

# Image-to-image translation for wavefront and point spread function estimation

Jeffrey Smith<sup>Ⓢ, a,\*</sup>, Jesse Cranney<sup>Ⓢ, b</sup>, Charles Gretton<sup>Ⓢ, a</sup>  
and Damien Gratadour<sup>Ⓢ, b,c</sup>

<sup>a</sup>Australian National University, School of Computing, Canberra,  
Australian Capital Territory, Australia

<sup>b</sup>Australian National University, Research School of Astronomy and Astrophysics,  
Advanced Instrumentation Technology Centre, Canberra, Australian Capital Territory,  
Australia

<sup>c</sup>University PSL, CNRS, Sorbonne Université, Université de Paris, LESIA,  
Observatoire de Paris, Meudon, France

**Abstract.** We develop and evaluate a new approach to phase estimation for observational astronomy that can be used for accurate point spread function reconstruction. Phase estimation is required where a terrestrial observatory uses an adaptive optics (AO) system to assist astronomers in acquiring sharp, high-contrast images of faint and distant objects. Our approach is to train a conditional adversarial artificial neural network architecture to predict phase using the wavefront sensor data from a closed-loop AO system. We present a detailed simulation study under different turbulent conditions, using the retrieved residual phase to obtain the point spread function of the simulated instrument. Compared to the state-of-the-art model-based approach in astronomy, our approach is not explicitly limited by modeling assumptions, e.g., independence between terms, such as bandwidth and anisoplanatism—and is conceptually simple and flexible. We use the open-source COMPASS tool for end-to-end simulations. On key quality metrics, specifically the Strehl ratio and Halo distribution in our application domain, our approach achieves results better than the model-based baseline. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: [10.1117/1.JATIS.9.1.019001](https://doi.org/10.1117/1.JATIS.9.1.019001)]

**Keywords:** generative adversarial networks; convolutional neural network; wavefront estimation; point spread function reconstruction; adaptive optics.

Paper 22082G received Aug. 19, 2022; accepted for publication Jan. 3, 2023; published online Jan. 19, 2023.

## 1 Introduction

Adaptive optics (AO) systems are an important component of astronomical imaging for large ground-based telescopes, enabling the capture of high-contrast images of faint objects. Aberrations due to Earth's atmospheric turbulence are a significant impediment to astronomical imaging, so the ability to estimate and compensate is critical. The scale of data requirements for this estimation problem increases quadratically with the telescope diameter, an ongoing problem while astronomers build larger telescopes to capture the light from fainter objects such as exoplanets or distant galaxies.

An AO system contains three main components: (i) a deformable mirror with a reflective surface that can be adjusted with an array of actuators to counteract some of the wavefront phase aberrations, (ii) a wavefront sensor (WFS) that collects information about the wavefront phase, and (iii) a controller that interprets the wavefront sensor information and computes a control solution to drive the actuators of the deformable mirror. These three components work in closed-loop in real-time, typically in the order of several kHz. In practice, the correction is not perfect, so the residual point spread function (PSF) after AO correction differs from the diffraction-limited PSF.

---

\*Address all correspondence to Jeffrey Smith, [jeffrey.smith@anu.edu.au](mailto:jeffrey.smith@anu.edu.au)

To deliver the most science from an astronomical observation, it is crucial that astronomers have access to an accurate estimate of the effective PSF during that observation window. State-of-the-art workflows<sup>1</sup> employ model-based techniques to estimate the residual phase and the corresponding PSF, enabling the quality of captured science images to be improved, e.g., by deconvolution.

The WFS is used to capture the instantaneous state of the phase into intensity variations in an image. In the case of the Shack–Hartmann (SH) concept, the telescope aperture is split into sub-regions, called sub-apertures, and an image of the reference guide source is created for each of these sub-apertures and captured by a camera. A centroider algorithm is then used to estimate the spot displacements, in each of these sub-apertures, with respect to a reference position. These displacements are directly related to the local slopes of the wavefront.<sup>2</sup> This information is used for mirror control, and is a significant part of the telemetry data used by existing PSF reconstruction techniques, the latter being our application here. A drawback of existing model-based wavefront phase estimation is that all non-linear, high-order wavefront information captured on the WFS is neglected. It is highly desirable for the non-linear information to be available to maximize the value of captured science assets.

## 1.1 Related Work

AO systems have been used to compensate for atmospheric turbulence since the late 1980s, when the available computer technology was first able to match the requirements for controlling the available deformable mirror technology. Since then, efforts to improve PSF and wavefront estimates have been ongoing both in model-based statistical estimation and other techniques using advances in artificial intelligence.

Model-based methods for wavefront and PSF estimation produce highly accurate estimates using fitting, reconstruction, and simulation methods.<sup>1</sup> The current state-of-the-art uses a comprehensive breakdown of error sources<sup>3</sup> and allows for frame-by-frame validation of end-to-end models using AO loop simulation tools such as COMPASS.<sup>4</sup> As these models are statistical in nature, they require thousands of iterations for wavefront estimation, which becomes demanding on computational equipment as the scale of the telescope increases.

AI-based techniques for improving AO systems have been investigated with many studies using slope estimates from centroider data<sup>5</sup> and typically making wavefront estimates by predicting the weights of a small number of low order, linearly independent Zernike modes<sup>6,7</sup> that can be added to create the wavefront phase image. Using the slope estimates from centroider algorithms limits WFS data to low-order information, as these wavefront slope estimates filter out higher-order information captured on the SH-WFS. (Wavefront) sensor-less methods have used convolutional neural networks (CNNs)<sup>8</sup> to estimate the wavefront from the PSF in the AO loop, rather than from a WFS. That approach can avoid the loss of some higher-order information but are best suited to low turbulence conditions and small telescope settings, and is not used for PSF reconstruction. Work on PSF reconstruction with deep neural networks has been presented by Guyon et al.,<sup>9</sup> however, the method is not described in detail, is applied to Pyramid WFS, and that work does not describe thorough testing in a range of operating conditions.

We describe a novel technique for PSF reconstruction based on wavefront phase estimation, by adapting an extremely general translational image(input)-to-image(output) artificial neural network (ANN). Our contributions are based on machine learning technology for general computer vision tasks, and thereby, leverage many decades of conceptual advances from investigations into supervised connectionist machine learning for a wide variety of specialized computer vision tasks. Specifically, generative networks of the type we investigate have their history in: (i) biologically inspired convolutional networks for interpreting visual scenes,<sup>10</sup> whose genesis and first successful applications lie in character recognition applications,<sup>11</sup> (ii) the development of general and robust activation functions that are compatible with automated differentiation, developed for modern ANNs,<sup>12,13</sup> (iii) deep feed-forward networks pursued by an active research program since their extraordinary representational power and applicability was understood,<sup>14,15</sup> and more directly relevant to our contribution, their practical development to a very broad range of computer vision tasks, such as biomedical image segmentation tasks motivating the convolutional UNet,<sup>16</sup> and (iv) fast parallelizable optimization procedures for learning the parameters of

multi-layered/deep architectures.<sup>17</sup> Furthermore, our contribution is based on the adversarial learning setting. Two ANNs are trained simultaneously, in our case with a feed-forward UNet trained to produce the desired phase estimates, and a Markov discriminator (i.e., a classifier) network that is trained to distinguish between real phase screens and those generated by the UNet. Usefully for PSF reconstruction, adversarial learning enables us to train a network that precisely represents fine/sharp detail, whereby generated artifacts exhibit the statistics of the training corpus.<sup>18</sup>

## 1.2 Contribution

In this paper, we develop a new method for phase estimation in AO that exhibits similar or better accuracy than the state-of-the-art model-based approach while being conceptually simpler and avoiding strong assumptions on the nature and properties of the stochastic process or system geometry, which solves an important engineering problem for practical implementations. We use a translational CNN to infer the phase directly from the SH-WFS image. Our method takes advantage of high-frequency information available in the SH-WFS that has not been used by existing estimation methods that instead use data-intensive statistical estimation methods from AO loop telemetry data (e.g., mirror control voltages). Our approach has immediate application in science image deconvolution workflows, and future application in the control of AO systems.

The paper is structured as follows. First, we discuss the AO setting and current methods. Second, we describe COMPASS—a state-of-the-art GPU-accelerated AO simulation package. Third, we describe our approach to using image-to-image CNNs for phase recovery in our setting. Fourth, we describe current methods in wavefront phase estimation and PSF analysis. Finally, we present a detailed analysis of experimental results comparing our approach to residual phase estimation with the state-of-the-art model-based method.

## 2 Background and Methodology

### 2.1 Adaptive Optics

The goal of AO is to obtain a sharp image of an observed target. Any aberrations of the incoming wavefront create perturbations in the image, which reduces the contrast of the observation—this translates as blur. A perfectly unaberrated image of a point source obtained through a telescope will produce a PSF that is diffraction limited, i.e., the image quality and resolution are only limited by the diffraction of the telescope aperture. Adding other sources of aberration, such as atmospheric turbulence, perturbs the wavefront by introducing optical path differences between the different points of the telescope aperture, affecting the PSF, and therefore image quality.

