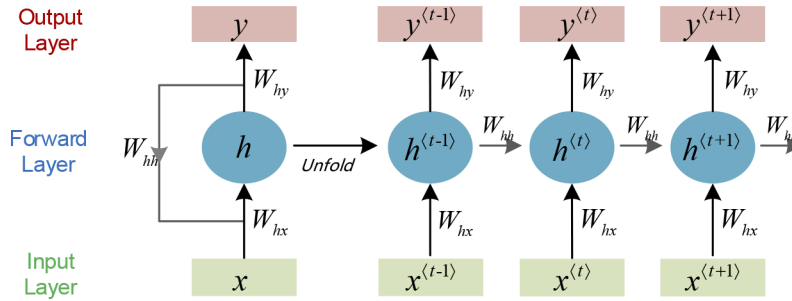


Fig. 1. Structure diagram of the unfold basic RNN model

In contrast to basic RNN, GRU network solves the gradient vanishing problem due to long term dependencies. It mainly refers to when the sequence input is long or the network structure is deep, the correlation between the data information of the previous and subsequent sequences decreases or even disappears, which causes the network to fail to learn the preorder important information on the network layer.

As shown in Fig. 2 (a), the hidden layer structure of the GRU model includes two gates, called an update gate $z^{(t)}$ and a reset gate $r^{(t)}$. The GRU model is presented in the form:

$$\begin{aligned} \hat{h}^{(t)} &= \psi_g \left(W_{hx}x^{(t)} + W_h \left(r^{(t)} * h^{(t-1)} \right) + b_h \right), \\ h^{(t)} &= \left(1 - z^{(t)} \right) * h^{(t-1)} + z^{(t)} * \hat{h}^{(t)}, \end{aligned} \tag{2}$$

and the two gates are presented as:

$$\begin{aligned} r^{(t)} &= \psi_g \left(W_{rx}x^{(t)} + W_{rh}h^{(t-1)} + b_r \right), \\ z^{(t)} &= \psi_g \left(W_{zx}x^{(t)} + W_{zh}h^{(t-1)} + b_z \right), \end{aligned} \tag{3}$$

where σ represents the sigmoid activation function, \tanh is the hyperbolic tangent function, \otimes is an element multiplication, $h^{(t-1)}$ is the hidden state at the previous moment, and $x^{(t)}$ is the current cell's input.

There are three gates and one memory cell in a LSTM block, which is shown in Fig. 2 (b). Input gate $I^{(t)}$, forget gate $F^{(t)}$ and output gate $O^{(t)}$ separately control the data flow from input, memory and output. The activations of input gate, forget gate and output gate depend on current input, previous memory and previous or current output. Thus, the total number of parameters (represented as matrices and bias vectors) for the 3 gates and the memory cell structure are four groups (W_x, W_h, b), which are equal to $4 \times (n^2+nm+n)$.

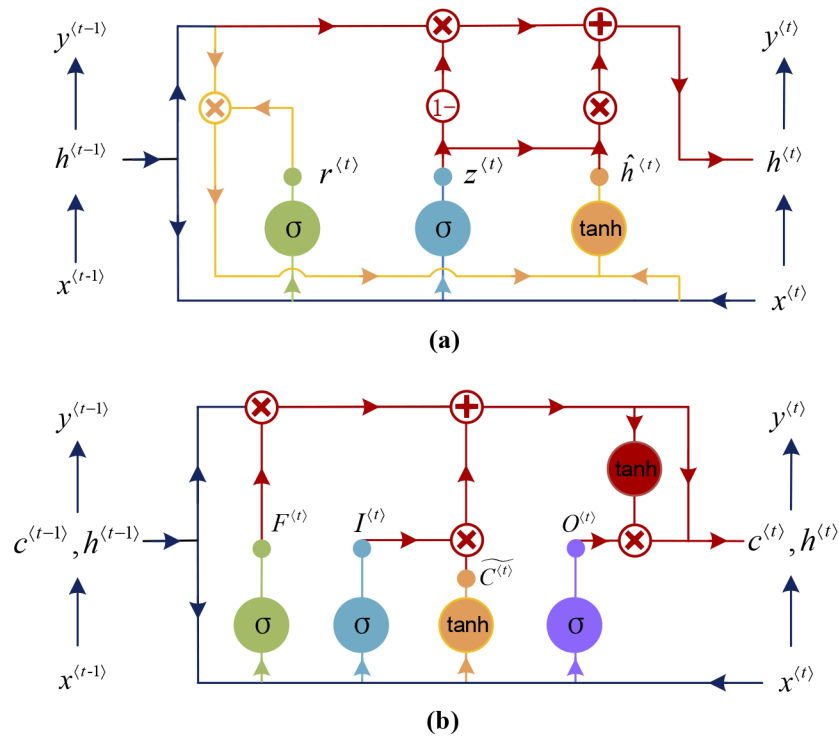


Fig. 2. Structure diagrams of gated recurrent unit (GRU) network (a) and long short-term memory (LSTM) network (b)

However, in GRU, the number of gates is reduced to two ($z^{(t)}, r^{(t)}$), and GRU gets rid of the cell state and uses the hidden state to transmit information. The activations of gates in GRU only depends on current input and previous output. In essence, the GRU has 3-folds increase in parameters in comparison to the simple RNN. Specifically, the total parameters in the GRU equals are equal to $3 \times (n^2 + nm + n)$.

Therefore, GRU has a simpler network structure, which not only reduces the risk of overfitting, but also converges quickly and requires less training time [42–44]. Most importantly, GRU can achieve the same training accuracy as LSTM. Therefore, we decide to construct and train a GRU network. For a particular image pattern, the intensity of pixels is not independent, and they have inherent relations. GRU network can make full use of these relations, when the images are decomposed into a series of small blocks and processed as a sequence.

In order to better correlate the sequence information, this paper applies the Bidirectional GRU (Bi-GRU) network model. The bidirectional GRU is composed of two GRUs in opposite directions, and a hidden layer is added. The structure is showed in Fig. 3. This structure provides the output layer with complete past and future information of each point in the input sequence, which can better identify the characteristics of the image.

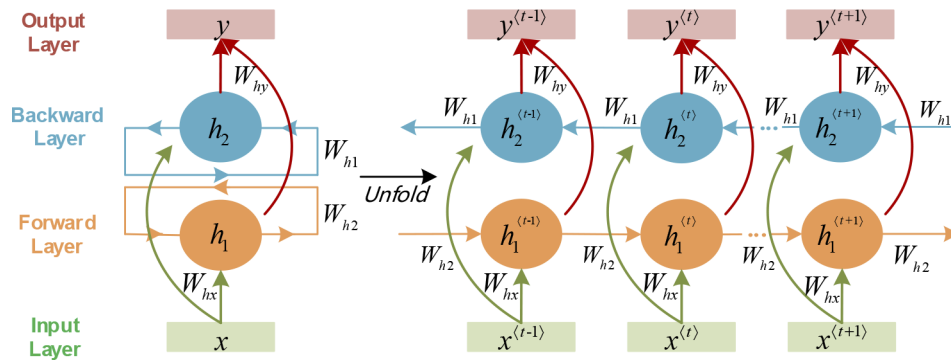


Fig. 3. Structure diagram of Bidirectional GRU model

3. Co-phasing of segmented telescopes using Bi-GRU neural work

3.1. Co-phasing using Bi-GRU neural work

As shown in Fig. 4, we showed the application procedure of how to use the Bi-GRU network to co-phasing of segmented telescopes. Specifically, for certain aberration coefficient ranges, a large number sets of aberration coefficients are randomly generated on each segmented mirror. We can use Fourier optics to calculate the corresponding PSFs at the defocused image plane (see section 3.2 for specific operations). The generated aberration coefficients and the defocused PSF images compose the output data sets and input data sets, respectively. The Bi-GRU network can then be trained using these data sets. After the Bi-GRU network is well trained, we can obtain the optimal weight value, bias value and save the network structure to test the distorted aberration in the real scene.

