Improving Astronomy Image Quality Through Real-time Wavefront Estimation

David Thomas[1], Joshua Meyers[2], Steven M. Kahn[1]
[1]Stanford University
[2]Lawrence Livermore National Laboratory

Abstract

We present a new framework for detecting telescope optics aberrations in real-time. The framework divides the problem into two subproblems that are highly amenable to machine learning and optimization. The first involves making local wavefront estimates with a convolutional neural network. The second involves interpolating the optics wavefront from all the local estimates by minimizing a convex loss function. We test our framework with simulations of the Vera Rubin Observatory. In a realistic mini-survey, the algorithm reduces the power in the optics wavefront by X%, the optics PSF FWHM by Y%, and the Strehl ratio by Z%. The resulting sharper images have the potential to boost the scientific payload for astrophysics and cosmology.

1. Introduction

The signal to noise ratio of most astronomical analyses critically depends on image quality. To maintain optimal image quality during operation, wide-field telescopes deploy active optics systems, which sense aberrations in the optics and correct them in realtime. The primary challenge for these systems is distinguishing the correctable aberrations due to the optics from the dominant and incorrectable aberrations due to atmospheric turbulence. Here we present a new machine learning framework that is capable of extracting the optics aberrations and improving image quality.

While the immediate application is improving image quality in present and future ground-based telescopes, there are also emerging use cases in space. The simultaneous demands for higher quality images and lighter payloads from space telescopes make large foldable mirrors attractive. These large, lightweight mirrors are more susceptible to environmental disturbances and would benefit from active optics control. Prototypes are already being explored [16, 15, 7, 31, 29, 26]. In this work we demonstrate that our method is capable of improving image quality in the challenging ground-based environment. We suspect its performance will improve in the space environment where atmospheric turbulence is absent.

The upcoming ground-based Vera Rubin Observatory (Rubin) has a 3.5 degree field of view and high dimensional optical model that make it the ideal stress test for our framework [12, 1]. The large scientific community behind the Rubin Observatory has developed a mature suite of simulation codes [23, 19, 6] which we used to train and test our model in realistic scenarios. The unlimited supply of virtual telescopes and observations allowed us to assess our method in a more comprehensive range of conditions than would be possible with a real instrument.

The input to our model comes from 4 curvature wavefront sensors in the corner of the Rubin focal plane, shown in Figure 1. Each of these sensors is split into two halfchips which are purposefully offset out of focus. The stars that fall on these sensors produce large ring-like "donut" images. The goal of our algorithm is to constrain optics aberrations attributed to the entrance pupil from the intensity patterns in all the donut images throughout the observation, which can number into the thousands. One of the key challenges is interpolating the optics wavefront across the entire focal plane from donut images in four concentrated regions which collectively cover less than 2% of the total focal plane.

The key breakthrough in our work is the realization that the wavefront sensing problem can be divided into two subproblems that are highly amenable to machine learning and optimization. The first problem is to estimate the local wavefronts, characterized by 18 Zernike coefficients, from individual donut images. The second problem is to interpolate the global optics wavefront, characterized by 56 double Zernike coefficients, from all the local estimates by minimizing a simple convex loss function. The main contributions of this work are:

1) We present a new mathematical framework for extracting the optics wavefront across the field of view.

2) We demonstrate that a convolutional neural network can make reasonable estimates of the local wavefront from donut images.