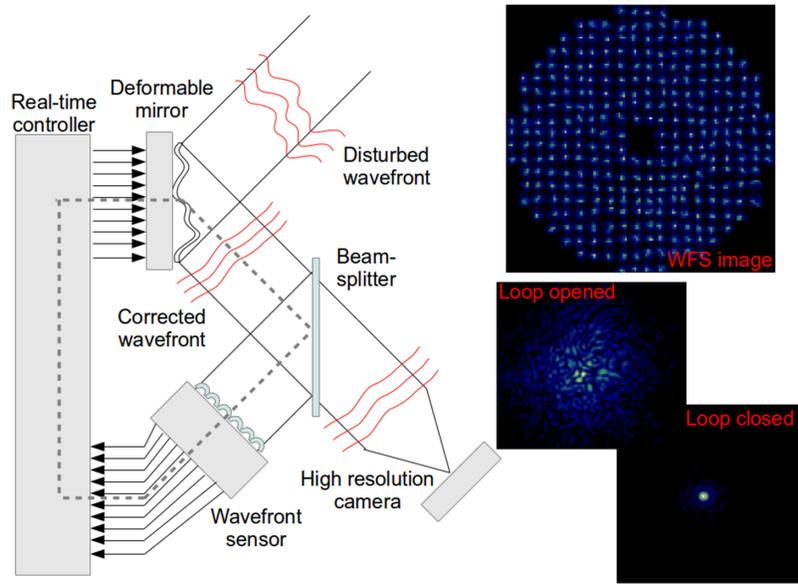
The PSF can be calculated from the wavefront phase as the absolute square of the Fourier transform of the complex electromagnetic field. This is an important relationship and is not invertible from PSF to phase, shown in Eq. (1)

$$\text{PSF} = |\text{FFT}(\text{amplitude} \cdot e^{i \cdot \text{phase}})|^2. \quad (1)$$

The AO loop is shown in Fig. 1 with each of the main components labeled. As the corrections made with the deformable mirror are imperfect, there will be some residual error passed on in the system and into telescope observations.

The residual error (AO loop error) is made up of several contributing sources and are dealt with in detail through an error budget estimation.<sup>3</sup> The error budget describes the total AO loop error in terms of components for bandwidth error, anisoplanatism, aliasing, noise, wavefront measurement error, mode filtering, and fitting error. All of these error sources contribute to the decrease in image quality at the output of the telescope, and are incorporated into the state-of-the-art PSF reconstruction technique described by Ferreira et al. in Ref. 3.

This numerical method of PSF reconstruction provides excellent results in simulation, though it relies heavily on knowledge of the system parameters available to the simulation. It also requires large buffers of data collected from the AO loop to estimate the wavefront error due to the statistical methods used and has a high computational cost. This creates a highly

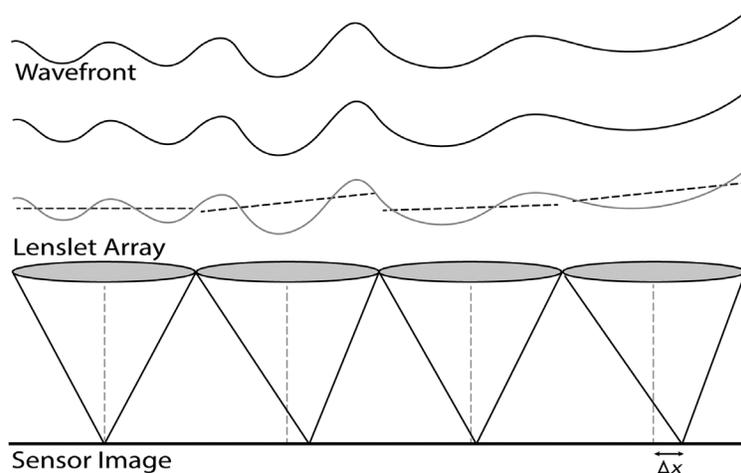


**Fig. 1** AO Loop diagram, showing the benefit of adaptive optics when switched on (loop closed) vs switched off (loop opened).

challenging engineering problem if the numerical methods are to be applied to on-sky data as the method is complex, and any parameters characterizing the real, on-sky observations (e.g., actual turbulence strength or wind speed) that are not perfectly estimated will propagate through the calculations and adversely affect the quality of the estimates. Other approaches<sup>19</sup> can be considered to solve this engineering problem with Fourier-based methods, although including a number of approximations on the PSF model. What we propose here is a high-fidelity estimate that is data driven, and so does not make explicit assumptions.

The SH-WFS used in AO systems is designed to take the wavefront phase information and encode it as an intensity image distributed over small sub-regions of the aperture. It does this with an array of small lenslets that focus the aperture sub-region onto a sensor, creating a spot that is tilted off axis by the average slope of the incoming wavefront sub-region in two dimensions.

Figure 2 shows the one-dimensional (1D) case and how the aberrated wavefront of a sub-region moves the focal point on the sensor off-center. This displacement gives an indication of the average slope of the area of the wavefront covered by the sub-aperture. From the sensor image for all sub-apertures, centroider algorithms are used to find the center of the spots and so a granular map of the slopes is created and passed on to the mirror control system.



**Fig. 2** SH WFS lenslet diagram showing displacement of spots due to wavefront perturbations.



**Fig. 3** SH WFS lenslet spots.

Slope measurements made from centroider data are lossy as there are limits to how small the sub-apertures can be with a limiting factor on the amount of light required per sub-aperture for effective measurements. Also, any non-linear information is lost when the algorithm picks the centroid of the spot for each sub-aperture, reducing the image to points on an  $x, y$  plane. Figure 2 shows the off-axis measurements ( $\Delta x$ ) that are used to measure the wavefront sub-aperture slopes, where the higher-order wavefront information is lost. The size of the lenslets limits the spatial frequency that can be measured, behaving like a low-pass filter.

The sensor image captured at each subaperture corresponds to a low fidelity PSF where the captured irregular patterns of light intensity correspond to a representation of higher-order information about the wavefront profile. A depiction of such patterns is given in Fig. 3, which shows a portion of a simulated SH-WFS image. For the intuition about what is being lost using centroider algorithms, Fig. 2 gives a simplified 1D schematic of the wavefront, lenslet array, and sensor image. The dashed lines drawn in the turbulent wavefront above lenslets represent the gradient measured by a centroider, here clearly missing out important details about high-frequency turbulence. It remains an open question in astronomy, to quantify exactly how much information is being lost in this setting depending on the actual instrument design (e.g., number of sub-apertures, number of pixels per sub-apertures, measurement wavelength, etc.) While we are investigating only the SH WFS in this paper, other sensors such as the Pyramid WFS and curvature sensors are used for wavefront sensing. In Ref. 20, we apply our methodology to Pyramid WFS.

### 2.1.1 COMPASS simulation software

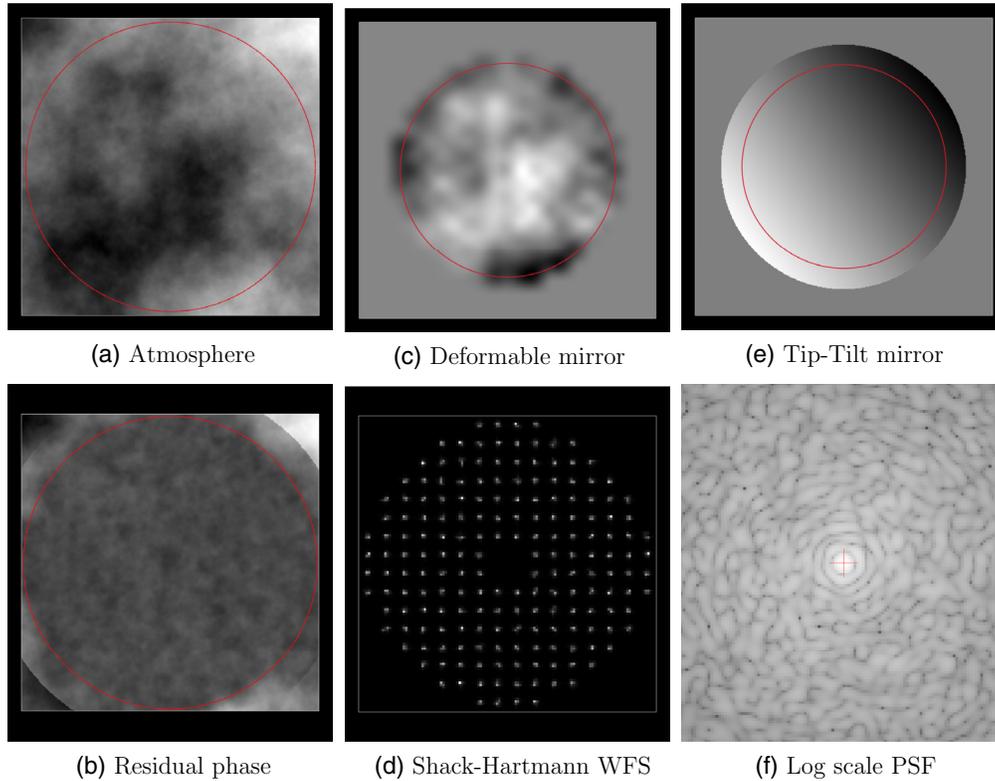
The COMPASS AO simulation software<sup>4</sup> simulates atmospheric conditions, telescope, and AO system to create accurate simulated residual wavefront and WFS images used to train the CNNs.

COMPASS is a GPU-accelerated AO loop simulator with a comprehensive application programming interface (API) that allows simple integration with python code. Highly detailed parameter information can be input to generate specific atmospheric conditions and other AO loop characteristics such as sensor noise and control loop delay. This is perfect for both generating training data for the CNNs and also for testing ranges of conditions for inference performance with trained models. A sample of the data is available in Fig. 4. Figures 4(b) and 4(d) show the residual wavefront phase the SH-WFS images, respectively, these are the two images we will use for training data.

