MIT Open Access Articles

Deep neural networks to improve the dynamic range of Zernike phase-contrast wavefront sensing in high-contrast imaging systems

Massachusetts Institute of Technology

PROCEEDINGS OF SPIE

Deep neural networks to improve the dynamic range of Zernike phase-contrast wavefront sensing in high-contrast imaging systems

Allan, Gregory, Kang, Iksung, Douglas, Ewan, N'Diaye, Mamadou, Barbastathis, George, et al.

Deep Neural Networks to Improve the Dynamic Range of Zernike Phase-Contrast Wavefront Sensing in High-Contrast Imaging Systems

Gregory Allan[a,d], Iksung Kang[a,d], Ewan S. Douglas[b], Mamadou N'Diaye[c], George Barbastathis[a], and Kerri Cahoy[a]

[a]Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA
[b]Department of Astronomy and Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
[c]Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, UMR 7293, CS 34229, 06304 Nice Cedex 4, France
[d]Equal contribution

ABSTRACT

In high-contrast imaging applications, such as the direct imaging of exoplanets, a coronagraph is used to suppress the light from an on-axis star so that a dimmer, off-axis object can be imaged. To maintain a high-contrast dark region in the image, optical aberrations in the instrument must be minimized. The use of phase-contrast-based Zernike Wavefront Sensors (ZWFS) to measure and correct for aberrations has been studied for large segmented aperture telescopes and ZWFS are planned for the coronagraph instrument on the Roman Space Telescope (RST). ZWFS enable subnanometer wavefront sensing precision, but their response is nonlinear. Lyot-based Low-Order Wavefront Sensors (LLOWFS) are an alternative technique, where light rejected from a coronagraph's Lyot stop is used for linear measurement of small wavefront displacements. Recently, the use of Deep Neural Networks (DNNs) to enable phase retrieval from intensity measurements has been demonstrated in several optical configurations. In a LLOWFS system, the use of DNNs rather than linear regression has been shown to greatly extend the sensor's usable dynamic range. In this work, we investigate the use of two different types of machine learning algorithms to extend the dynamic range of the ZWFS. We present static and dynamic deep learning architectures for single- and multi-wavelength measurements, respectively. Using simulated ZWFS intensity measurements, we validate the network training technique and present phase reconstruction results. We show an increase in the capture range of the ZWFS sensor by a factor of 3.4 with a single wavelength and 4.5 with four wavelengths.

Keywords: Wavefront sensing, High-contrast imaging, Machine learning, Deep neural networks

1. INTRODUCTION

Optical instruments for high-contrast imaging of exoplanets require exceptionally small wavefront error in order operate.[1] In observing systems, error can arise from atmospheric aberrations, deflections of telescope optics, and other mechanical disturbances. These errors are commonly detected and corrected by an adaptive optics (AO) system. The AO system is made up of one or more wavefront sensors (WFS) and an actuator component such as a deformable mirror (DM). The system is used in a closed loop, where phase measurements from the WFS are used to apply corrective displacements to the DM. There are many different wavefront sensor technologies, but the Zernike Wavefront Sensor (ZWFS) has been shown to have optimal sensitivity in the presence of photon noise.[2]

Zernike Wavefront Sensors (ZWFS) use a focal-plane phase-contrast mask with a small region of phase discontinuity that creates a phase-shifted reference wave. The reference wave interferes with the measured

---

Send correspondence to Gregory Allan:
E-mail: gregallan@mit.edu.

wavefront in a re-imaged pupil, converting phase errors into intensity variations that can be measured with a sensor. This technique was pioneered for microscopy of biological samples,[3] but has more recently been refined for use in astronomical telescopes.[4,5] The sensor's nonlinear intensity response to phase displacements in the pupil can be analytically modelled and inverted.[5] This method has shown success as a secondary wavefront sensor for calibration of quasi-static errors on the Spectro-Polarimetric High-contrast Exoplanet REsearch (SPHERE) instrument on the Very Large Telescope (VLT).[6,7] However, the analytical approach breaks down when large unknown wavefront errors are present in the pupil. The range over which the sensor response remains monotonic is -0.14λ to 0.36λ.[6] And with large RMS wavefront error, there is significant cross-talk between wavefront sensor pixels. In other words, the intensity in any one pixel in the re-imaged pupil is not solely dependent on the wavefront displacement in the corresponding area of the pupil plane.

The ZWFS has been proposed for use as a low-order wavefront sensors on the next generation of space telescopes.[8,9] However, in their commissioning phase, space telescopes can be subject to large wavefront errors, which fall outside the capture range of conventional ZWFS methods. One solution is to include additional, high-dynamic-range wavefront sensors as Shack-Hartmann WFS to facilitate coarse alignment.[10,11] For segmented mirrors such as the James Webb Space Telescope (JWST) primary, a complex multi-stage measurement and alignment procedure is required,[12] involving single-purpose optical elements in the form of the Dispersed Hartmann Sensor.[13] This added complexity poses a problem for space systems, which are very limited in volume, mass, and power available on board. Use of the ZWFS for segment co-phasing has been proposed,[14] potentially simplifying the process, but its use is limited to small wavefront errors unless a more complex stepped phase mask is used, along with precise linear actuation.[15] Extending the usable capture range of the standard ZWFS would simplify telescope alignment.