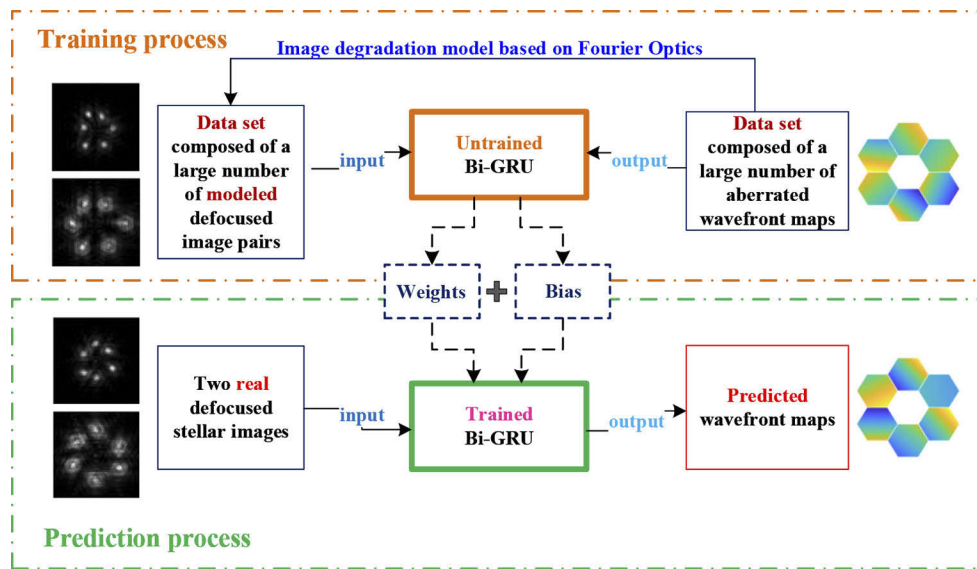


Fig. 4. Sketch map of co-phasing approach using Bi-GRU network

Note that Bi-GRU networks are mainly used to process sequences and the obtained defocused PSF images cannot directly be taken as the input. Therefore, we need decompose two PSF images into a series of patches which can be regarded as a sequence. Specifically, stitch two defocused

images of size $N \times N$ into an image of $N \times 2N$, and then split the stitch image into $2n$ n -dimensional vectors in rows. These vectors can be regarded as $2n$ interrelated timing inputs, and these split vectors are sequentially input into the network. The output of the Bi-GRU networks is still a sequence. This sequence is then taken as a vector, which serves as the input of a fully-connected layer and finally we can obtain a set of aberration coefficients through the Bi-GRU network. We use the aberration coefficients predicted by the Bi-GRU network to recover the wavefront, and the residual root-mean-square errors (RMSEs) between the recovered phase and original phase are used as the evaluation criteria.

3.2. Image degradation model for segmented telescopes

In this paper, the mathematic model of the segmented telescope's primary mirror refers to the JWST telescope, which is composed of 6 smaller-diameter hexagonal sub-mirrors arranged in a certain order in space. The position of each sub-mirror is adjusted by a precise driving device to achieve optical co-phasing. The structure is shown in Fig. 5. This paper numbers each sub-mirror and takes the center of the primary mirror as the origin, establishes the corresponding rectangular coordinate system.

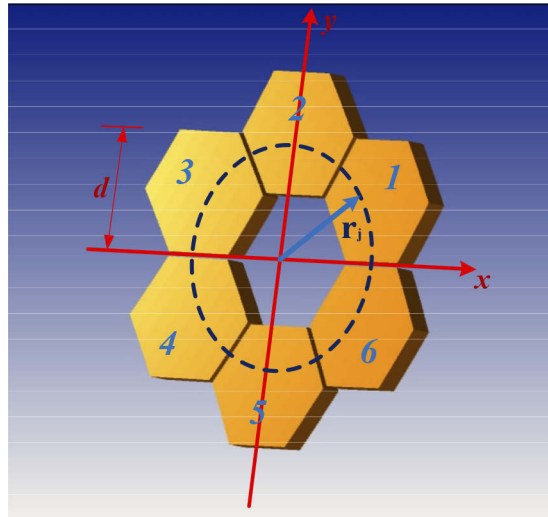


Fig. 5. Structural diagram of 6-segmented primary mirror mathematical model

As shown in Fig. 5, d is the width of each sub-mirror, \vec{r}_j is the center position of the j th sub-mirror. According to the arrangement rule of the segmented sub-mirrors in the Fig. 5, the coordinates of the center position \vec{r}_j is as follows:

$$x_j = \frac{\sqrt{3}d}{2} \cdot l_x, \quad y_j = \frac{d}{2} \cdot l_y. \quad (4)$$

Among them, l_x and l_y are integers, and their values satisfy two conditions: $l_x + l_y$ is an even number; $0 < l_x^2 + l_y^2 < 4$.

The generalized pupil function of the segmented primary mirror can be written as

$$P(x, y) = \sum_{j=1}^N P_j(x - x_j, y - y_j) \exp[i\phi_j(x - x_j, y - y_j)], \quad (5)$$

(x, y) are pupil coordinates of the telescope, (x_j, y_j) are the pupil center coordinates of each sub-mirror. $P_j(x - x_j, y - y_j)$ is the shape function of the submirror and can be given by

$$P_j(x - x_j, y - y_j) = \delta(x - x_j, y - y_j) = \begin{cases} 1, & \text{inside the } j\text{th hexagon} \\ 0, & \text{outside the } j\text{th hexagon} \end{cases}, \quad (6)$$

$\phi_j(x - x_j, y - y_j)$ is the aberration corresponding to the j th submirror and can be expressed as a linear combination of Zernike polynomials. When considering the aberration of piston, tip/tilt in both directions and other high order aberrations, can be written as

$$\phi_j(x - x_j, y - y_j) = \frac{2\pi}{\lambda} (\alpha_{j1}Z_{j1} + \alpha_{j2}Z_{j2} + \alpha_{j3}Z_{j3} + \cdots + \alpha_{jn}Z_{jn}). \quad (7)$$

In the above Eq. (7), Z_{j1} is the piston error of the j th sub-mirror along the optical axis, Z_{j2} and Z_{j3} are the tilt errors of the j th sub-mirror around the X axis and Y axis, and some high order aberrations Z_{jn} are added. α_{jn} is the corresponding aberration coefficient. Therefore, the generalized pupil function of the segmented primary mirror can be written as

$$P(x, y) = \sum_{j=1}^N P_j(x - x_j, y - y_j) \exp \left[i \frac{2\pi}{\lambda} (\alpha_{j1}Z_{j1} + \alpha_{j2}Z_{j2} + \alpha_{j3}Z_{j3} + \cdots + \alpha_{jn}Z_{jn}) \right]. \quad (8)$$