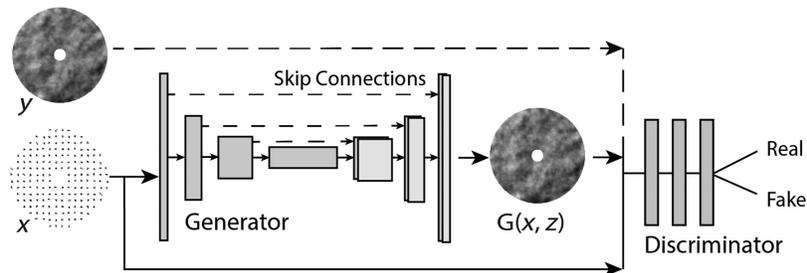
## 3 Our Network-Based Approach

We estimate the residual phase by adapting an ANN for image-to-image translation,<sup>21</sup> where Fig. 5 gives a visual breakdown of the network. This design is a conditional generative adversarial network (cGAN), with the translational encoding performed by a UNet and the adversarial training performed using a Markov discriminator. The network learns to take an input of an SH WFS image and output the inferred wavefront phase.

To motivate the UNet generator component of the network, it can be compared to the similar and widely used auto-encoder—a CNN—that is used for image transformation. An auto-encoder encodes an image to some latent variable through successive convolutional layers, and then



**Fig. 4** Typical data available through COMPASS simulator: (a) The simulated atmospheric phase screen, (b) residual phase screen, (c) deformable mirror shape, (d) Shack-Hartmann wavefront sensor image, (e) ‘tip-tilt’ mirror shape, (f) point-spread function in log scale.



**Fig. 5** cGAN architecture – UNet/PatchGAN.

through deconvolutional steps generates a new image from the latent variable. An auto-encoder learns to transform images minimizing a reconstruction loss and can be used for several applications. For our purposes an auto-encoder is not ideal, because it is deterministic by design, and because we need to preserve some structure from the original image in our application, such as spatial relationships for translation. The UNet design adds skip connections, where information from layers of the encoder is transported to corresponding decoder layers via concatenation, allowing for some structure from the input image to be preserved. Because we aim to translate WFS images from sensor data with incomplete information we cannot map from image-to-image in a deterministic manner as there will be many possible wavefront phase images that could be represented by each input image. To avoid the deterministic nature of the auto-encoder (and UNet) structure, some stochastic process needs to be added to allow for variability in the output. This is accomplished by introducing noise to the network via network dropout ( $z$ ), where Gaussian noise is ineffective because this approach learns to filter it.

Considering the variational auto-encoder (VAE) as an alternative—it also has the ability to create varied translations from an input image and does not have deterministic outcomes due to

encoding and sampling from distributions. The output images from a VAE tend to be blurry and faint, which is not ideal for our application, as we find this occurs in important regions of the wavefront phase. The loss function for a VAE must be carefully designed, which is additionally very difficult to do in practice. By contrast, a GAN<sup>18</sup> has the benefit of learning a loss function, and so simplifies the loss design problem associated with VAEs, as well as tending toward sharper output images, while adding complexity to the network with the addition of a CNN classifier forming the discriminator network that is trained simultaneously with the generator.

A conditional GAN improves the performance of the GAN by including a ‘semantic-image’ in the discriminator as a paired image with either the “real” or “fake” image, which acts as a label for the distribution generated, adding supervision which further improves the sharpness and accuracy of the translated image.

The discriminator in Fig. 5 is a PatchGAN discriminator, also known as a Markov discriminator.<sup>22</sup> This discriminator architecture operates by classifying local image regions, and is broadly motivated in computer vision applications due to the speed of inference (i.e., local inference is relatively fast), and the quality of PatchGAN architectures in preserving complex image detail, such as texture. The discriminator component is a convolutional classifier, trained simultaneously with the generator, with the objective to maximize the value of Eq. (2).

The overall objective function [Eq. (3)] combines the cGAN loss [Eq. (2)], with the L1 reconstruction loss terms [e.g., of the form in Eq. (4)] that provide strong guidance to learning of low-frequency structure. It is noteworthy that a second reconstruction loss term is added in Eq. (3) to reinforce the reconstruction loss where the network underestimates the upper and lower extreme phase values  $G_M$  (or G “masked”). These regions of the wavefront with the largest perturbations substantially impact the important metrics in our setting, and our ablation testing of this parameter shows this term is required. The approach is to simultaneously train: (i) a generator,  $G(x, z)$ , that models the distribution of wavefront phases that are consistent with the input WFS image  $x$ —i.e.,  $z$  is a noise term, specifically dropout noise, and (ii) a discriminator,  $D(x, y)$ , that estimates the probability that a pair  $(x, y)$ , comprising a wavefront phase image  $y$  and a corresponding WFS image  $x$ , are “real.”

$$\mathcal{L}_{\text{cGAN}}(G, D) = E_{x,y}[\log(D(x, y))] + E_{x,z}[\log(1 - D(x, G(x, z)))], \quad (2)$$

$$G^* = \arg \min_G \max_D \mathcal{L}_{\text{cGAN}}(G, D) + \lambda \mathcal{L}_{L1}(G) + \lambda_M \mathcal{L}_{L1}(G_M), \quad (3)$$

$$\mathcal{L}_{L1}(G) = E_{x,y,z}[\|y - G(x, z)\|_1]. \quad (4)$$

It is worth noting that, by removing the cGAN loss term from the loss function and the discriminator component, the cGAN schema can very easily be modified into a standard feed-forward CNN setting. Specifically, here we are thereby reduced to the scenario of training a UNet, the generator component within the cGAN schema. Removal of the PatchGAN discriminator has the effect of limiting capacity to learn about the fine detail of the phase image. Consequently, the trained network is only good at estimating low-frequency parts of the signal. The contribution of the PatchGAN is to enable a trained network to model the distribution of high-frequency information in the phase images, and thereby generate phase estimates that contain accurate high-frequency information from the target distribution. We demonstrate this effect in the following results section by contrasting the UNet results to the complete cGAN network, validating our motivation for using a cGAN for PSF reconstruction.

### 3.1 Network Architecture

We adapted our network architecture and code from Ref. 21, where much of the architectural details remains the same, and we will follow the same labeling conventions. Ck denotes a convolution-batchnorm-ReLU layer with  $k$  filters and CDk denotes a convolution-batchnorm-dropout-ReLU layer. Dropout rates, stride, downsample scaling, and upsample scaling, are all determined as per the literature mentioned above. Choice of hyper-parameters, where not specified in literature, were made through experimentation.

### 3.1.1 Generator architecture

Our network uses a UNet generator, consisting of an encoder, a decoder, and skip connections between all layers as shown in the following layer structures:

UNet encoder:

*C64-C128-C256-C512-C512-C512-C512*

UNet decoder:

*CD1024-CD1024-CD1024-C1024-C1024-C512-C256-C128*

### 3.1.2 Discriminator architecture

The Markovian discriminator architecture by layer:

*C64-C128-C256-C256*

## 3.2 Data Transformations

The raw data from astronomical instruments and simulators requires some transformation to be amenable to the translational architecture we have just discussed. First, the piston mode—i.e., a constant phase shift across the full aperture—is removed from the residual phase data, since it is not measured by the WFS. This is done by subtracting the average value of the wavefront phase that is inside the pupil from the phase array. Experiments where the training data was not adjusted to remove piston experienced early mode collapse that prevented effective training. Second, the residual phase data is normalized to sit in the accepted range  $[0,1]$  for the network by dividing through by a constant value, with null-piston fixed at 0.5, so that the amplitude is restored for application by multiplying any inferred image from the network by the same constant, reversing the normalisation.

The normalisation factor is a tuning parameter, with high values scaling small wavefront errors too much, leading to mode collapse. When normalized perfectly, so the largest value in the training data is exactly 1, we find the trained network does not perform well in generalization evaluations, e.g., where we infer a residual phase in turbulence unseen during training. In our work we have set this factor to 10, leaving some headroom over the minimum required value,  $\sim 7$ . This normalized range allows for inference of wavefront phase estimates with stronger turbulence ( $r_0$ ) than the training dataset to stay within the range  $[0,1]$ , while not reducing the amplitude of the lower turbulence data to the point of being ignored during training. As the wavefront phase amplitude is related to the degree of turbulence ( $r_0$ ), this is a key relationship for the network to learn. Inference of WFS data that exceeds the normalization range will decrease estimation accuracy by generating artifacts in the wavefront phase estimates.

The WFS image is also normalised, again dividing by a constant value which is slightly above the maximum. The WFS scale is preserved because the WFS amplitude information, along with the shape of the WFS spots, is the additional nonlinear information captured using our estimation method. In all cases, our networks use a constant scaling factor of 1.2 million for WFS images. This value is chosen according to the magnitude of the guide star and optical throughput to the WFS subapertures, with our simulations using a fixed guide star of magnitude 10.