New processing methods for ZWFS data with a larger dynamic range are needed. The method must accurately retrieve a phase aberration map from intensity information recorded with a detector. This is a type of inverse problem, where the process of optical propagation and intensity measurement must be inverted computationally. More specifically, this is a phase retrieval problem, as occurs in the context of many types of physical systems, e.g. holography,[16,17] transport-of-intensity equation (TIE),[18–21] and coherent diffraction imaging (CDI).[22–24] Iterative reconstruction algorithms have often been used to address phase retrieval, such as the Gerchberg-Saxton-Fienup (GSF) algorithm.[25–27] Recently, static neural network techniques have been found to be effective for retrieving phase information from measured intensity[28] even when the photon level of the measurement is extremely low.[29–31] These techniques have also been demonstrated in the context of the Lyot Low-order Wavefront sensor commonly used in high-contrast imaging systems.[32]

In this study, both static and dynamic machine learning algorithms based on Deep Neural Networks (DNNs) are applied to the inverse problem of the ZWFS. Here, a dynamic machine learning algorithm is one which incorporates recurrence of a trainable cell, in contrast to a static neural network which does not use recurrence. The recurrent cell is used to receive multiple inputs along an axis where some dynamics are defined, typically the temporal axis. In this case, we make use of the additional information that can be provided by collecting ZWFS data at multiple wavelengths and implementing recurrence of the network along the wavelength axis. In the following sections, we describe DNN techniques for single- and multi-wavelength ZWFS measurements in the presence of large wavefront aberrations. We evaluate these techniques using simulated ZWFS data. In Sec. 2 we describe the optical model used to generate training and testing data for the neural networks, and in Sec. 3 we describe the DNN methods in detail. Section 4 contains the results of network training and testing, while Sec. 5 focuses on the techniques with the best performance, and compares them to the analytical methods currently in use.[5]

2. OPTICAL MODEL AND SIMULATION

To observe the behavior of the ZWFS in the presence of large wavefront error, and to generate data for the training and testing of neural network methods, a ZWFS system is numerically simulated. The simulation is performed using the Physical Optics Propagation in PYthon (POPPY) library.[33] The optical system shown in Fig. 1 consists of a circular pupil 6 meters in diameter and a focal plane phase-contrast mask. This mask has a circular central area with a diameter of $1.06\lambda/D$ which applies a phase step of $\pi/2$ relative to the rest of the mask. A wavefront with a chosen optical path difference (OPD) error map is provided at the input of the system
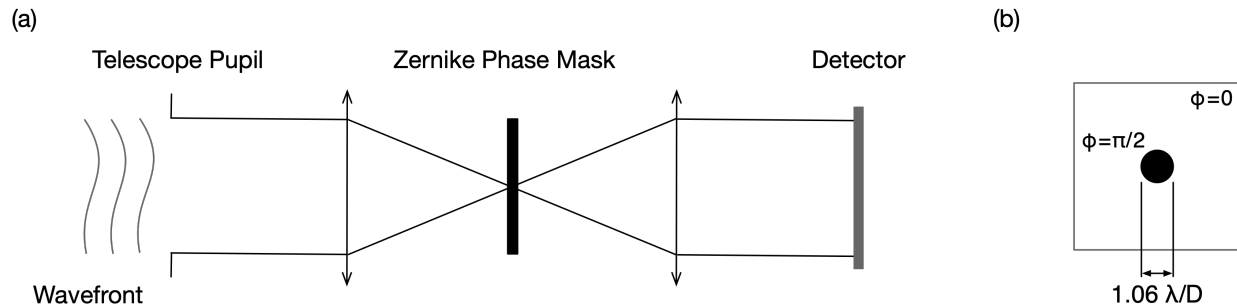
Figure 1. Diagram of the simulated ZWFS system (a). A phase screen is placed at the entrance pupil to simulate a wavefront with known aberrations, and a phase contrast mask is placed at a focal plane. The resulting intensity at a subsequent pupil plane is simulated. The phase mask (b) has a circular area of phase shift $\pi/2$ with diameter $1.06\lambda/D$.

and propagated through the mask to a re-imaged pupil, where the resulting intensity is recorded. We rely on the Matrix Fourier Transform (MFT) method described in[34] to efficiently perform propagation through the small central feature of the focal plane mask with adequate sampling. The phase mask OPD is set to $\pi/2$ for a center wavelength of 1.0 µm, but the output intensity is simulated at four wavelengths of at 0.96 µm, 0.98 µm, 1.0 µm, and 1.2 µm for use in network training.

To provide representative examples of ZWFS response in the presence of wavefront errors, random wavefronts were generated for input into the optical model. The wavefronts were constructed of low-order terms comprising the first 15 Zernike polynomials. The overall wavefront map in polar coordinates $W(\rho,\phi)$ can be given as the sum of the individual Zernike terms $Z_k(\rho,\phi)$ as

$$W(\rho,\phi) = \sum_{k=1}^{K} c_k Z_k(\rho,\phi),\tag{1}$$

where the terms Z_k are normalized to 1 nm RMS wavefront error, and c_k are the weighting coefficients. The resulting RMS wavefront error across the pupil for a chosen set of c_k can be calculated by

$$E^{\mathrm{RMS}} = \sqrt{\overline{\left(W(\rho,\phi)\right)^2}} = \sqrt{\sum_{k=1}^{K} c_k^2},\tag{2}$$

where the over-line notation indicates an average over the polar coordinates ρ and ϕ within the limits of the pupil.