In the optical system, the point spread function corresponding to the intensity distribution of the ideal focal plane image can be obtained by the inverse Fourier transform of the generalized pupil function:

$$PSF(x, y) = |FT^{-}[P(x, y)]|^2. \quad (9)$$

Similarly, the point spread function corresponding to the defocus plane can be expressed as:

$$\begin{aligned} PSF_d(x, y) &= |FT^{-}[P_d(x, y)]|^2, \\ P_d(x, y) &= \sum_{j=1}^N P_j(x - x_j, y - y_j) \exp\{i[\phi_j(x - x_j, y - y_j) + \phi_d(x - x_j, y - y_j)]\}, \\ \phi_d(x, y) &= \alpha_4 Z_4. \end{aligned} \quad (10)$$

In the above Eq. (10), FT^{-} is a two-dimensional inverse Fourier transform operation. $\phi_d(x, y)$ is the introduced known defocus amount, which can be expressed by the fourth term representing the defocus amount in the Zernike polynomial.

4. Simulations

In this section, according to the principle of Fourier optics, MATLAB is used to model the imaging system, generate a training set, and verify the effectiveness of the Bi-GRU network. Then we compare the nonlinear fitting capabilities of Bi-GRU networks and other networks.

4.1. Generation of data set and Bi-GRU implementation

The application procedure of the deep learning fine phasing approach presented above is presented below:

- (1) Determination of the system parameters and generation PSF images

System parameters are the premise for us to generate the data set needed for training the network. In this paper, we set the aperture of the main mirror to 1 m, the target image size to

64×64 pixels, the observation wavelength to 632 nm, the focal length to 10 m, the CCD pixel size to be 5.5μm, and 5 mm defocus distance between two PSFs.

We then use MATLAB to model the optical system. A pair of defocus PSF images can be generated according to the principle of Fourier optics. Common phase error considered here are simply 1st~3th Fringe Zernike coefficients corresponding to piston and tip-tilt, which are randomly generated within the range of $[-0.5\lambda, 0.5\lambda]$. The size of each PSF image is 64×64. After a large number of the repetition of the above process, we generate 10,000 pairs of degraded PSF images and their corresponding aberration coefficients as the input data set and output data set.

(2) Addition of noise and other aberration errors caused by environmental disturbances

In order to approximate the actual imaging environment, 40 dB noise is randomly introduced into the simulated PSF image (the evaluation standard is PSNR, see below). Besides, in order to get closer to the actual wavefront, some high-order aberrations are added within the range of $[-0.2\lambda, 0.2\lambda]$.

To simulate noise, we model each image to have Gaussian CCD read noise with a standard deviation of 15 e- and a dark current of 0.1 e- /s over a 1s integration time. The photon noise which is dependent on intensity follows a Poisson distribution. The peak pixel signal-to-noise ratio (PSNR) is defined as:

$$PSNR = 20\log_{10} \left(\frac{S_{peak}}{\sqrt{S_{peak} + \sigma_{read}^2 + \sigma_{dark}^2}} \right), \quad (11)$$

where S_{peak} is the peak pixel value of the noise-free image, σ_{read}^2 and σ_{dark}^2 are the variances associated with the readout noise and the dark current noise at each pixel, respectively. The peak value of the PSF is set to 10,000 photons, which is limited to the number of fully trapped electrons. Then the final peak pixel PSNR is approximately equal to 40 dB.

(3) Training of the Bi-GRU network

Then, we input the above 1W training sets into the Bi-GRU network, and set the following network parameters: initial learning rate is 0.0003, learning rate attenuation is 0.99, optimization algorithm is Adam algorithm, hidden layer node number is 128, batch size is 200, and the network loss function adopts RMSE, that is, the root mean square error between the output aberration value and the true value, and L2 regularization is added. The CPU used for network training is Intel Core i7-8700, the frequency is 3.20 GHz, and the GPU is NVIDIA Quadro P2000. The software version of Tensorflow is tensorflow-gpu-1.14.1, and the software version of Python is 3.6.

(4) Testing of the effectiveness of the Bi-GRU network

After training, we use MATLAB again to randomly generate 1W test sets according to the above optical system parameters to test the network performance. Under the above same environment and parameters, the network calculates the absolute error between the network output value and the true value for 10,000 sets of test set data, and the average RMSE is 0.0065λ . The histogram is shown in Fig. 6 (a). We can see that the trained Bi-GRU network can accurately establish the non-linear mapping between pairs of degraded PSF images and corresponding phase aberrations.

4.2. Comparison between Bi-GRU network and other networks as fine phasing tool

In this paper, Bi-GRU network is selected as a mathematical tool to predict the wavefront aberration coefficients of the segmented telescope. In fact, other neural networks can also do this

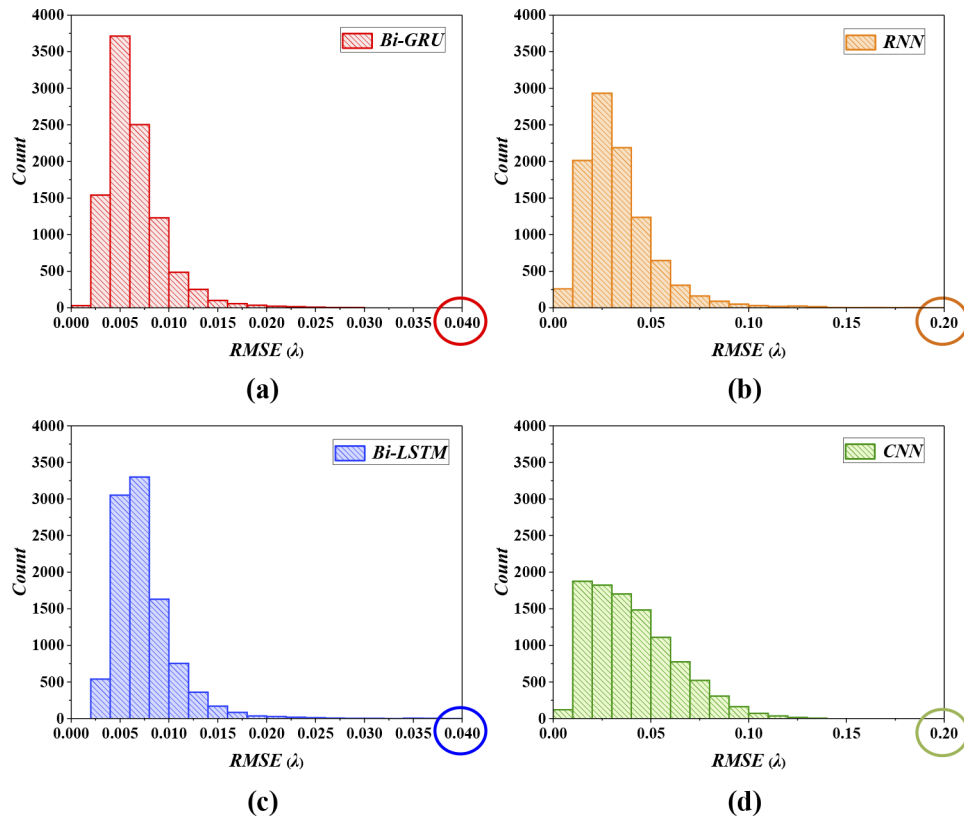


Fig. 6. Comparison of RMSE distribution of each network for test sets of 6-aperture imaging system. We can see that (a) Bi-GRU and (c) deep LSTM can accurately fit the non-linear mapping between the pairs of PSF image and the phase aberrations. The fitting accuracy of (b) standard RNN and (d) deep CNN is much less than Bi-GRU for predicting phase distribution in this paper.

job. The reason why we choose Bi-GRU network as a nonlinear fitting tool is mainly because in response to the problems encountered in this article, the Bi-GRU network not only has higher nonlinear fitting accuracy than other network models, but also has a simpler network structure and the convergence speed is faster.