## 4 Simulation Results

### 4.1 COMPASS Parameters

Parameters for simulation were selected to demonstrate performance for realistic large telescope AO loop scenarios. The degree of turbulence is defined by the so-called Fried parameter,

**Table 1** Simulation parameters for training data.

Telescope parameters	
Diameter	8 m
Simulated atmospheric parameters	
Number of layers	1
$r_0$	0.093 to 0.400 m
Wind velocity	10 ms <sup>-1</sup>
Target parameters	
Wavelength $\lambda_t$	1.65 $\mu\text{m}$
WFS parameters	
Number of sub-apertures	16 $\times$ 16
Number of pixels per sub-aperture	8 $\times$ 8
Wavelength $\lambda_{wfs}$	0.5 $\mu\text{m}$
AO parameters	
Loop frequency	500 Hz
Delay	2 frames
Integrator gain	0.4
DM parameters	
Number of DM actuators	17 $\times$ 17
1 tip-tilt mirror	

$r_0$ , which is a measure of the coherence scale of the turbulence<sup>2</sup> and depends on the observing wavelength. Typically the real-world operating conditions for  $r_0$  are in the range of 0.16 m—at visible light wavelengths for the lower range of atmospheric turbulence—to 0.05 m for very extreme conditions. For the purposes of this study, we focus training around the typical  $r_0$  value of 0.093 m, and when interrogating a trained network for robustness to a range of atmospheric turbulence we have  $r_0$  ranging from 0.08 to 0.18.

AO loop data has been simulated for a typical wind speed of 10 ms<sup>-1</sup>. About 50 k sample image pairs are generated for each of the  $r_0$  values in [0.093, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40]. We thereby are able to train the network with a range of turbulence scenarios so that it learns to robustly estimate wavefronts in a range of turbulent conditions that would be expected for on-sky operating conditions. The training and evaluation atmospheres are different. In particular, when interrogating network model performance, in a control setting below and otherwise, we use simulations that are seeded uniquely, and therefore are of atmospheres not seen during training. See Table 1 for simulation parameters.

## 4.2 Network Parameters

We adapted our network architecture and code from Isola et al.,<sup>21</sup> where much of the architecture details remain the same. Dropout rates, stride, downsample scaling, and upsample scaling, are all determined as per the literature mentioned above.

Both generator and discriminator networks used 64 filters. The generator performs better with 64 filters over trials run with 32 or 16, however, this comes with a computational cost as the number of parameters is significantly increased, which in turn increases training time

and hardware memory requirements. Datasets consisted of 350 k image pairs selected in random order with a batch size of 1, and validation is done by varying the pseudo-random seed for atmospheric generation in COMPASS.

Our adaptation of the cGAN architecture includes an additional weighted loss parameter, which we optimized for our setting. We found that the original loss configuration from the literature is unable to yield a network that accurately reconstructs extreme maximum and minimum values. This is a concern because the consequent discrepancies in generated phase images have a substantial effect on the overall performance in our application, and specifically in PSF reconstruction. To compensate for this, a second  $L1$  loss term was introduced for the extreme values (the upper and lower 10% of phase amplitude), increasing the effect of the  $L1$  loss term on these regions of the training image by masking out all but the extreme phase errors. This extends Eq. (3) with the additional masked  $L1$  loss term, where  $G_M$  defines the region of extremes masked in the generated image  $G$ , and  $\lambda_M$  is the weighting coefficient for the added masked loss term. All cGAN simulation results in this paper use a single trained network with hyper-parameters  $\lambda = 150$  and  $\lambda_M = 30$ .

### 4.3 Evaluation Metrics

Assessing the predictions of the trained models will follow methods used by the state-of-the-art in long-exposure PSF reconstruction.<sup>1,3</sup> Our network infers a short-exposure wavefront image that is converted to a short-exposure PSF as per Eq. (1), and then averaged over 20  $k$  samples. The simulated long-exposure PSF is generated in COMPASS and compared with the inferred long-exposure PSF, and also a PSF reconstructed using the state-of-the-art method<sup>3</sup> using identical simulation parameters. By training on simulated data for given telescope parameters, we hope to create a network that can infer accurate wavefront phase estimates from real WFS data in future work. In this paper, we are using simulated telescopes and so cannot be totally confident that this will work for real WFS data (Figs. 6 and 7).

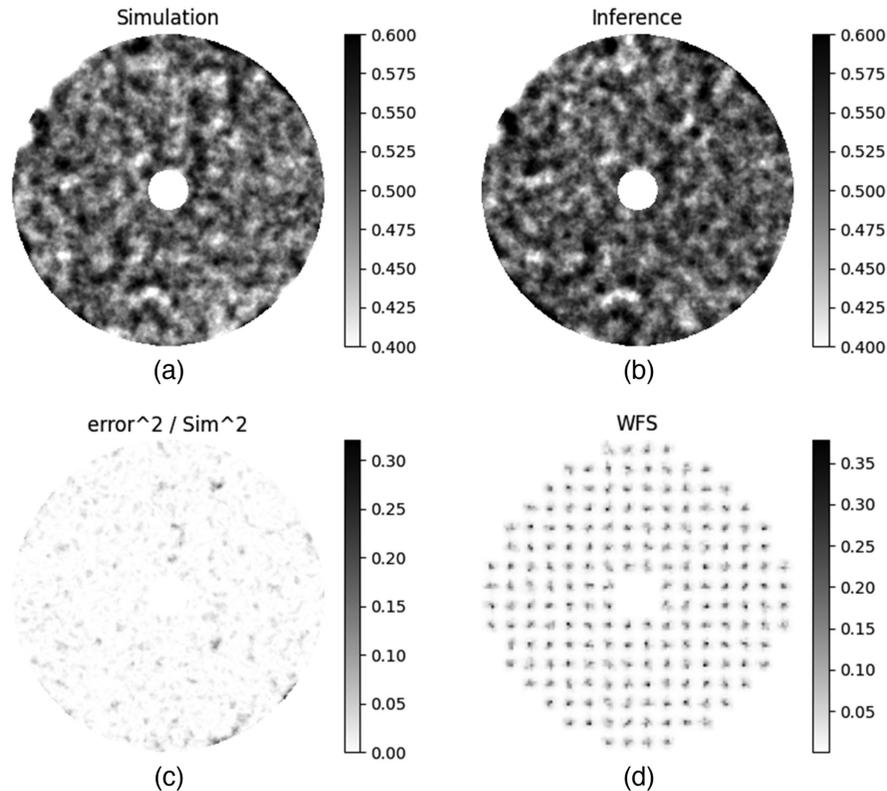
The PSFs are compared (Figs. 8 and 9) in log scale due to the characteristic shape of the PSF having a central peak that rapidly falls off towards the edges of the image in the range of [0, 1]. In the COMPASS simulator the flux from the PSF is normalised, so that the central peak maximum corresponds to the Strehl ratio (SR), a measurement of the intensity of the image against the one obtained on an ideal system. SR is used to assess image quality and performance of AO systems. Our ability to accurately estimate the SR is a key metric. A second important feature of PSF comparison used here is the overall fit of the PSF shape outside of the central peak, i.e., PSF halo. The residual phase error information is contained in the halo, with the low spatial frequency information close to the central peak and the higher spatial frequency information radiating outwards. The majority of the residual will be in the low-order information near the center, so this is the most important region to match correctly. Higher spatial frequencies will affect the contrast in the halo. To compare both of these elements of the PSF estimates, an  $X, Y$  cross-section view is compared in logarithmic scale.

### 4.4 Results and Analysis

Using the methodology description in the previous section, we run experiments using a variety of Nvidia GPU hardware (such as V100 GPUs) and PyTorch v1.8.0 to train our networks, using numerical models in the work of Ferreira et al.<sup>3</sup> using an end-to-end error breakdown model of the PSF as a reference benchmark. (Code available at <https://github.com/GANs4AO/>.)

This is an exciting result, as it is visible in the inferred residual phase image, as evidenced by accurate features showing much finer detail than just the average slopes across the width of the lenslets. This is evidence that there is high-order information being interpreted by the network from a single SH-WFS image.

The PSFs are broken down by cross-section in Fig. 8 where the SR can be compared, and the profile of the halo can be observed. We use the methods discussed previously to compare long-exposure PSFs shown in Fig. 9, including the model-based reconstruction (reference) and network-based (inferred) results with the simulated ground truth (COMPASS).



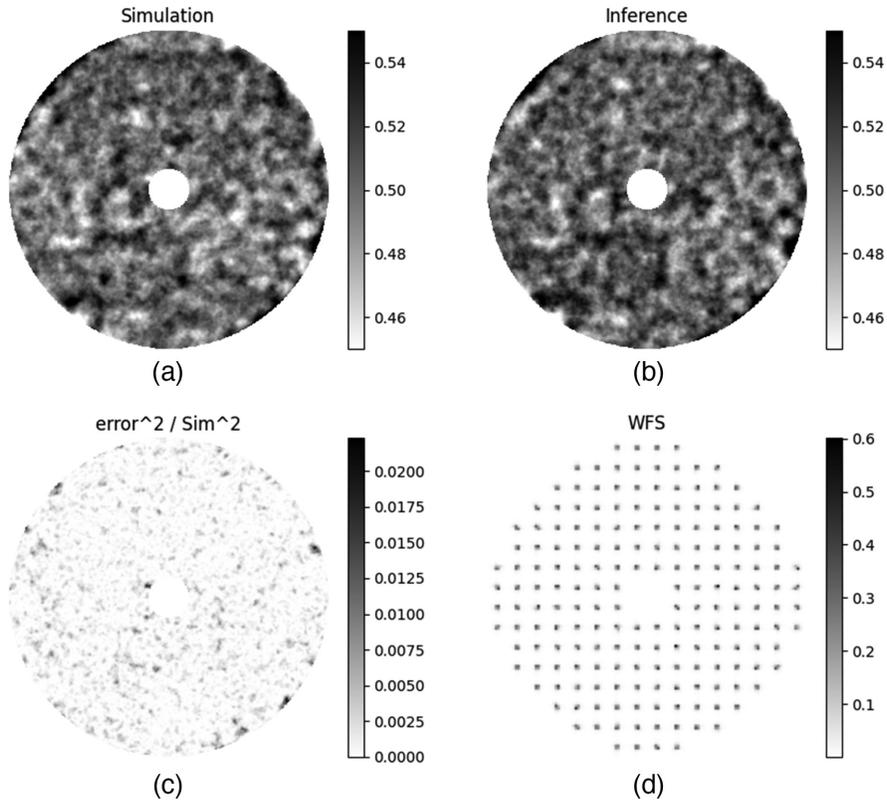
**Fig. 6** (a) High turbulence residual wavefront simulated and (b) inferred, with error measurement shown as square of difference relative to (c) square of true amplitude. (d) Shown with corresponding WFS image.