3) We show that fitting the global wavefront from a mul-

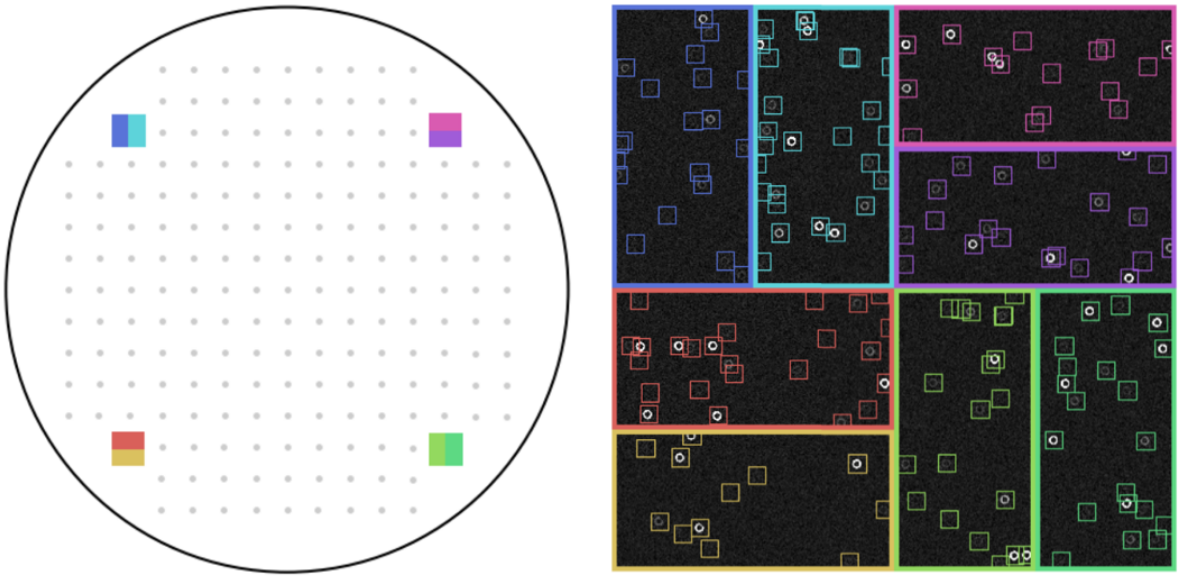

Figure 1. Example of a short caption, which should be centered.

titude of local wavefront estimates can suppress the atmospheric contribution.

4) We run our framework on a realistic mini-survey where it reduces the power in the optics wavefront by X%, the optics PSF FWHM by Y%, and the Strehl ratio by Z%.

Finally, we emphasize that while this work focuses on the context of the Rubin Observatory, our framework extends to all wide-field telescopes with curvature wavefront sensors, and potentially to future space telescopes.

2. Related Work

The potential for neural networks to learn the non-linear mapping between intensity patterns and aberrations in the pupil plane was first recognized in 1990 [4]. Shortly afterwards this potential was realized as neural networks were deployed to detect turbulence induced distortion on the Multiple Mirror Telescope [25] and to detect aberrations in the primary mirror of the Hubble Space Telescope [5]. Others expanded this concept to predict more wavefront components [13], incorporate temporal history [18, 20], compare reconstruction methods [9], and better characterize atmospheric turbulence [30].

In the past decade, convolutional neural networks (CNN) [17] have re-emerged and spurred dramatic advances in computer vision [14, 27, 24, 11]. This has opened up new possibilities for wavefront sensing in astronomy. CNNs can complement conventional iterative estimation approaches by providing good starting estimates [22]. They can even solve the full estimation problem in various scenarios [21, 10, 28]. Combining these networks with CMOS sensors opens up the possibility that guide stars could be used for adaptive optics on 2-4m class telescopes [3]. They also have the potential to sense and correct low-order wavefront aberrations for high-contrast astronomical imaging [2].

Most previous work focuses on sensing the dominant source of wavefront error, atmospheric turbulence or simple distortions, on short timescales, typically 10s of milliseconds. In this work, we study the harder problem of sensing subdominant sources of wavefront error, small alignment and mirror surface perturbations, from images integrated over an entire 15 second exposure. This presents new challenges such as how to best aggregate intensity information from throughout the field of view to suppress the dominant and spatially correlated atmospheric turbulence contribution.