For training and testing of the neural network, example wavefronts were generated by drawing values of c_k such that the resulting values of E_{RMS} are uniformly distributed between 0 and 500 nm. For each of 8000 training examples, 1600 validation examples, and 3000 testing examples, all four wavelengths were simulated.

3. STATIC AND DYNAMIC NEURAL NETWORK TECHNIQUES

The static neural network that we implement for this study is the Phase Extraction Neural Network (PhENN)[28] as shown in Fig. 2a. It is designed to be a U-net structure[35] with a convolutional ResNet encoder and decoder.[36] We find that this particular architecture is helpful for retrieving a phase aberration map on the telescope pupil, shown in Fig. 3, from the intensity measurement recorded on the detector. The network is originally designed to learn a one-to-one correspondence between the input and output; thus we use the network for a single-wavelength application. We also show multi-wavelength reconstructions using this same architecture, for comparison with the dynamic technique described below. Here we do not change the PhENN architecture, but instead augment the input and output datasets by pairing the intensity measurements from multiple wavelengths with a single phase aberration map.
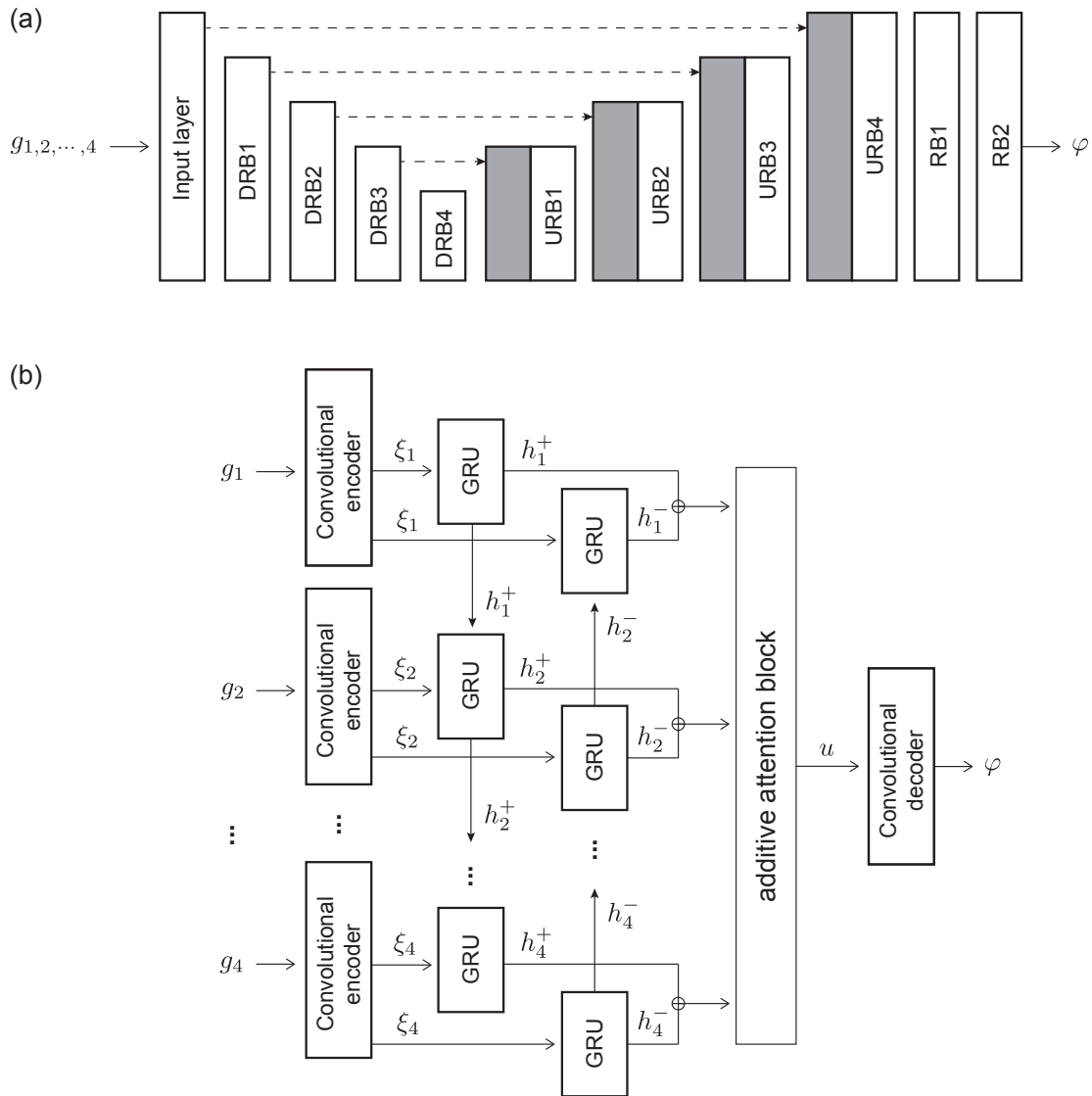
Figure 2. (a) The static neural network is adopted from Sinha et al.,[28] and is designed for a single-wavelength input-output relationship. There are four down-residual blocks (DRB), up-residual blocks (URB), and two residual blocks (RB). Dashed lines denote skip connections as an implementation of the U-net architecture. For training in a multi-wavelength application, the input layer receives N different sensor images ($g_{1,2,\cdots,N}$), and the same corresponding phase map (φ) is replicated N times. (b) A dynamic neural network is proposed for better performance in the multi-wavelength application. It receives multiple sensor images in separate paths, first encoding them individually with convolutional encoders, and further processing them with bi-directional Gated Recurrent Units (GRUs). Hidden features from each GRU are concatenated. The additive attention block merges them into a single representation u, and the convolutional decoder restores it to the original dimensions to generate a phase map φ. The convolutional encoder consists of 4 DRBs, and convolutional decoder is composed of 4 URBs and 2 RBs.