From the result in Fig. 8, the inferred PSF has slightly better accuracy in SR than the model-based reference PSF, observable in the relative errors of the inferred (solid green line) and the model-based reference (dashed orange line) and on comparison in the tabulated SR data shown in Table 2. (It is noteworthy that Fig. 8 shows results for  $r_0 = 0.093$  m. The range of charts corresponding to tabular values are available in Appendix.) To verify the robustness of our trained translational network we run the same long-exposure experiments for a range of  $r_0$  values over the same network, including some that were not available in the training dataset. Table 2 shows the resulting SR for the simulated ground truth, the referenced numerical model benchmark and our inferred results. As shown by the data, the inferred results from our network demonstrate a remarkable robustness to changing atmospheric conditions and in most cases improve on the numerical model benchmark, over much of the typical range of real-world atmospheric viewing conditions.

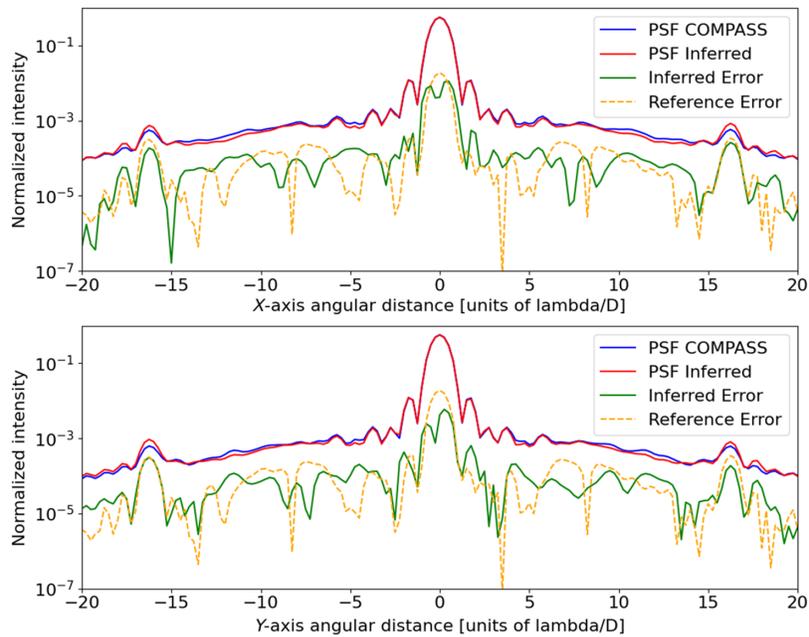
Figure 8 shows a direction-dependent error, which ends up comparable for both estimation methods, for the specific direction chosen in this case. The difference between our network-based approach and the model-based approach becomes stark when we examine the direction independent error. The circular average plot of Fig. 9, which is a direction independent estimate of the error, shows how the symmetry assumptions of the model-based approach have a profound impact on the overall error. The bump highlighted in the top-right plot from Fig. 9 indicates a potential sensitivity loss of  $10^3$  with the model-based approach. Practically, it means that if the astronomer is looking for a faint companion (e.g., an exoplanet) in this area of the image, the sensitivity of the observations after post-processing using this PSF estimate will be  $10^3$  times lower if using the model-based approach.

#### 4.5 Comparison with Feed-Forward Networks

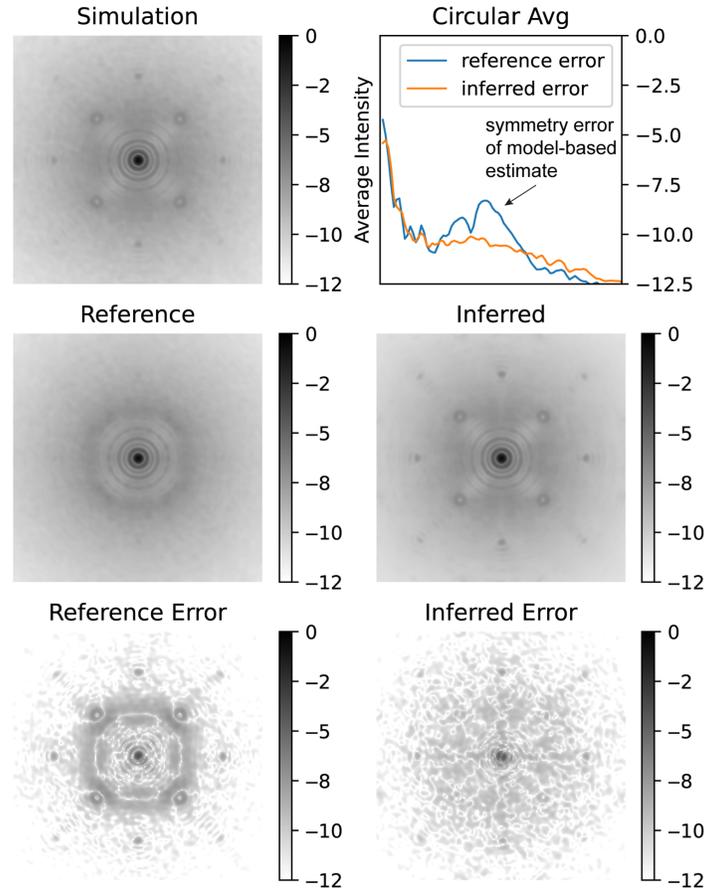
Removing the discriminator and the cGAN loss term from the loss function we transform our cGAN into a UNet. Training for this network is a great deal faster, only requiring a few epochs to



**Fig. 7** (a) Low turbulence residual wavefront simulated and (b) inferred, with error measurement shown as (c) square of difference relative to square of true amplitude. (d) Shown with corresponding WFS image (bottom right).



**Fig. 8** Long-exposure PSF cross-section comparison in units of log normalized intensity ( $r_0 = 0.093$  m).



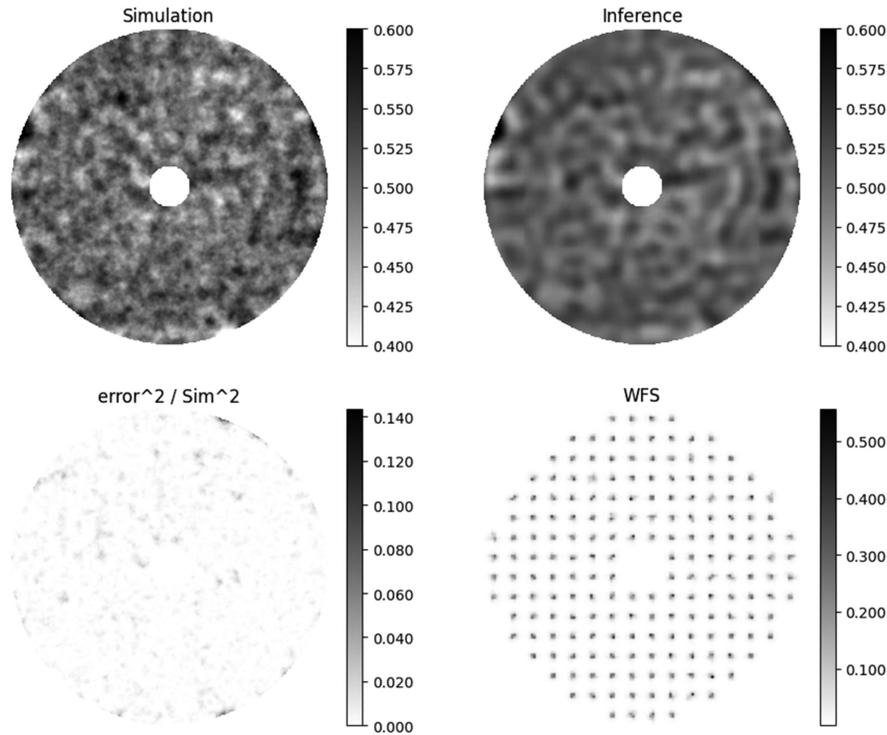
**Fig. 9** Long-exposure PSF ( $r_0 = 0.093$  m) in log scale. Top: Simulated PSF (left), circular average of PSF errors (right). Middle: Reconstruction from literature (left) and inferred result (right). Bottom: Error as absolute difference.