3. Wavefront Sensing Framework

In each observation, we are given n donut image and position pairs $(D_1, \vec{r}_1), \ldots, (D_n, \vec{r}_n) \in \mathbb{R}^{256\times256} \times \mathbb{R}^3$, like those shown in figure X. The goal is to estimate the global wavefront $\mathcal{G} \in \mathbb{R}^{k\times m}$, where k is the dimension of the interpolation basis and m is the dimension of the wavefront basis, which represents the complete optical aberration of the image system. We divide this problem into two subproblems that are highly amenable to machine learning.

Estimating the local wavefront. The first subproblem is to estimate the physical wavefront $w \in \mathbb{R}^m$, at the position r, from the corresponding donut image and position pair (D, r). The entrance pupil is the domain of the wavefront. We use the standard optics convention of representing

the wavefront with m low order Zernike polynomial coefficients. We demonstrate that a convolutional neural network $\varphi : \mathbb{R}^{256\times256} \times \mathbb{R}^3 \to \mathbb{R}^m$ can learn to infer the wavefront from a donut image and position so that $\varphi(D_i, r_i) \approx w_i$.

Interpolating the global wavefront. The second subproblem is to interpolate the global wavefront $\mathcal{G}$ from the estimated local wavefronts and their positions w_i, r_i. Let $W_{\text{est}} \in \mathbb{R}^{k\times m}$ be defined

$$W_{\text{est}} = \begin{bmatrix} \varphi(D_1, r_1)^T \\ \vdots \\ \varphi(D_n, r_n)^T \end{bmatrix} \tag{1}$$

and $Z \in \mathbb{R}^{n\times k}$ be defined

$$Z = \begin{bmatrix} z_1(r_1) & \dots & z_k(r_1) \\ \vdots & \vdots & \vdots \\ z_1(r_n) & \dots & z_k(r_n) \end{bmatrix} \tag{2}$$

where $z_i(r)$ is the ith Zernike polynomial evaulated at r. Then $\mathcal{G}$ can be estimated by minimizing the loss $\sum_{j=1}^{m} \ell(Z\mathcal{G}_j, W_{\text{est},j})$, where $\mathcal{G}_j$ and $W_{\text{est},j}$ are the jth columns of $\mathcal{G}$ and W_{est} respectively. We study the Manhattan ℓ_1, Euclidean ℓ_2, and Huber ℓ_h loss functions. We also explore alternative ways of aggregating the coefficients, including by taking the median over each chip, and clipping outliers by chip.

There are multiple benefits to dividing the problem in this manner. First, we can use specialized techniques for each subproblem. Second, the intermediate variable passed between the two subproblems, the local wavefront, is a physically meaningful quantity. Tracking this adds an additional layer of transparency and robustness to the system. Finally, by training on small donut crops in the first subproblem, as opposed to the full wavefront sensor images, we can reduce the compute required to simulate a training sample by three orders of magnitude. These savings allow us to generate large enough simulated datasets to use deep learning, which would otherwise be intractable.

Algorithm 1: Two-stage algorithm for estimating the global wavefront from donut images.

function Estimate Global Wavefront (I, R)
Input : Full image $I \in \mathbb{R}^{N\times N}$; donut positions $R \in \mathbb{R}^{n\times 3}$
Output: Global wavefront $\mathcal{G}^* \in \mathbb{R}^{k\times m}$
initialize local wavefront estimate $W_{\text{est}} \in \mathbb{R}^{n\times m}$
initialize global Zernike basis $Z \in \mathbb{R}^{n\times k}$
for *donut* i *in* $1 \dots n$ **do**
  $r_i = R[i]$
  $D_i = \text{Crop}(I, r_i)$
  $W_{\text{est}}[i] = \varphi(D_i, r_i)$
  $Z[i] = [z_1(r_i) \ \dots \ z_k(r_i)]$
end
initialize global wavefront $\mathcal{G}^* \in \mathbb{R}^{n\times m}$
for *local Zernike* j *in* $1 \dots m$ **do**
  $\mathcal{G}^*[:, j] = \text{argmin}_{\mathcal{G}[:,j]} \ell(Z\mathcal{G}[:, j], W_{\text{est}}[:, j])$
end
return $\mathcal{G}^*$

4. Experiments and Analysis

Rubin Observatory: m=18,k=3.
4 paragraphs

4.1. Datasets

Donut Bank: comprised of 600,147 simulated 256×256 pixel Rubin Observatory donut images (see Figure 3). Each donut image has a corresponding position and true local wavefront label. This dataset is used to train the neural network to estimate the local wavefront.