To take better advantage of multi-wavelength measurements, we implement a dynamic machine learning algorithm using the concept of the recurrent neural network, which uses recurrent trainable units along an axis corresponding to real-world dynamics, as described in Sec. 1. The recurrent unit is trained to filter out a static feature, in our case the phase aberration map, across the dynamic axis by retaining useful information to reconstruct as close to the ground truth OPD map and forgetting any redundancies between the recurrent instances. Typical RNN applications in imaging and computer vision include video frame prediction,[37–40] depth map prediction,[41] three-dimensional object reconstruction[42–45] and segmentation[46, 47] from multi-view images. In this implementation, the recurrence is defined along the wavelength axis as shown in Fig. 2b. By varying the wavelength with four different values $(\lambda_1, \cdots, \lambda_4)$, we get four intensity patterns $(g_1, \cdots, g_4)$ each corresponding to a wavelength, which are deeply correlated with each other. The RNN is designed to exploit this correlation among the measurements by using a recurrent unit called Gated Recurrent Unit (GRU).[48] It is a variant of another commonly-used recurrent unit, called Long-Term Short Memory (LSTM),[49] which has more complex computational paths. Despite the comparative simplicity, the GRU does not compromise its performance partly thanks to the reduction of trainable parameters. Additionally, we implement the GRU block in a bi-directional RNN[50] fashion. In this arrangement, the connections between hidden representations are unbiased, matching the unbiased physical relationship between neighboring wavelengths. That is, the information for each wavelength is interpreted in the context of information from *both* of its neighboring wavelengths, taking full advantage of their physical relationship.

The bi-directional recurrent unit forms the core of our dynamic machine learning algorithm, but a convolutional ResNet encoder and decoder are additionally attached. The former is at the GRU input, to form lower-dimensional representations of the ZWFS images for the GRU to operate on, and the latter to restore the output to the original dimensions of the OPD map. Lastly, there is an additive attention block before the convolutional decoder to merge resulting hidden representations from each GRU into a single estimated OPD map u. The block is defined following the convention in ref.[51] as:

$$u = \sum_{n=1}^{N} \alpha_n h_n, \tag{3}$$

a weighted sum over hidden representations h_n, which are each the concatenation of forward and backward hidden features h_n^+ and h_n^-, respectively. The normalized weights α_n are calculated from attention scores e_n, as

$$e_n = \tanh\left(W_e h_n\right),$$
$$\alpha_n = \text{softmax}\left(e_n\right) = \frac{\exp(e_n)}{\sum_{n=1}^{N} \exp(e_n)}, \tag{4}$$
$$n = 1, 2, \ldots, N \, (= 4).$$

The resulting output is effectively the expected hidden feature, in this case an encoding of the correct OPD map, incorporating information from each recurrent unit.

4. RESULTS

In Fig. 4, we show results of the reconstruction for a single example wavefront from the testing dataset. The reconstructed wavefront for each of the reconstruction methods is shown. OPD is given in units of µm for a center wavelength of 1 µm We see that for the all methods tested, the reconstructed wavefront is substantially similar to the true OPD map. For the RNN method, the reconstruction improves as more wavelengths are added, but this is not the case for PhENN. The areas of highest error correspond to those with the largest initial OPD, especially the areas at the edges of the pupil, where edge-effects may also contribute to error.