**Table 2** Long-exposure SR robustness of translation results to atmospheric turbulence variation (Fried parameter – ' $r_0$ '), referenced to numerical model benchmark and simulated truth.

$r_0$ (m)	Long exposure SR		
	Simulation	Inferred	Reference
0.080	0.490	0.514	0.463
0.093	0.572	0.568	0.553
0.100	0.608	0.599	0.592
0.120	0.690	0.689	0.679
0.150	0.771	0.773	0.763
0.180	0.823	0.829	0.818

train. This represents a dramatic reduction in training time to that of a cGAN. However, there are drawbacks to the UNet, as without the discriminator the ability to learn high-frequency information about the wavefront is lost.

GAN training proceeds according to an adversarial loss regime. We employ a Markovian discriminator specifically, otherwise known as a PatchGAN. In that setup, the trained generator



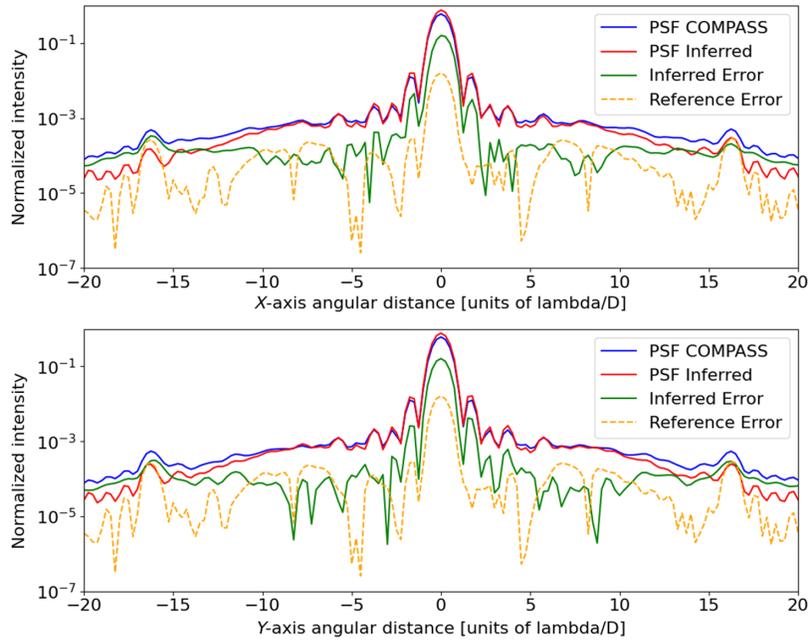
**Fig. 10** Residual wavefront simulated (top left) and UNet-only generated inference (top right), with error measurement shown as square of difference relative to square of true amplitude (bottom left). Shown with corresponding WFS image (bottom right). Notice that the UNet can only capture low-frequency wavefront information evident from the lack of detail in the estimate.

models a distribution of phase images that locally matches the statistics of the training corpus. A generator—in our case feed-forward UNet—trained without this adversarial loss would be optimized according to the  $L1$  loss terms alone. In this section we showcase the importance of the adversarial loss, showing that training the feed-forward network using  $L1$  terms alone fails to yield phase estimates with realistic or accurate fine detail.

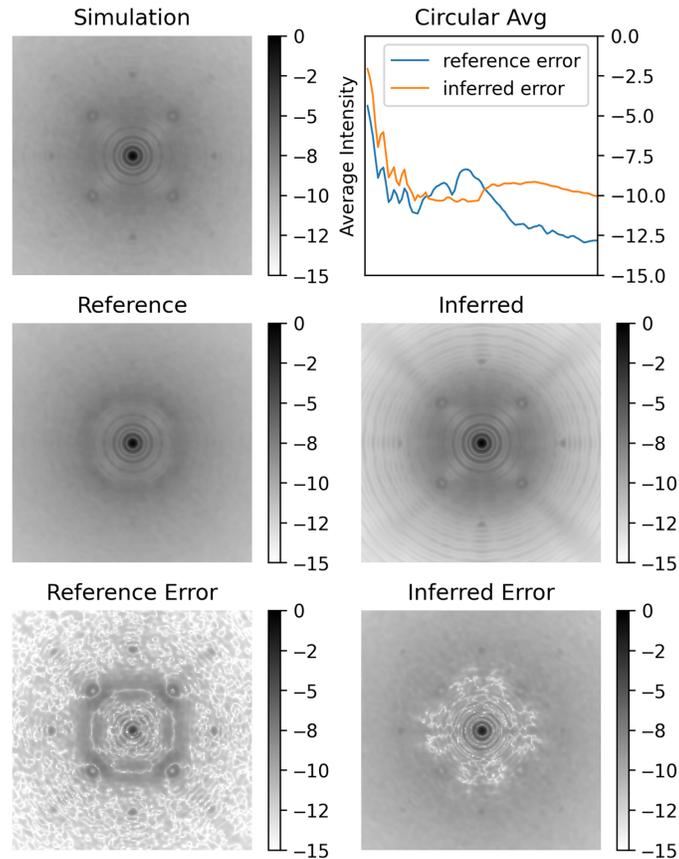
By inspecting the phase estimate inferred by the UNet in Fig. 10, it is clear that there is a lack of detail in the image. This lack of detail has the effect of lowering the fidelity of the phase estimate and when transforming to a PSF [Eq. (1)]. The high-frequency information is missing, altering the shape of the PSF as seen in Figs. 11 and 12. This effect of filtering the high-frequency information out of the estimate hampers the PSF reconstruction. Clearly a cGAN is a better choice for PSF reconstruction with image-to-image translation, because the network can form a more accurate phase estimate by learning about: (i) the low-frequency structure of the phase via the  $L1$  loss, and (ii) the high-frequency information with the cGAN loss. The GAN approach leads to good PSF reconstruction using the estimated wavefront information.

## 5 Discussion and Future Work

We have shown that our approach can utilize high order information from the SH-WFS in an AO loop that is inaccessible with current centroider algorithms applied by current practical methods and theoretical models for wavefront estimation. With this high order data, our translational network can accurately estimate the wavefront from just the WFS image, with no loss of accuracy when compared with best case theoretical models. Using network estimates we are able to improve the quality and accuracy of corresponding long-exposure PSFs, with the consequence in our application being the ability to increase the value of astronomical images according to deconvolutional post-processing workflows. This work is a proof of concept based on a very realistic end-to-end simulator, but applicability on real systems will need fine tuning and will require new



**Fig. 11** UNet-only generated long-exposure PSF cross-section comparison in units of log normalized intensity ( $r_0 = 0.10$  m). Notice the poor accuracy of the UNet estimate for both central peak and also for higher angular distance.



**Fig. 12** Long-exposure PSF ( $r_0 = 0.10$  m) in log scale. Top: Simulated PSF (left) and circular average of PSF errors (right). Middle: Reconstruction from literature (left) and UNet-only generated inference result (right). Bottom: Error as absolute difference. Note that the PSF reconstruction from the UNet-generated PSF is clearly unable to match the accuracy of the literature (or the cGAN method).

dedicated strategies for optimal training. In future work, we will investigate the accuracy of these networks with a more rigorous statistical analysis.

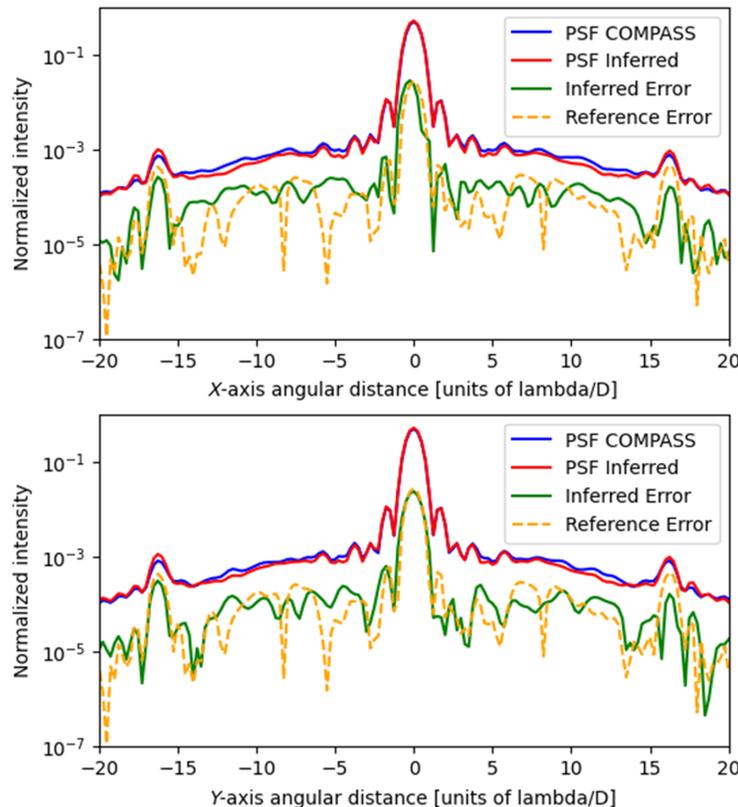
Using a translational network is simpler than a model-based approach, which solves an engineering problem. We do not require strong assumptions on the nature of the stochastic process or system geometry, nor a large database to be processed to lead to comparable or better reconstruction accuracy. It is noteworthy that storing and archiving telemetry data is an existing and significant problem for post-processing workflows.