We compare the fitting accuracy of Bi-GRU, standard RNN, deep Long Short-Term Memory (LSTM, another common variant of RNN) and CNN for the fitting problem of this paper. The number of training sets and test sets for all networks are 10,000 groups. Other conditions for training the network are the same as training Bi-GRU network. Bi-GRU, RNN, and Bi-LSTM all use the single Layer network structure, have the same number of neurons, and the activation functions are all tanh functions. The number of network layers of CNN is 9, including input layer, convolutional layer, pooling layer and fully-connected layer. The activation function uses the relu function (a commonly used activation function for CNN).

Comparing Fig. 6(b), (d) with Fig. 6(a), we can see that the RMSE of using Bi-GRU network is an order of magnitude smaller than that of using standard RNN or CNN network. CNN network has the worst fitting accuracy. The underlying reason may be that the PSF image in this paper contains some sparse and scattered points and it may be a bit difficult to further extract features through convolution. The down-sampling operation of the pooling layer causes the loss of a large amount of valuable original information; and the accuracy of aberration fitting will be

limited. Comparing Fig. 6(c) with Fig. 6(a), we can see that the residual fitting error of using LSTM network is not much different from that of using Bi-GRU network. However, because Bi-GRU has a simpler network structure, its network training convergence speed is much faster than LSTM. Specifically, for the test set in Fig. 6(b)(c)(d), the average absolute error between the network's output and the real value is 0.0733λ , 0.0074λ , and 0.0401λ , respectively.

5. Experiments

In this part, we verify the effectiveness of the proposed method through experiments. The sketch of the experimental setup is shown in Fig. 7. A collimator is used to generate parallel light, which passes through the optical system and a PSF can be obtained with a detector. The optical system is composed of a primary mirror, a secondary mirror and a set of lenses (not shown in Fig. 7) used to expand the field of view of the system. The primary mirror contains three mirror segments. The primary mirror and secondary mirror are placed on PI hexapods which are used to accurately adjust the position and pose of the primary mirror segments and secondary mirror to the accuracy of micron. Piezoelectric ceramics are further used to adjust the position of primary segments to the accuracy of nanometer, and therefore they can be used to correct phasing errors. The optical system is well aligned before this experiment and in each experiment coarse phasing is performed using dispersed fringe sensing before deep learning is applied to fine phasing. Meanwhile, two defocused PSF images serve as the input of Bi-GRU neural network (one PSF is collected before the focus, and the other is collected after the focus), and the defocusing length between two PSF images is 2 mm. In this work, we mainly solve the phasing error between two segments (in other words, in this experiment we only use two mirror segments). Some other parameters of this experimental system are shown in Table 1, which are needed when generating data set for training the Bi-GRU network.

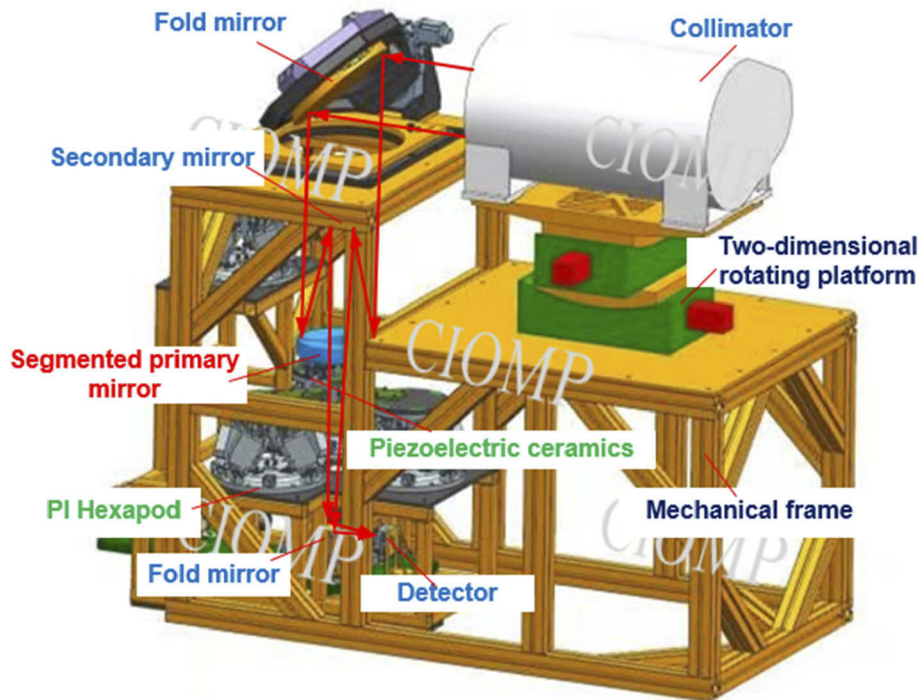


Fig. 7. Sketch map of the experimental setup

Table 1. Some parameters of the experimental system

Parameter	Value
Entrance pupil diameter	Pupil_size = 200mm
Focal length	f = 2000mm
Central wavelength	$\lambda = 550\text{nm}$
Defocusing length	Delta_z = 2mm
CCD pixel size	Pixel_size = 5.5 μm

According to the system parameters shown in Table 1, we simulate and get 10000 images for training and 10000 images for testing. Then a Bi-GRU neural network can be trained for detecting the phasing error between two mirror segments. The training result for the test set is shown in Fig. 8, and it can be recognized that the average RMS error between the predicted value and the actual input value is about 0.00293λ . This result demonstrates that the Bi-GRU network can accurately establish the non-linear mapping between defocused images and phasing error. Note that the influence of airflow and vibration during the experiment cannot be neglected. Therefore, when generating the data set for training the Bi-GRU network, some additional random high-order aberration terms should also be added into the image degradation model (apart from the phasing error).

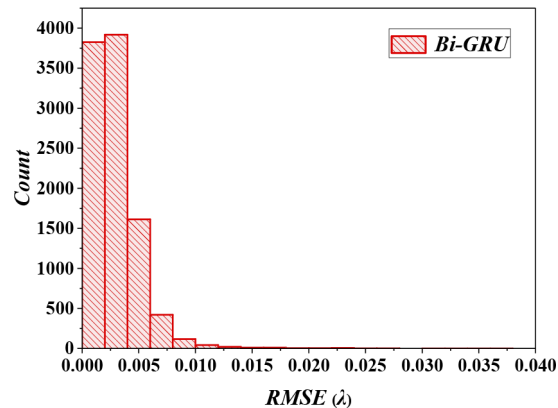


Fig. 8. RMSE distribution for test sets simulated with the experimental system parameters after training

After obtaining a Bi-GRU network through training, it is applied to correct the phasing error between two mirror segments. Three procedures are conducted to demonstrate the effectiveness of the proposed fine phasing approach using Bi-GRU network.