It is desirable to assess and compare the performance of each method across all 3000 test examples. One useful metric that we can apply is the RMS wavefront error of the phase aberration after subtracting the estimate from the true wavefront OPD map. This corresponds to the residual error that would be present after a single
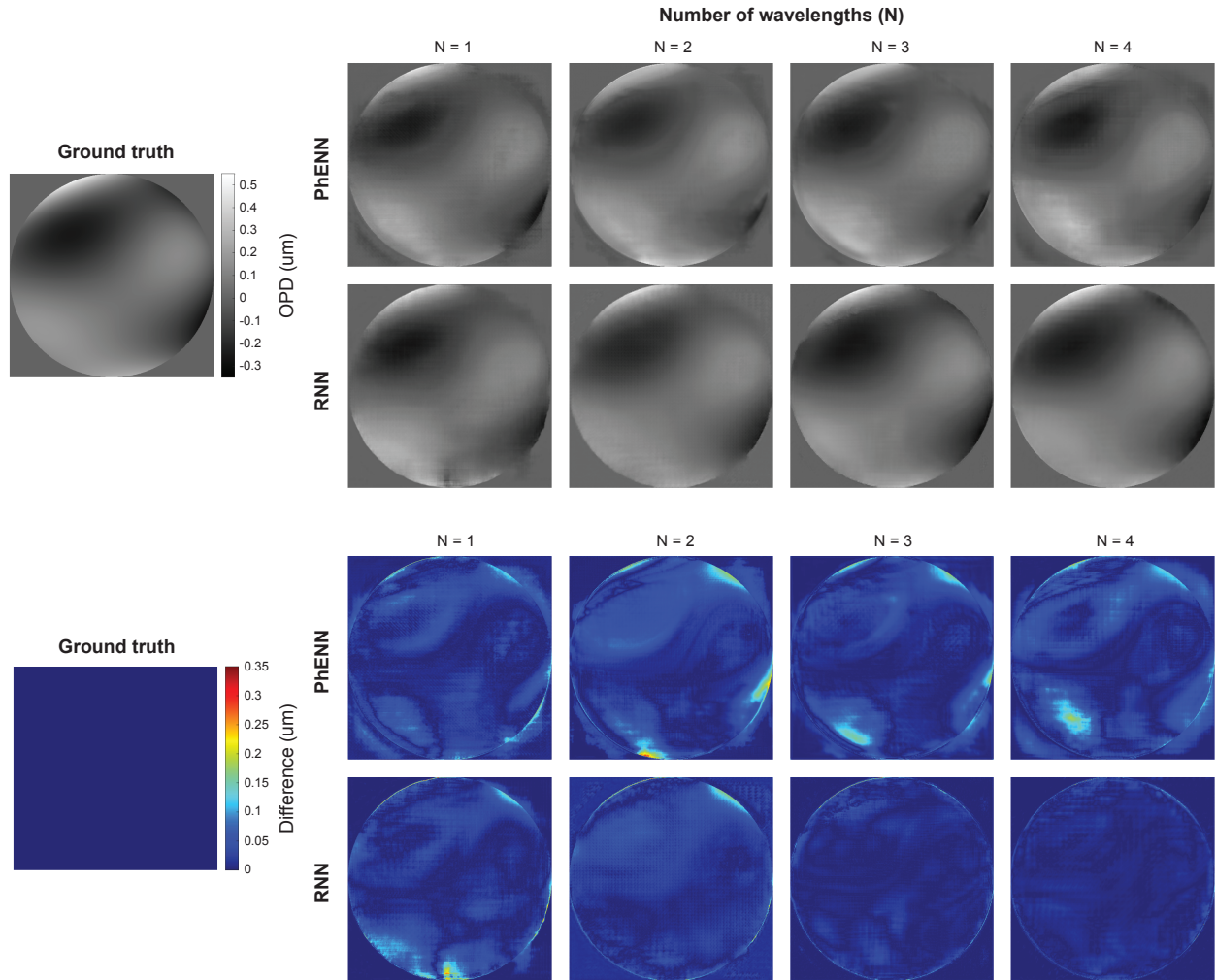
Figure 3. Comparison of reconstructions for a single example wavefront using both the PhENN and RNN methods with different numbers of wavelength samples. The top two rows show the reconstructed wavefronts, while the bottom show the residual after subtracting the true OPD map.

iteration of a well-calibrated AO control system. We can define this residual RMS wavefront error similarly to E_{RMS} in Eq. 2:

$$E_{\mathrm{RMS}}^{\mathrm{res}} = \sqrt{\overline{\left(W(\rho,\phi) - \hat{W}(\rho,\phi)\right)^2}} \, . \tag{5}$$

We also utilize the Pearson correlation coefficient (PCC), which is a measure of linear correlation between the true and estimated wavefront maps. In Fig. 4 we show the values of both of these metrics for both the PhENN and RNN methods, with each number of wavelengths. We see that when using a single ZWFS image for training, PhENN performs better, but significant performance increases are realized when using the RNN to incorporate images taken at multiple wavelengths. No appreciable improvement in wavefront estimation is obtained by using images at multiple wavelengths with PhENN.