In future work, we shall investigate a range of network architectures and loss regimes, such as Wasserstein GANs,<sup>23,24</sup> and approaches using maximum mean discrepancy.<sup>25–28</sup> An ideal approach would have very low sensitivity to data preparation parameters related to scaling, and to some extent network hyperparameters, while achieving the excellent high quality reconstructions we have demonstrated.

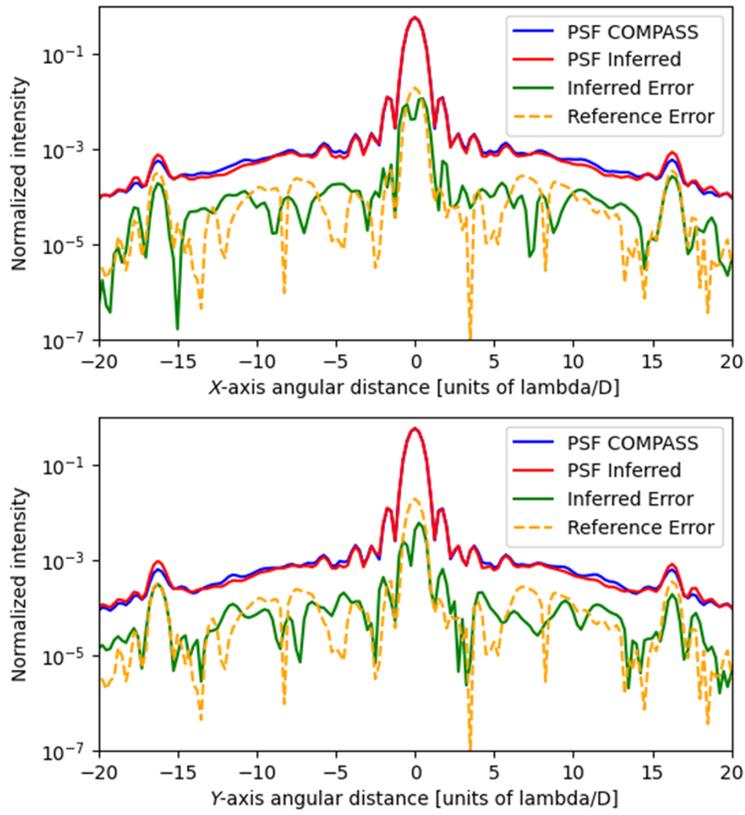
We are also examining applications of our phase reconstruction in real-time control with impressive early results,<sup>20,29</sup> and are encouraged about its applicability given we are measuring an inference time of 0.34 ms on retail desktop GPUs. With optimization and dedicated hardware<sup>30,31</sup> and the COSMIC framework platform,<sup>32</sup> wavefront inference from pre-trained networks has potential for hard real-time AO loop control, and not just in postprocessing workflows via PSF reconstruction. Future lab experiments will allow for verification of loop control with real sensor equipment. The speed and simplicity, combined with demonstrated accuracy of our method is of great interest in the AO community and shows a lot of promise for the in-construction ELT.<sup>33</sup>

## 6 Appendix: Long-exposure PSF Cross-sections for Robustness to Turbulence

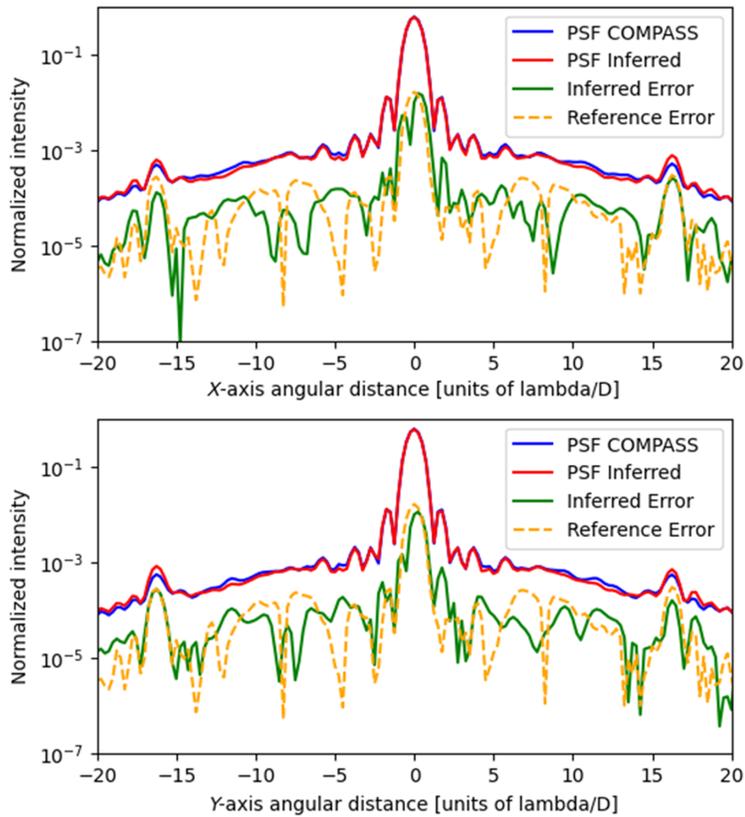
Figures 13–18 show the robustness of cGAN phase estimates for a single trained model over a range of atmospheric turbulence parameters ( $r_0$ ), inferred from WFS image alone.



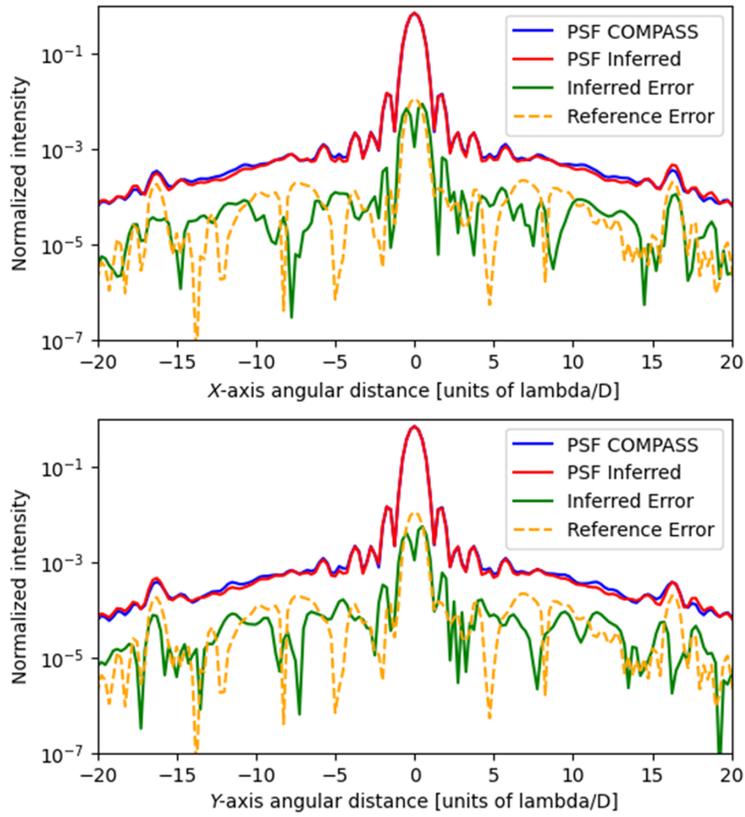
**Fig. 13** Long-exposure PSF ( $r_0 = 0.080$  m).



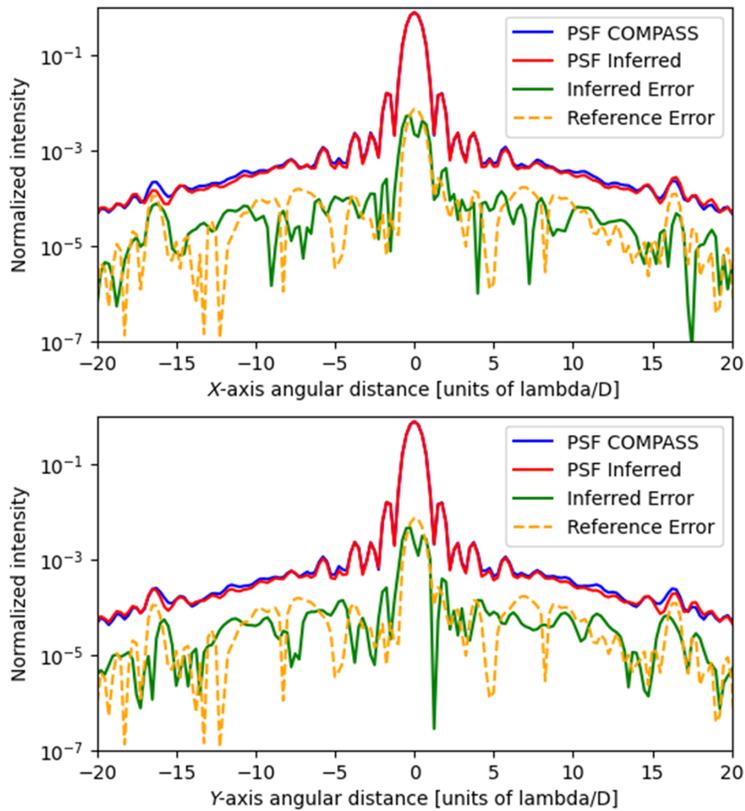
**Fig. 14** Long-exposure PSF ( $r_0 = 0.093$  m).