- (1) Reconstruct the input defocused images.

With the two defocused PSF images serving as the input, the Bi-GRU neural network can output the wavefront (containing phasing error) corresponding to each PSF image. We can use the recovered aberration coefficients to re-generate a pair of defocused PSF images according to Fourier optics. As shown in Fig. 9, the first column presents those actually collected defocused PSF images, and the second column provides the reconstructed defocused PSF images using Bi-GRU network. We can intuitively recognize the consistency between the collected PSF images and reconstructed PSF images, which qualitatively demonstrate effectiveness of the Bi-GRU network in phasing error sensing.

- (2) Analyze the focal plane PSF images after correcting phasing error.

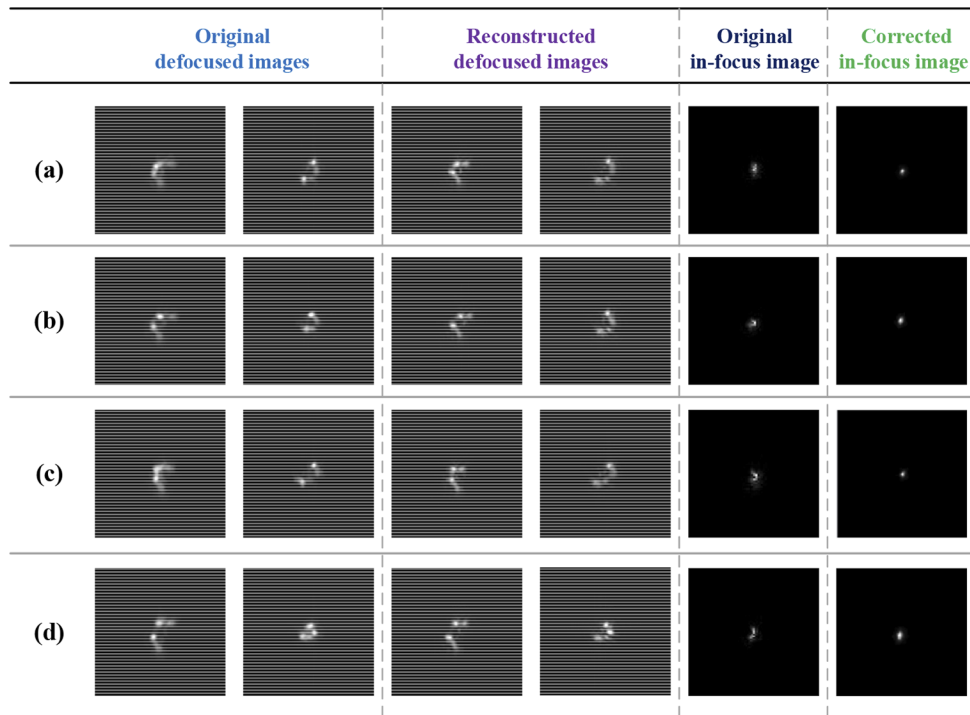


Fig. 9. Four independent sets of experimental data ((a)~(d)) for demonstrating the effectiveness of the proposed phasing error sensing approach using Bi-GRU network. The consistency between the original defocused PSF images (first column) and the reconstructed images (second column) indicates that Bi-GRU network can accurately predict phase aberrations. The in-focus PSF images (last column) after correcting phasing error using actuators become much sharper than original in-focus PSF images (third column), which further demonstrate the effectiveness of the proposed approach in phasing error sensing.

Since we can sense the phasing error between two mirror segments using Bi-GRU network, we can then correct this error using the actuator (piezoelectric ceramics) under each mirror segment. The focal plane PSF images after correcting phasing error with actuator is shown in the last column of Fig. 9. Comparing the last column of Fig. 9 with the third column which presents those focal plane PSFs before correcting phasing error, we can see that the PSFs become much sharper after correcting phasing errors. The Strehl ratio of the original in-focus image is (a) 0.3457, (b) 0.4388, (c) 0.3689 and (d) 0.3017 respectively. After correcting phasing error with actuator, the Strehl ratio of the corrected in-focus image is (a) 0.7116, (b) 0.6482, (c) 0.6989 and (d) 0.6034 respectively, which indicates the wavefront error of the optical system is effectively reduced.

(3) Test the wavefront error of the segmented optical system using interferometer.

The pose of the fold mirror above secondary mirror can be automatically adjusted. When it is horizontally positioned, the wavefront of the optical system can be measured using interferometer, as shown in Fig. 10. The results of the interferometer corresponding to the four sets of experimental data are shown in Fig. 11. It can be seen that the RMS wavefront errors after correcting phasing error are below 0.1 waves, and the residual wavefront errors are mainly higher order components due to manufacture or initial alignment. While single wavelength interferometer cannot break 2π ambiguity, it can still serve as a side validation for the effectiveness of the proposed approach.

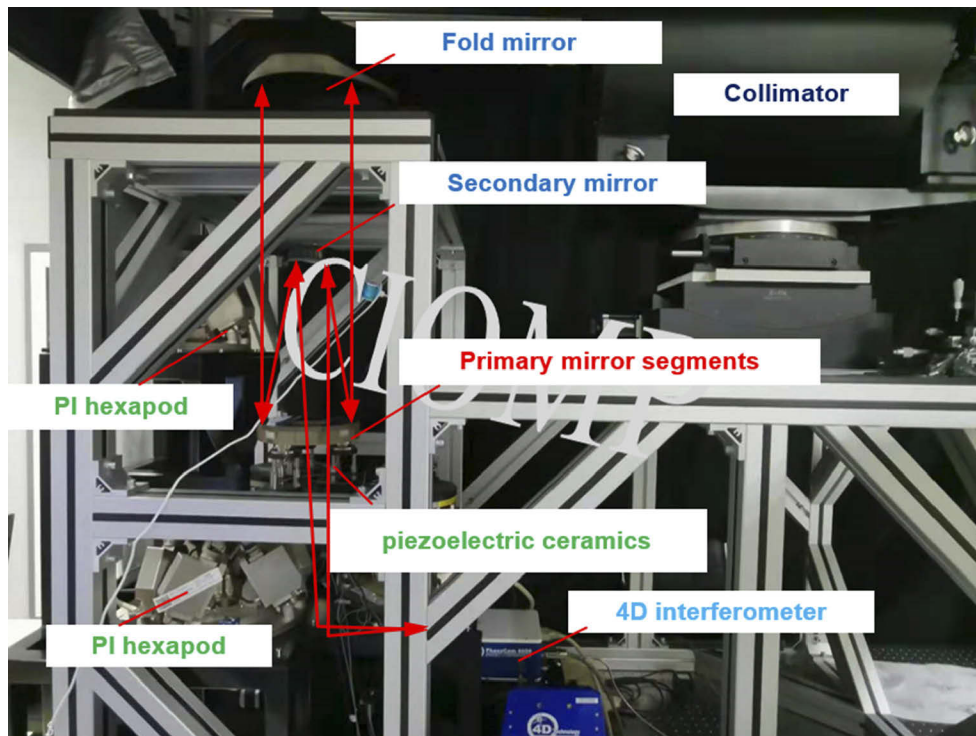


Fig. 10. Autocollimation interference optical path used to test the wavefront error the optical system.