5. DISCUSSION

In this section we discuss the performance of two selected methods and compare them to the second-order analytical approximation described by N'Diaye et al.[5] The chosen methods are PhENN with a single wavelength and the RNN with 4 wavelengths. We again use the residual RMS wavefront error $E_{\mathrm{RMS}}^{\mathrm{res}}$, but here we examine
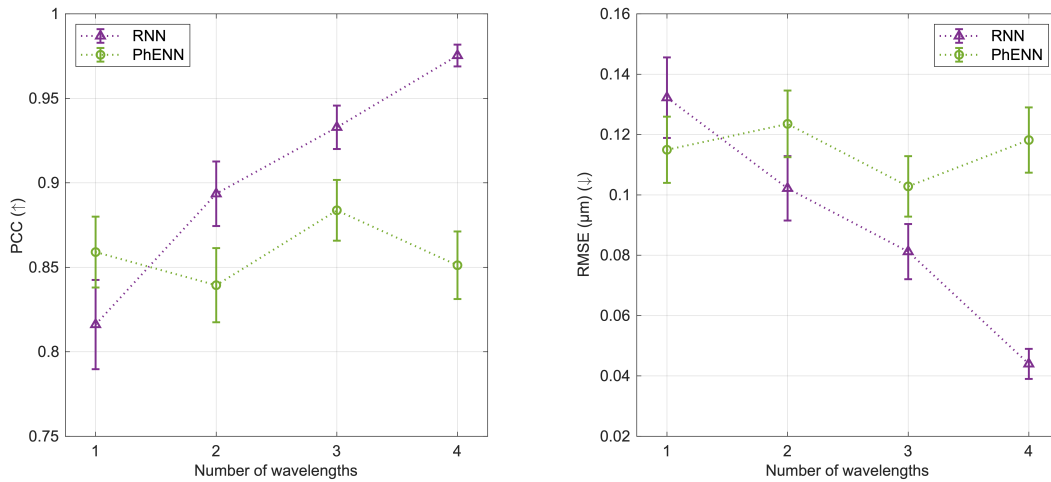
Figure 4. Each data point in this figure shows the mean $\pm$ s.e.m. (95% confidence interval, $P < 0.05$) of 400 test pairs. Figure (left) shows the PCC over a different number of wavelengths considered during the training and testing, and Fig. (right) shows the RMSE in µm.
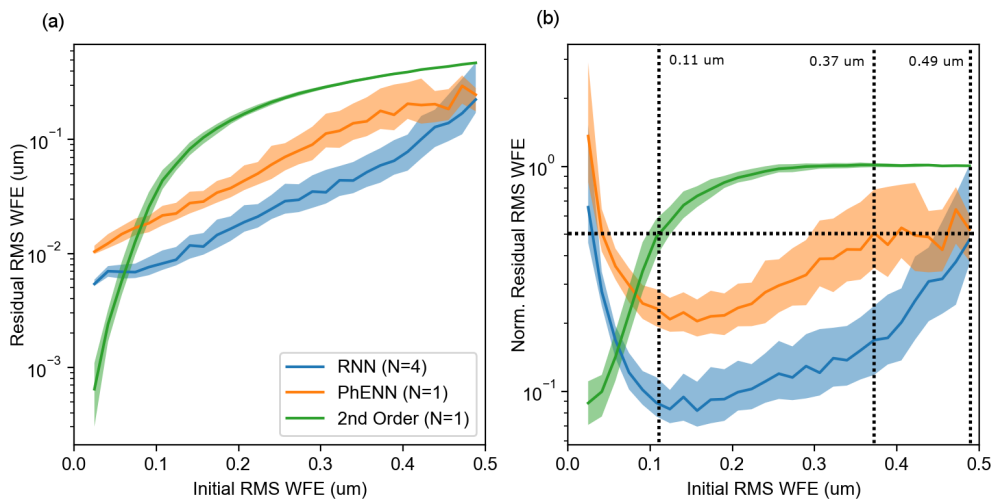

Figure 5. Comparison of overall residual wavefront error after subtracting estimates using three selected methods: the RNN using four wavelength images, the PhENN using a single image, and the existing 2nd-order analytical approach,[5] which uses a single image. In both plots, the solid lines mark mean performance and the shaded areas show the 20-80 percentile performance. (a) shows RMS reconstruction error, while (b) shows error normalized by the initial RMS error of the wavefront. The horizontal dotted line indicates 50% error reduction, and the corresponding vertical lines show the dynamic range for each method. The DNN based approaches are seen to greatly exceed performance of the traditional method for large wavefront errors.

the relationship between this estimation error for each example and the value of the initial RMS wavefront error E_{RMS}, as shown in Fig. 5. In Fig. 5a we see the mean value of $E_{\mathrm{RMS}}^{\mathrm{res}}$ binned by E_{RMS} for each of the three methods. Figure 5b, plots a normalized residual RMS error $E_{\mathrm{RMS}}^{\mathrm{res}}/E_{\mathrm{RMS}}$. We see that the two DNN-based methods exceed the performance of the semi-analytical method, except for cases with small wavefront error where the semi-analytical method has previously been demonstrated to be accurate. For the purposes of this analysis, we define the dynamic range of the sensor as the maximum initial E_{RMS} for which $E_{\mathrm{RMS}}^{\mathrm{res}}$ is less than half that value. This represents a condition for likely convergence in a closed loop. Defined as such, the dynamic

range is 110 nm for the analytical method, 370 nm for PhENN with one wavelength, and 490 nm (roughly the entire range of test data) for the RNN with four wavelengths. PhENN therefore extends the dynamic range of the ZWFS by a factor of $\sim$3.4, and the RNN by a factor of $\sim$4.5. Training the network with a wider range of examples could extend the dynamic range even further.

We see that the performance characteristics of the DNN methods are complementary to the second-order analytical method. The DNN methods trained here on data from a wide range of wavefront conditions have a limit to their absolute error performance, and so perform poorly for weakly aberrated wavefronts in terms of normalized error. In contrast, the analytical approximation method becomes more accurate as the wavefront error becomes smaller. It is therefore likely that a bi-modal system incorporating both of these methods could be used in a closed loop AO system to reduce wavefront error to the noise limits of the analytical method. Future work is necessary to assess the closed-loop stability of such a system.

This study relies on simulated data, and the techniques described must still be validated on a physical ZWFS system. We note the practical considerations in applying deep techniques to a physical system, namely that a large amount of training data is required. This would necessitate the application of known wavefront errors to the system and recording of the resulting ZWFS images. The generation of training data in this way is likely to be time-consuming, but could be sped up by leveraging transfer learning techniques where the network is first trained on simulated data and then fine-tuned on a smaller amount of real-world data.[52]

REFERENCES

[1] Traub, W. A. and Oppenheimer, B. R., [*Direct imaging of exoplanets*], University of Arizona Press, Tucson (2010).

[2] Guyon, O., "Limits of adaptive optics for high-contrast imaging," *The Astrophysical Journal* **629**(1), 592 (2005).

[3] Zernike, F., "Diffraction theory of the knife-edge test and its improved form, the phase-contrast method," *Mon. Notices Royal Astron. Soc.* **94**, 377–384 (1934).