The sources are chosen to be as realistic as possible. We started by drawing 5,000 r-band observations from a simulated Rubin Observatory observing schedule [6]. For each of these visits we queried the Gaia DR2 catalog for sources that would fall on the wavefront sensors [8]. Then we sampled 200 stars, with replacement, to simulate from each visit. We simulated an additional 100,147 blends - donut images with multiple stars overlapping - so that the network could learn to handle these complicated cases as well.

The simulations start by drawing photons from a blackbody distribution based on the star temperature and magnitude from the catalog. We then propagate these through the atmosphere with the help of the GalSim Python package [23]. We use frozen phase screens to represent low spatial frequency turbulence and apply a randomly drawn 'second kick' to account for high frequency turbulence. Then we use the Batoid Python raytracing package to generate differently perturbed Rubin optics instances and trace the photons into the detector [19]. Finally, we use the GalSim to model the sensor readout. We incorporate custom functions throughout this pipeline to account for additional physical effects such as: chromatic seeing, differential chromatic refraction, charge diffusion in the sensors, bad pixels, and astrometric errors.

The local wavefront labels are calculated with Batoid. For each perturbed telescope instance, a grid of rays are traced from the entrance pupil through the corresponding field position to the exit pupil. Then Zernike polynomials are fit to the optical path differences between the rays. These coefficients are the entries of the labels.

[stress that donuts have no relation to each other in transition]

Baseline Survey: consists of 499 Rubin visits, each containing hundreds to thousands of simulated donuts. This dataset is used for testing the full framework, which ingests

	Count		MSE	
Sources	Train	Test	Train	Test
Stars	498,071	1,708	4.5 ± 3.2	4.4 ± 3.5
Blends	100,028	340	9.5 ± 20.0	9.6 ± 22.0

Table 1. Train and test results on stars and blends. The mean-squared-error (MSE) is in units of thousandths of waves.

Sources	Loss	$\|\|\mathcal{G}^r\|\|_1 < \|\|\mathcal{G}^t\|\|_1$	$\|\|\mathcal{G}^r\|\|_1 / \|\|\mathcal{G}^t\|\|_1$
Stars and Blends	ℓ_1	99.6	0.482 ± 0.126
	ℓ_2	99.8	0.488 ± 0.124
	ℓ_{huber}	100.0	0.480 ± 0.120
	$\ell_{1,\text{median}}$	97.8	0.670 ± 0.138
	$\ell_{2,\text{median}}$	100.0	0.464 ± 0.121
Stars	ℓ_1	99.8	0.435 ± 0.111
	ℓ_2	100.0	0.430 ± 0.100
	ℓ_{huber}	100.0	0.430 ± 0.103
	$\ell_{1,\text{median}}$	97.2	0.635 ± 0.135
	$\ell_{2,\text{median}}$	99.8	0.414 ± 0.105
Brightest Stars	ℓ_1	99.6	0.369 ± 0.127
	ℓ_2	100.0	0.342 ± 0.120
	ℓ_{huber}	100.0	0.343 ± 0.122
	$\ell_{1,\text{median}}$	97.2	0.599 ± 0.156
	$\ell_{2,\text{median}}$	100.0	0.348 ± 0.121
Labels	ℓ_1	100.0	0.127 ± 0.053
	ℓ_2	100.0	0.060 ± 0.018
	ℓ_{huber}	100.0	0.079 ± 0.037

Table 2. Train and test results on stars and blends. The mean-squared-error (MSE) is in units of thousandths of waves.

all the donuts in a given visit and produces a global wavefront estimate.