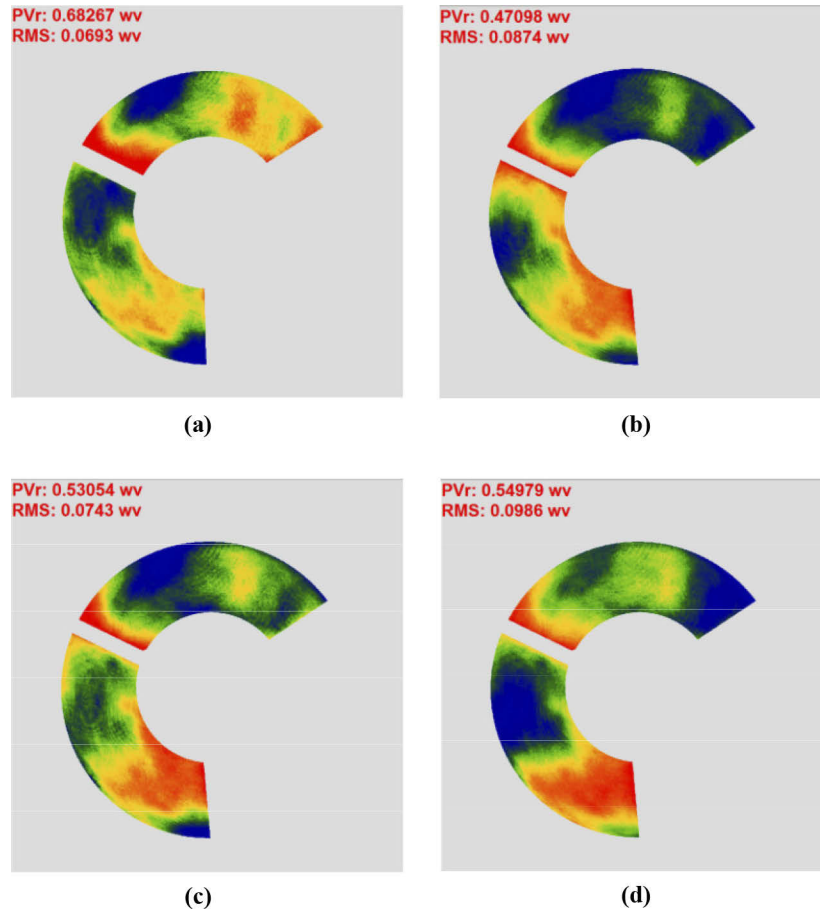


Fig. 11. Wavefront maps for the four cases shown in Fig. 9 after correcting the phasing errors between two mirror segments. We can recognize that the phasing error between two segments are effectively reduced and residual wavefront error are high-order sub-aperture aberrations due to mirror figure manufacture.

6. Other discussions

In this section, we will further discuss phasing error sensing for the case of segmented primary mirror composed of 18 sub-mirrors using the four different kinds of networks, and study the influence of training conditions on the accuracy of the recovered phasing errors. We model an 18-aperture imaging system using Fourier optics, which is shown in Fig. 12. It should be extra explained that the exit pupil diameter of the system is kept unchanged, and the spacing between the sub-apertures is kept constant when the diameter of each-aperture is reduced proportionally.

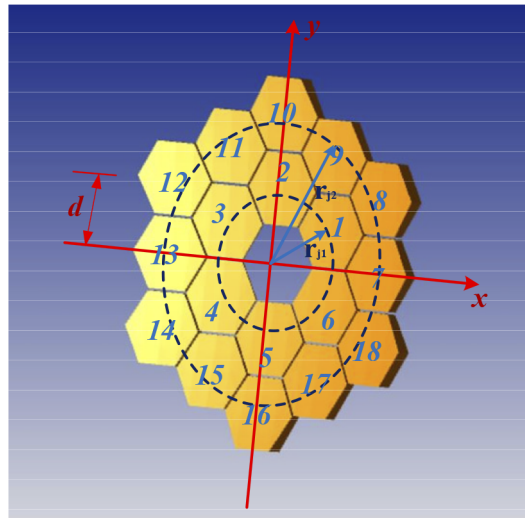


Fig. 12. Structural diagram of 18-segmented primary mirror mathematical model

Specifically, we consider 2 cases: First, we test whether the Bi-GRU model could be used for the fine phasing of the 18-aperture imaging system under the condition that the size of the data set does not change. We collected 10,000 PSF images for training and another 10,000 for testing. We use the same four network structures as above for training, and the phase prediction accuracy is shown on the left side of Fig. 13. On the other hand, we will further qualitatively discuss the impact of increasing the data set and network depth on the accuracy of the restoration of aberration coefficients. Specifically, we use four PSFs image-phase training pairs, and collect 50,000 training data sets for training, and another 10,000 sets for testing.

Considering that the aberration nonlinear mapping relationship established for the 18-aperture imaging system network is much complicated, we further improve the Bi-GRU network by stacking the basic Bi-GRU network repeatedly, and we call it a deep Bi-GRU network. The deep Bi-GRU network can better extract the deep information of the PSF images. After a large number of simulation tests, this paper chooses to use a three-layer stacking method. This paper also made the same improvement to the RNN and LSTM network. For CNN, we increase the number of convolutional layers and neurons in each layer to increase the depth of the network. Bi-GRU, RNN, and Bi-LSTM all use a 3-layer network stack structure, have the same number of neurons, and the activation functions are all tanh functions. CNN also increased the number of network layers accordingly, with the number of network layers being 15. The activation function is the relu function. The co-phasing error accuracy is shown on the right side of Fig. 13, and we compared the performance of the four networks at the same time.

The following conclusions can be drawn from Fig. 13:

- (1) Under the same training conditions as the 6-aperture imaging system, the proposed co-phasing error sensing method based on the Bi-GRU network does not have high prediction