[4] Wallace, J. K., Crawford, S., Loya, F., and Moore, J., "A phase-shifting zernike wavefront sensor for the palomar p3k adaptive optics system," in [*Adaptive Optics Systems III*], **8447**, 84472K, International Society for Optics and Photonics (2012).

[5] N'Diaye, M., Dohlen, K., Fusco, T., and Paul, B., "Calibration of quasi-static aberrations in exoplanet direct-imaging instruments with a zernike phase-mask sensor," *Astronomy & Astrophysics* **555**, A94 (2013).

[6] N'Diaye, M., Vigan, A., Dohlen, K., Sauvage, J.-F., Caillat, A., Costille, A., Girard, J., Beuzit, J.-L., Fusco, T., Blanchard, P., et al., "Calibration of quasi-static aberrations in exoplanet direct-imaging instruments with a zernike phase-mask sensor-ii. concept validation with zelda on vlt/sphere," *Astronomy & Astrophysics* **592**, A79 (2016).

[7] Vigan, A., N'diaye, M., Dohlen, K., Sauvage, J.-F., Milli, J., Zins, G., Petit, C., Wahhaj, Z., Cantalloube, F., Caillat, A., et al., "Calibration of quasi-static aberrations in exoplanet direct-imaging instruments with a zernike phase-mask sensor-iii. on-sky validation in vlt/sphere," *Astronomy & Astrophysics* **629**, A11 (2019).

[8] Shi, F., Balasubramanian, K., Bartos, R., Hein, R., Lam, R., Mandic, M., Moore, D., Moore, J., Patterson, K., Poberezhskiy, I., et al., "Low order wavefront sensing and control for wfirst coronagraph," in [*Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave*], **9904**, 990418, International Society for Optics and Photonics (2016).

[9] Pueyo, L., Zimmerman, N., Bolcar, M., Groff, T., Stark, C., Ruane, G., Jewell, J., Soummer, R., Laurent, K. S., Wang, J., et al., "The luvoir architecture" a" coronagraph instrument," in [*UV/Optical/IR Space Telescopes and Instruments: Innovative Technologies and Concepts VIII*], **10398**, 103980F, International Society for Optics and Photonics (2017).

[10] Mendillo, C. B., Brown, J., Martel, J., Howe, G. A., Hewawasam, K., Finn, S. C., Cook, T. A., Chakrabarti, S., Douglas, E. S., Mawet, D., et al., "The low-order wavefront sensor for the picture-c mission," in [*Techniques and Instrumentation for Detection of Exoplanets VII*], **9605**, 960519, International Society for Optics and Photonics (2015).

[11] Mendillo, C. B., Hewawasam, K., Howe, G. A., Martel, J., Cook, T. A., and Chakrabarti, S., "The picture-c exoplanetary direct imaging balloon mission: first flight preparation," in [*Techniques and Instrumentation for Detection of Exoplanets IX*], **11117**, 1111707, International Society for Optics and Photonics (2019).

[12] Perrin, M. D., Acton, D. S., Lajoie, C.-P., Knight, J. S., Lallo, M. D., Allen, M., Baggett, W., Barker, E., Comeau, T., Coppock, E., et al., "Preparing for jwst wavefront sensing and control operations," in [*Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave*], **9904**, 99040F, International Society for Optics and Photonics (2016).

[13] Shi, F., King, B. M., Sigrist, N., and Basinger, S. A., "Nircam long wavelength channel grisms as the dispersed fringe sensor for jwst segment mirror coarse phasing," in [*Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter*], **7010**, 70102E, International Society for Optics and Photonics (2008).

[14] Janin-Potiron, P., N'Diaye, M., Martinez, P., Vigan, A., Dohlen, K., and Carbillet, M., "Fine cophasing of segmented aperture telescopes with zelda, a zernike wavefront sensor in the diffraction-limited regime," *Astron. Astrophys.* **603**, A23 (2017).

[15] Jackson, K., Wallace, J. K., and Pellegrino, S., "Co-phasing primary mirror segments of an optical space telescope using a long stroke zernike wfs," in [*Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave*], **9904**, 99046D, International Society for Optics and Photonics (2016).

[16] Rivenson, Y., Zhang, Y., Günaydın, H., Teng, D., and Ozcan, A., "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light Sci. Appl.* **7**(2), 17141–17141 (2018).

[17] Wu, Y., Luo, Y., Chaudhari, G., Rivenson, Y., Calis, A., de Haan, K., and Ozcan, A., "Bright-field holography: cross-modality deep learning enables snapshot 3d imaging with bright-field contrast using a single hologram," *Light Sci. Appl.* **8**(1), 1–7 (2019).

[18] Streibl, N., "Phase imaging by the transport equation of intensity," *Opt. Commun.* **49**(1), 6–10 (1984).

[19] Waller, L., Tian, L., and Barbastathis, G., "Transport of intensity phase-amplitude imaging with higher order intensity derivatives," *Opt. Express* **18**(12), 12552–12561 (2010).

[20] Waller, L., Kou, S. S., Sheppard, C. J., and Barbastathis, G., "Phase from chromatic aberrations," *Opt. Express* **18**(22), 22817–22825 (2010).

[21] Waller, L., Tsang, M., Ponda, S., Yang, S. Y., and Barbastathis, G., "Phase and amplitude imaging from noisy images by kalman filtering," *Opt. Express* **19**(3), 2805–2815 (2011).

[22] Nugent, K. A., "Coherent methods in the x-ray sciences," *Advances in Physics* **59**(1), 1–99 (2010).