All the donuts in a visit are simulated under the same conditions, namely atmospheric turbulence and sky background. The visits are drawn from Rubin scheduler simulations and the sources correspond to Gaia queries. Then the simulations proceed as described above. In addition to the positions and local wavefront labels for each donut, there is also a global wavefront for the entire visit. We use the batoid package to compute the relevant double Zernike polynomial basis from the telescope optics.

4.2. Architecture and Training

Question: is the network memorizing, or learning? Train-test similarity and qualitative test suggests it is actually learning.

	Center FWHM	Corner FWHM
Before	0.288 ± 0.034	0.314 ± 0.045
After	0.211 ± 0.005	0.215 ± 0.009

Table 3. Todo.

4.3. Local Wavefront Results

4.4. Global Wavefront Results

5. Conclusion

The exceptional simplicity and robustness of our approach make it well suited for space, where robustness and reliability are at a premium.

References

[1] P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, S Bailey, et al. Lsst science book, version 2.0. 2009. 1

[2] Gregory Allan, Iksung Kang, Ewan S. Douglas, George Barbastathis, and Kerri Cahoy. Deep residual learning for low-order wavefront sensing in high-contrast imaging systems. *Optics Express*, 28(18):26267, Aug. 2020. 2

[3] Torben Andersen, Mette Owner-Petersen, and Anita Enmark. Neural networks for image-based wavefront sensing for astronomy. *Optics Letters*, 44(18):4618, Sept. 2019. 2

[4] J. R. P. Angel, P. Wizinowich, M. Lloyd-Hart, and D. Sandler. Adaptive optics for array telescopes using neural-network techniques. *Nature*, 348(6298):221–224, Nov. 1990. 2

[5] Todd K. Barrett and David G. Sandler. Artificial neural network for the determination of Hubble Space Telescope aberration from stellar images. *Applied Optics*, 32(10):1720–1727, Apr. 1993. 2

[6] Francisco Delgado and Michael A. Reuter. The LSST Scheduler from design to construction. In *Observatory Operations: Strategies, Processes, and Systems VI*, volume 9910 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 991013, July 2016. 1, 3

[7] Nicholas Devaney, Fiona Kenny, Alexander V. Goncharov, Matthias Goy, and Claudia Reinlein. Development of a prototype active optics system for future space telescopes. *Applied Optics*, 57(22):E101, Aug. 2018. 1

[8] Gaia Collaboration, A. G. A. Brown, A. Vallenari, T. Prusti, et al. Gaia Data Release 2. Summary of the contents and survey properties. *Astronomy and Astrophysics*, 616:A1, Aug. 2018. 3

[9] Hong Guo, Nina Korablinova, Qiushi Ren, and Josef Bille. Wavefront reconstruction with artificial neural networks. *Optics Express*, 14(14):6456–6462, July 2006. 2

[10] Hongyang Guo, Yangjie Xu, Qing Li, Shengping Du, Dong He, Qiang Wang, and Yongmei Huang. Improved machine learning approach for wavefront sensing. *Sensors*, 19(16), 2019. 2