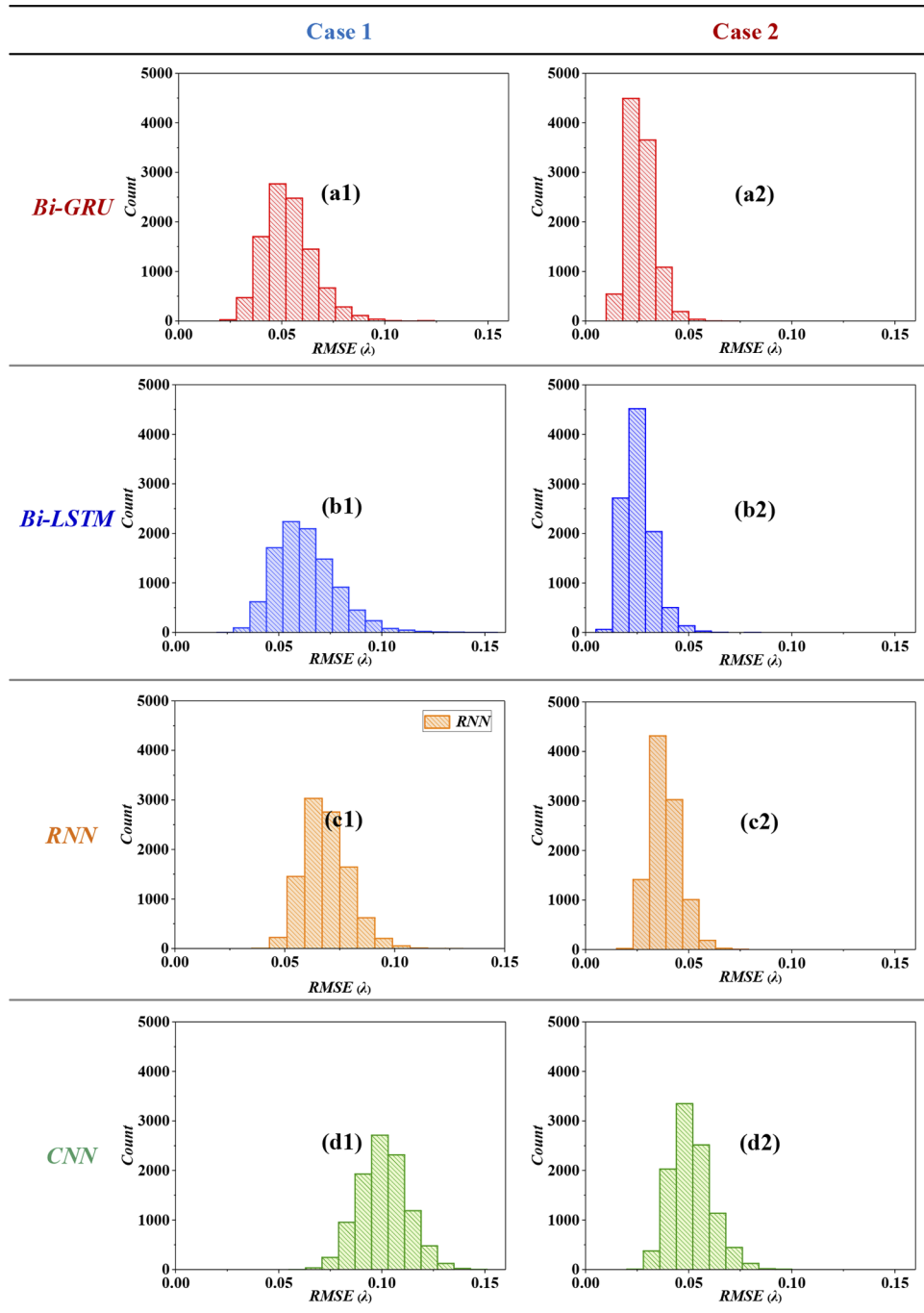


Fig. 13. Comparison of RMSE distribution between different deep learning methods for 18-aperture imaging system under two different training conditions (Case 1 and Case 2). In Case 1, we use two defocused PSF images and collect 10,000 training data sets for training. In Case 2, we increase the data set by using four defocused PSF images and collecting 50,000 training data sets for training. Besides, we also deepen the corresponding each network depth. Result clearly demonstrates that by increasing the data set and network depth, the fine phasing accuracy of the 18-aperture imaging system is effectively improved.

accuracy in the 18-aperture imaging system. The nonlinear mapping relationship between the PSF images and the phase aberration become more complicated when the number of aperture increases, and more aberration coefficients need to be predicted.

- (2) By increasing the data set and network depth, the fine phasing accuracy of the 18-aperture imaging system with the application of the different type of networks is effectively improved. As shown in the column of Case 2, the RMSE between the outputs of the Bi-GRU network and the true aberration value is 0.02953λ . This paper also tested RNN, LSTM and CNN networks, and the RMSE is 0.03842λ , 0.02564λ and 0.05113λ , respectively.
- (3) Bi-GRU network still has advantages over other networks in terms of fitting accuracy and training speed. Compared with CNN network, RNN network can predict the phase error more effectively. Besides, the training time of the CNN network is twice that of the RNN network. As two variants of RNN, GRU and LSTM can effectively solve the problem of gradient disappearance caused by long-term dependence of RNN. Therefore, it can be seen that the phase fitting residuals of the Bi-GRU network and the LSTM network are much better than the RNN network. The predicted phase accuracy of these two networks is not much different. However, due to the simpler structure of the GRU network, GRU consumes less computer resources and converges faster in network training than LSTM under the same conditions.

7. Conclusion

This paper systematically discusses the application of Bi-GRU network to fine phasing of segmented mirrors, including the introduction of Bi-GRU network, the application procedure of using Bi-GRU network for phasing error sensing, and the validation for the effectiveness of the Bi-GRU network (simulations and experiments). Some other discussions are also presented for the case of 18-aperture segmented systems. Some conclusions are presented below:

- (1) Bi-GRU network is superior to other networks in fitting accuracy and training speed. We compare the fitting accuracy of Bi-GRU, standard RNN, deep LSTM and deep CNN for fine phasing problem encountered in this paper under the same conditions. The result show that the former has a higher fitting accuracy and need less training time.
- (2) Real experiments are performed to verify that Bi-GRU network can be used for phasing error sensing and realize end-to-end wavefront sensing and control for segmented telescopes. Even in the presence of some residual higher-order aberrations due to manufacture or initial alignment, the RMS wavefront errors after correcting phasing error are below 0.1 waves.
- (3) The Bi-GRU network can be successfully applied to the co-phasing error sensing of 18-aperture segmented mirrors. By increasing the data set and network depth, we have effectively improved the fine phasing accuracy of the Bi-GRU network on the 18-aperture imaging system.

This work will contribute to the application of deep learning to image-based wavefront sensing and high-resolution imaging.

Funding. National Natural Science Foundation of China (61905241, 62005279, 61705220); Youth Innovation Promotion Association of the Chinese Academy of Sciences (2020221, 2019219).

Disclosures. The authors declare that there are no conflicts of interest related to this article.