[23] Horisaki, R., Takagi, R., and Tanida, J., "Learning-based imaging through scattering media," *Opt. Express* **24**, 13738–13743 (2016).

[24] Miao, J., Sandberg, R. L., and Song, C., "Coherent x-ray diffraction imaging," *IEEE J. Sel. Top. Quantum Electron.* **18**(1), 399–410 (2011).

[25] Gerchberg, R. W., "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik* **35**, 237–246 (1972).

[26] Saxton, W., [*Computer techniques for image processing in electron microscopy*], vol. 10, Academic Press (2013).

[27] Fienup, J. R., "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**(15), 2758–2769 (1982).

[28] Sinha, A., Lee, J., Li, S., and Barbastathis, G., "Lensless computational imaging through deep learning," *Optica* **4**(9), 1117–1125 (2017).

[29] Goy, A., Arthur, K., Li, S., and Barbastathis, G., "Low photon count phase retrieval using deep learning," *Phys. Rev. Lett.* **121**(24), 243902 (2018).

[30] Deng, M., Li, S., Goy, A., Kang, I., and Barbastathis, G., "Learning to synthesize: Robust phase retrieval at low photon counts," *Light Sci. Appl.* **9**(1), 1–16 (2020).

[31] Kang, I., Zhang, F., and Barbastathis, G., "Phase extraction neural network (PhENN) with coherent modulation imaging (CMI) for phase retrieval at low photon counts," *Opt. Express* **28**(15), 21578–21600 (2020).

[32] Allan, G., Kang, I., Douglas, E. S., Barbastathis, G., and Cahoy, K., "Deep residual learning for low-order wavefront sensing in high-contrast imaging systems," *Opt. Express* **28**(18), 26267–26283 (2020).

[33] Perrin, M., Long, J., Douglas, E., Sivaramakrishnan, A., and Slocum, C., "Poppy: Physical optics propagation in python," *Astrophysics Source Code Library (ASCL)* (2016).

[34] Soummer, R., Pueyo, L., Sivaramakrishnan, A., and Vanderbei, R. J., "Fast computation of lyot-style coronagraph propagation," *Optics Express* **15**(24), 15935–15951 (2007).

[35] Ronneberger, O., Fischer, P., and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," in [*Medical Image Computing and Computer-Assisted Intervention (MICCAI)*], 234–241, Springer (2015).

[36] He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition," in [*Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*], 770–778 (2016).

[37] Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-c., "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in [*Adv. Neural Inf. Process. Syst. (NIPS)*], 802–810 (2015).

[38] Wang, Y., Jiang, L., Yang, M.-H., Li, L.-J., Long, M., and Fei-Fei, L., "Eidetic 3D LSTM: A model for video prediction and beyond," in [*International Conference on Learning Representations (ICLR)*], (2018).

[39] Wang, Y., Long, M., Wang, J., Gao, Z., and Philip, S. Y., "PredRNN: Recurrent neural networks for predictive learning using spatiotemporal LSTMs," in [*Adv. Neural Inf. Process. Syst. (NIPS)*], 879–888 (2017).

[40] Wang, Y., Gao, Z., Long, M., Wang, J., and Yu, P. S., "PredRNN++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning," *arXiv preprint arXiv:1804.06300* (2018).

[41] CS Kumar, A., Bhandarkar, S. M., and Prasad, M., "Depthnet: A recurrent neural network architecture for monocular depth prediction," in [*Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*], 283–291 (2018).

[42] Liu, J. and Ji, S., "A novel recurrent encoder-decoder structure for large-scale multi-view stereo reconstruction from an open aerial dataset," in [*Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*], 6050–6059 (2020).

[43] Choy, C. B., Xu, D., Gwak, J., Chen, K., and Savarese, S., "3D-R2N2: A unified approach for single and multi-view 3D object reconstruction," in [*European Conference on Computer Vision (ECCV)*], 628–644, Springer (2016).

[44] Kang, I., Goy, A., and Barbastathis, G., "Limited-angle tomographic reconstruction of dense layered objects by dynamical machine learning," *arXiv preprint arXiv:2007.10734* (2020).

[45] Cierniak, R., "A new approach to image reconstruction from projections using a recurrent neural network," *International Journal of Applied Mathematics and Computer Science* **18**(2), 147–157 (2008).

[46] Le, T., Bui, G., and Duan, Y., "A multi-view recurrent neural network for 3D mesh segmentation," *Comput. Graph.* **66**, 103–112 (2017).

[47] Stollenga, M. F., Byeon, W., Liwicki, M., and Schmidhuber, J., "Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation," in [*Adv. Neural Inf. Process. Syst. (NIPS)*], 2998–3006 (2015).

[48] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078* (2014).

[49] Hochreiter, S. and Schmidhuber, J., "Long short-term memory," *Neural Comput.* **9**(8), 1735–1780 (1997).

[50] Schuster, M. and Paliwal, K. K., "Bidirectional recurrent neural networks," *IEEE Trans. on Signal Process.* **45**(11), 2673–2681 (1997).

[51] Bahdanau, D., Cho, K., and Bengio, Y., "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473* (2014).

[52] Wang, W., Ni, H., Su, L., Hu, T., Ren, Q., Gerstoft, P., and Ma, L., "Deep transfer learning for source ranging: Deep-sea experiment results," *J. Acoust. Soc. Am.* **146**(4), EL317–EL322 (2019).