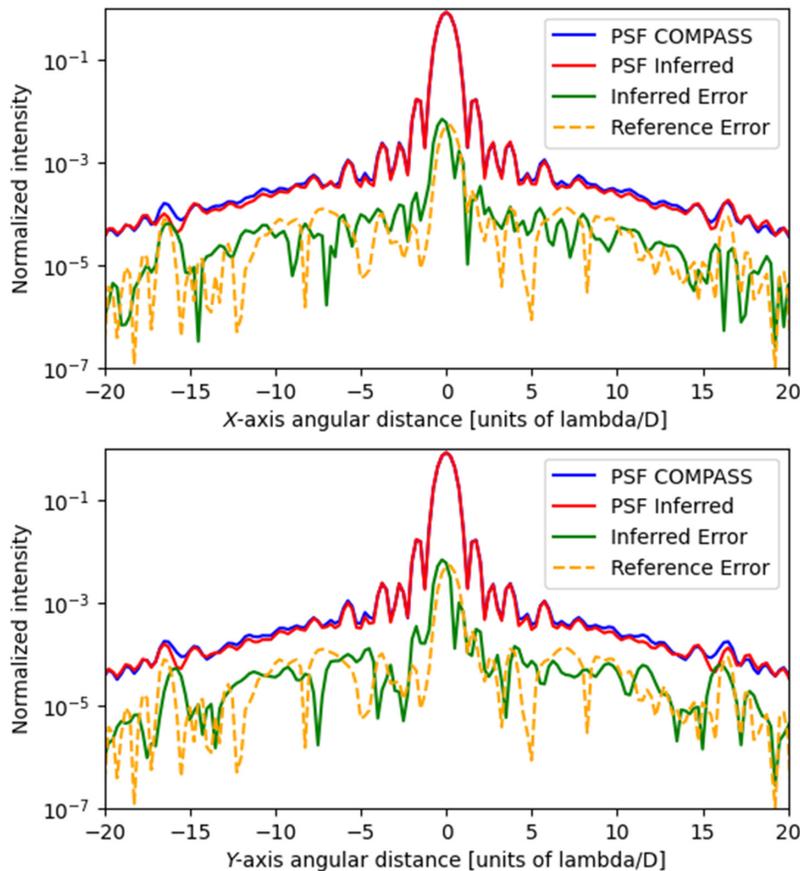
**Fig. 15** Long-exposure PSF ( $r_0 = 0.100$  m).



**Fig. 16** Long-exposure PSF ( $r_0 = 0.120$  m).



**Fig. 17** Long-exposure PSF ( $r_0 = 0.150$  m).



**Fig. 18** Long-exposure PSF ( $r_0 = 0.180$  m).

## Acknowledgments

Many thanks to Florian Ferreira for donating his time and knowledge assisting with COMPASS, Felipe Trevizan for guidance on manuscript editing and Mark Burgess for reviews and feedback. Thanks also to Bartomeu Pou Mulet and Hao Zang for numerous discussions that enhanced knowledge in this field of research. This work was supported in part by Oracle Cloud credits and related resources provided by the Oracle for Research program. This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI Australia), an NCRIS-enabled capability supported by the Australian Government. This paper was originally published in Proceedings – Adaptive Optics Systems VIII, SPIE Astronomical Telescopes and Instrumentation 2022.<sup>34</sup>

## References

1. R. Wagner et al., “Overview of PSF determination techniques for adaptive-optics assisted ELT instruments,” in *AO4ELT 2019—Proc. 6th Adapt. Opt. for Extremely Large Telesc.* (2019).
2. F. Roddier, Ed., *Adaptive Optics in Astronomy*, Cambridge University Press (1999).
3. F. Ferreira et al., “Numerical estimation of wavefront error breakdown in adaptive optics,” *Astron. Astrophys. (Berlin)* **616**, A102 (2018).
4. F. Ferreira et al., “COMPASS: an efficient GPU-based simulation software for adaptive optics systems,” in *Proc. – 2018 Int. Conf. High Perform. Comput. and Simul., HPCS 2018* (2018).
5. R. Swanson et al., “Closed loop predictive control of adaptive optics systems with convolutional neural networks,” *Mon. Not. R. Astron. Soc.* **503**(2) (2021).

6. H. Guo et al., “Wavefront reconstruction with artificial neural networks,” *Opt. Express* **14**(14), 6456–6462 (2006).
7. S. Weddell and R. Webb, *A Neural Network Architecture for Reconstruction of Turbulence Degraded Point Spread Functions*, University of Canterbury, Electrical and Computer Engineering (2007).
8. H. Guo et al., “Improved approach machine learning for wavefront sensing,” *Sensors (Basel)* **19**(16), 3533 (2019).
9. O. Guyon et al., “High contrast imaging at the photon noise limit with self-calibrating WFS/C systems,” *Proc. SPIE* **11823**, 1182318 (2021).
10. K. Fukushima, “Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,” *Biol. Cybern.* **36**, 193–202 (1980).
11. Y. LeCun et al., “Backpropagation applied to handwritten zip code recognition,” *Neural Comput.* **1**, 541–551 (1989).
12. V. Nair and G. E. Hinton, “Rectified linear units improve restricted Boltzmann machines,” in *Proc. 27th Int. Conf. Mach. Learn. (ICML-10)*, Haifa, June 21–24, 2010, J. Fürnkranz and T. Joachims, Eds., Omnipress, pp. 807–814 (2010).
13. A. Maas, A. Hannun, and A. Ng, “Rectifier nonlinearities improve neural network acoustic models,” in *Proc. Int. Conf. Mach. Learn.* (2013).
14. A. G. Ivakhnenko, “Polynomial theory of complex systems,” *IEEE Trans. Syst. Man, Cybern. SMC-1*(4), 364–378 (1971).
15. G. Cybenko, “Continuous valued neural networks with two hidden layers are sufficient,” tech. rep., Department of Computer Science, Tufts University, Medford, MA (1988).
16. O. Ronneberger, P. Fischer, and T. Brox, “U-net: convolutional networks for biomedical image segmentation,” *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
17. D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” in *3rd Int. Conf. Learn. Represent., ICLR 2015, Conf. Track Proc.*, San Diego, California, May 7–9, Y. Bengio and Y. LeCun Eds. (2015).
18. I. Goodfellow et al., “Generative adversarial nets,” in *Adv. Neural Inf. Process. Syst.*, Z. Ghahramani et al., Eds., Vol. 27, Curran Associates, Inc. (2014).
19. B. Neichel et al., “TIPTOP: a new tool to efficiently predict your favorite AO PSF,” *Proc. SPIE* **11448**, 114482T (2021).
20. B. Pou et al., “Model-free reinforcement learning with a non-linear reconstructor for closed-loop adaptive optics control with a pyramid wavefront sensor,” *Proc. SPIE* **12185**, 121852U (2022).
21. P. Isola et al., “Image-to-image translation with conditional adversarial networks,” in *Proc. – 30th IEEE Conf. Comput. Vision and Pattern Recognit., CVPR 2017* (2017).
22. C. Li and M. Wand, “Precomputed real-time texture synthesis with markovian generative adversarial networks,” *Lect. Notes Comput. Sci.* **9907**, 702–716 (2016).
23. M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein generative adversarial networks,” in *34th Int. Conf. Mach. Learn., ICML 2017*, Vol. 1 (2017).
24. I. Gulrajani et al., “Improved training of Wasserstein GANs,” in *Adv. Neural Inf. Process. Syst.* (2017).
25. C. L. Li et al., “MMD GAN: towards deeper understanding of moment matching network,” in *Adv. Neural Inf. Process. Syst.* (2017).
26. M. G. Bellemare et al., “The cramer distance as a solution to biased Wasserstein gradients,” <https://arxiv.org/pdf/1705.10743.pdf> (2017).
27. T. Unterthiner et al., “Coulomb GANs: provably optimal nash equilibria via potential fields,” in *6th Int. Conf. Learn. Represent., ICLR 2018, Conf. Track Proc.*, Vancouver, BC, April 30 – May 3, [OpenReview.net](https://openreview.net) (2018).
28. M. Binkowski et al., “Demystifying MMD GANs,” in *6th Int. Conf. Learn. Represent., ICLR 2018—Conf. Track Proc.* (2018).
29. J. Smith et al., “Enhanced adaptive optics control with image to image translation,” in *Proc. Thirty-Eighth Conf. Uncertain. Artif. Intell., UAI 2022*, 1–5 August 2022, Eindhoven, The Netherlands, Proc. Mach. Learn. Res., J. Cussens and K. Zhang, Eds., PMLR, Vol. 180, pp. 1846–1856 (2022).

30. D. Perret et al., “Bridging FPGA and GPU technologies for AO real-time control,” *Proc. SPIE* **9909**, 99094M (2016).
31. D. Gratadour et al., “MAVIS real-time control system: a high-end implementation of the COSMIC platform,” in *Adapt. Opt. Syst. VII*, L. Schreiber, D. Schmidt, and E. Vernet, Eds., SPIE, p. 114482M (2020).
32. F. Ferreira et al., “Hard real-time core software of the AO RTC COSMIC platform: architecture and performance,” *Proc. SPIE* **11448**, 1144815 (2020).
33. M. Kasper et al., “PCS—a roadmap for exoearth imaging with the ELT,” *The Messenger* **182**, 38–43 (2021).
34. J. Smith et al., “Image-to-image translation for wavefront and PSF estimation,” *Proc. SPIE* **12185**, 121852L (2022).

Biographies of the authors are not available.