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

References

1. P. A. Lightsey, C. Atkinson, M. Clampin, and L. D. Feinberg, "James Webb Space telescope: large deployable cryogenic telescope in space," *Opt. Eng.* **51**(1), 011003 (2012).
2. M. Clampin, "Status of the James Webb Space Telescope Observatory," *Proc. SPIE* **8442**, 84422A (2012).
3. R. Fienup J, "Phase retrieval algorithms: a personal tour," *Appl. Opt.* **52**(1), 45–56 (2013).
4. L. Zhao, J. Bai, Y. Hao, H. Jing, C. Wang, B. Lu, Y. Liang, and K. Wang, "Modal-based nonlinear optimization algorithm for wavefront measurement with under-sampled data," *Opt. Lett.* **45**(19), 5456–5459 (2020).
5. L. Zhao, H. Yan, J. Bai, J. Hou, Y. He, X. Zhou, and K. Wang, "Simultaneous reconstruction of phase and amplitude for wavefront measurements based on nonlinear optimization algorithms," *Opt. Express* **28**(13), 19726–19737 (2020).
6. G. Paxman R, J. Thelen B, and H. Seldin J, "Phase-diversity correction of turbulence-induced space-variant blur," *Opt. Lett.* **19**(16), 1231–1233 (1994).
7. X. Qi, G. Ju, and S. Xu, "Efficient solution to the stagnation problem of the particle swarm optimization algorithm for phase diversity," *Appl. Opt.* **57**(11), 2747–2757 (2018).
8. Q. An, X. Wu, X. Lin, M. Ming, J. Wang, T. Chen, J. Zhang, H. Li, L. Chen, J. Tang, R. Wang, and H. Zhao, "Large segmented sparse aperture collimation by curvature sensing," *Opt. Express* **28**(26), 40176–40187 (2020).
9. T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Deep learning approach for Fourier ptychography microscopy," *Opt. Express* **26**(20), 26470–26484 (2018).
10. S. Jiang, K. Guo, J. Liao, and G. Zheng, "Solving Fourier ptychographic imaging problems via neural network modeling and TensorFlow Biomed," *Opt. Express* **9**(7), 3306–3319 (2018).
11. Y. F. Cheng, M. Strachan, Z. Weiss, M. Deb, D. Carone, and V. Ganapati, "Illumination pattern design with deep learning for single-shot Fourier ptychographic microscopy," *Opt. Express* **27**(2), 644–656 (2019).
12. G Dardikman and T Shaked N, "Phase Unwrapping Using Residual Neural Networks," *Imaging and Applied Optics* (2018).
13. T Sinha A, J Lee, L Shuai, and G Barbastathis, "Solving inverse problems using residual neural networks," *Digital Holography & Three-dimensional Imaging*, (2017).
14. E. Spoorthi G, S. Gorthi, and K. S. S. Gorthi R, "PhaseNet: A Deep Convolutional Neural Network for Two-Dimensional Phase Unwrapping," *IEEE Signal Process. Lett.* **26**(1), 54–58 (2019).
15. Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," *Optica* **4**(11), 1437–1443 (2017).
16. B. Manifold, E. Thomas, A. T. Francis, A. H. Hill, and D. Fu, "Denoising of stimulated Raman scattering microscopy images via deep learning," *Biomed. Opt. Express* **10**(8), 3860–3874 (2019).
17. T. Shimobaba, Y. Endo, T. Nishitsuji, T. Takahashi, Y. Nagahama, S. Hasegawa, M. Sano, R. Hirayama, T. Kakue, A. Shiraki, and T. Ito, "Computational ghost imaging using deep learning," *Opt. Commun.* **413**, 147–151 (2018).
18. C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," *Sci. Rep.* **8**(1), 2369 (2018).
19. S. Feng, Q. Chen, G. Gu, T. Tao, L. Zhang, Y. Hu, W. Yin, and C. Zuo, "Fringe pattern analysis using deep learning," *Adv. Photonics* **1**(02), 1 (2019).
20. K. Wang, Y. Li, Q. Kema, J. Di, and J. Zhao, "One-step robust deep learning phase unwrapping," *Opt. Express* **27**(10), 15100–15115 (2019).
21. A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica* **4**(9), 1117–1125 (2017).
22. A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, "Using deep neural networks for inverse problems in imaging :beyond analytical methods," *IEEE Signal Process. Mag.* **35**(1), 20–36 (2018).
23. Ç. Işıl, F. S. Oktem, and A. Koç, "Deep iterative reconstruction for phase retrieval," *Appl. Opt.* **58**(20), 5422–5431 (2019).
24. G. Zhang, T. Guan, Z. Shen, X. Wang, T. Hu, D. Wang, Y. He, and N. Xie, "Fast phase retrieval in off-axis digital holographic microscopy through deep learning," *Opt. Express* **26**(15), 19388–19405 (2018).
25. Y. Xue, S. Cheng, Y. Li, and L. Tian, "Reliable deep-learning-based phase imaging with uncertainty quantification," *Optica* **6**(5), 618–629 (2019).
26. Y. Rivenson, Y. Zhang, H. Gunaydin, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light: Sci. Appl.* **7**(2), 17141 (2018).
27. H. Cao, J. Zhang, F. Yang, Q. An, and Y. Wang, "Extending Capture Range for Piston Error in Segmented Primary Mirror Telescopes Based on Wavelet Support Vector Machine With Improved Particle Swarm Optimization," *IEEE Access* **99**, 111585 (2020).
28. G. Ju, X. Qi, H. Ma, and C. Yan, "Feature-based phase retrieval wavefront sensing approach using machine learning," *Opt. Express* **26**(24), 31767–31783 (2018).
29. L. Mckl, N. Petrov P, and E. Moerner W, "Accurate phase retrieval of complex point spread functions with deep residual neural networks," *Appl. Phys. Lett.* **115**(25), 251106 (2019).
30. D. Guerra-Ramos, J. Trujillo-Sevilla, and J. M. Rodríguez-Ramos, "Towards Piston Fine Tuning of Segmented Mirrors through Reinforcement Learning," *Appl. Sci.* **10**(9), 3207 (2020).
31. Y. Nishizaki, R. Horisaki, K. Kitaguchi, M. Saito, and J. Tanida, "Analysis of non-iterative phase retrieval based on machine learning," *Opt. Rev.* **27**(1), 136–141 (2020).

32. Q. Tian, C. Lu, B. Liu, L. Zhu, X. Pan, Q. Zhang, L. Yang, F. Tian, and X. Xin, "DNN-based aberration correction in a wavefront sensorless adaptive optics system," *Opt. Express* **27**(8), 10765–10776 (2019).
33. Y. Nishizaki, M. Valdivia, R. Horisaki, K. Kitaguchi, M. Saito, J. Tanida, and E. Vera, "Deep learning wavefront sensing," *Opt. Express* **27**(1), 240–251 (2019).
34. Q. Xin, G. Ju, C. Zhang, and S. Xu, "Object-independent image-based wavefront sensing approach using phase diversity images and deep learning," *Opt. Express* **27**(18), 26102–26119 (2019).
35. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition," IEEE, 2818–2826 (2016).
36. S. W. Paine and J. R. Fienup, "Machine learning for improved image-based wavefront sensing," *Opt. Lett.* **43**(6), 1235–1238 (2018).
37. D. Li, H. Xu, D. Wang, and D. Yan, "Large-scale piston error detection technology for segmented optical mirrors via convolutional neural networks," *Opt. Lett.* **44**(5), 1170–1173 (2019).
38. Guerra-Ramos Dailos, Díaz-García Lara, Trujillo-Sevilla Jose Juan, Manuel, and Rodríguez-Ramos, "Piston alignment of segmented optical mirrors via convolutional neural networks," *Opt. Lett.* **43**(17), 4264–4267 (2018).
39. X. Ma, Z. Xie, H. Ma, Y. Xu, and Y. Liu, "Piston sensing of sparse aperture systems with a single broadband image via deep learning," *Opt. Express* **27**(11), 16058–16070 (2019).
40. X. Ma, Z. Xie, H. Ma, Y. Xu, and G. Ren, "Piston sensing for sparse aperture systems with broadband extended objects via a single convolutional neural network," *Opt. Lasers. Eng.* **128**, 106005 (2020).
41. M. Hui, W. Li, M. Liu, L. Dong, L. Kong, and Y. Zhao, "Object-independent piston diagnosing approach for segmented optical mirrors via deep convolutional neural network," *Appl. Opt.* **59**(3), 771–778 (2020).
42. L. Deng and D. Yu, "Deep learning: methods and applications," *FNT in Signal Processing* **7**(3-4), 197–387 (2014).
43. K Cho, B Van Merriënboer, C Gulcehre, D Bahdanau, F Bougares, H Schwenk, and Y Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation." Computer Science, 2014.
44. N Group, "R-NET: Machine Reading Comprehension with Self-matching Networks." 2017.