[11] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Com-*

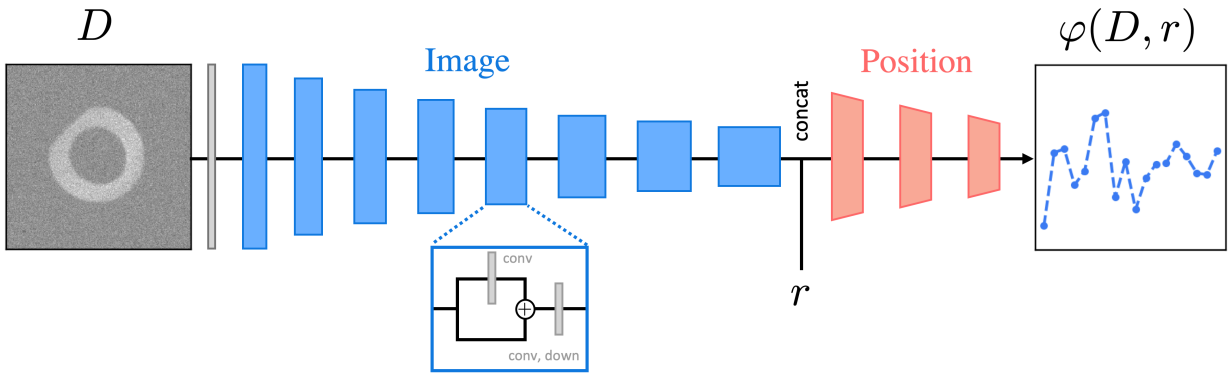

Figure 2. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

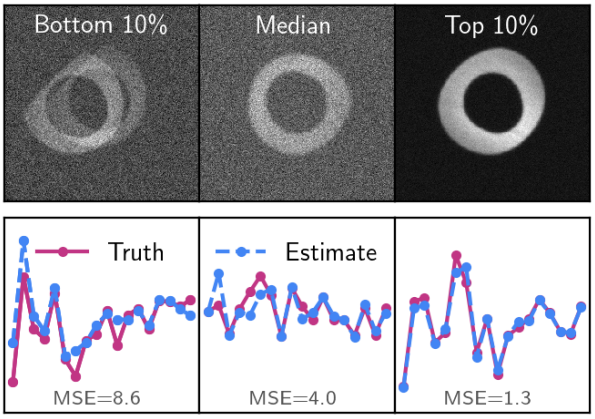

Figure 3. Foo.

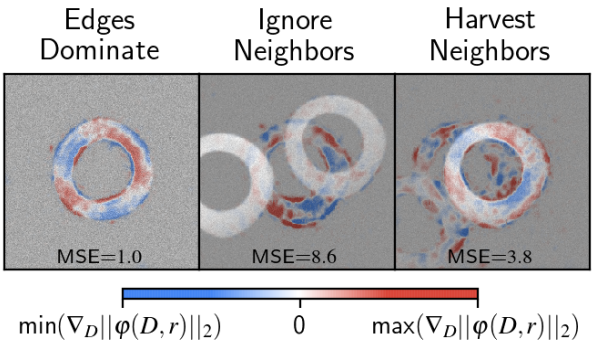

Figure 4. Foo.

puter Vision and Pattern Recognition (CVPR), pages 770–778, 2016. 2

[12] Ž. Ivezić, J. A. Tyson, E. Acosta, R. Allsman, S. F. Anderson, J. Andrew, J. R. P. Angel, T. S. Axelrod, J. D. Barr, A. C. Becker, et al. Lsst: from science drivers to reference design and anticipated data products. 2008. 1

[13] Mark B. Jorgenson and George J. Aitken. Neural network prediction of turbulence-induced wavefront degradations with applications to adaptive optics. In Firooz A. Sadjadi, editor, *Adaptive and Learning Systems*, volume 1706 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 113–121, Aug. 1992. 2

[14] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, page 10971105, Red Hook, NY, USA, 2012. Curran Associates Inc. 2

[15] Marie Laslandes, Claire Hourtoule, Emmanuel Hugot, Marc Ferrari, Christophe Devilliers, Arnaud Liotard, Céline Lopez, and Frédéric Chazallet. Last results of MADRAS, a space active optics demonstrator. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 10564 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 1056413, Nov. 2017. 1

[16] Marie Laslandes, Emmanuel Hugot, Marc Ferrari, Claire Hourtoule, Christian Singer, Christophe Devilliers, Céline Lopez, and Frédéric Chazallet. Mirror actively deformed and regulated for applications in space: design and performance. *Optical Engineering*, 52:091803, Sept. 2013. 1

[17] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. 2

[18] M. Lloyd-Hart and P. McGuire. Spatio-temporal prediction for adaptive optics wavefront reconstructors. In *European Southern Observatory Conference and Workshop Proceedings*, volume 54 of *European Southern Observatory Conference and Workshop Proceedings*, page 95, Jan. 1996. 2

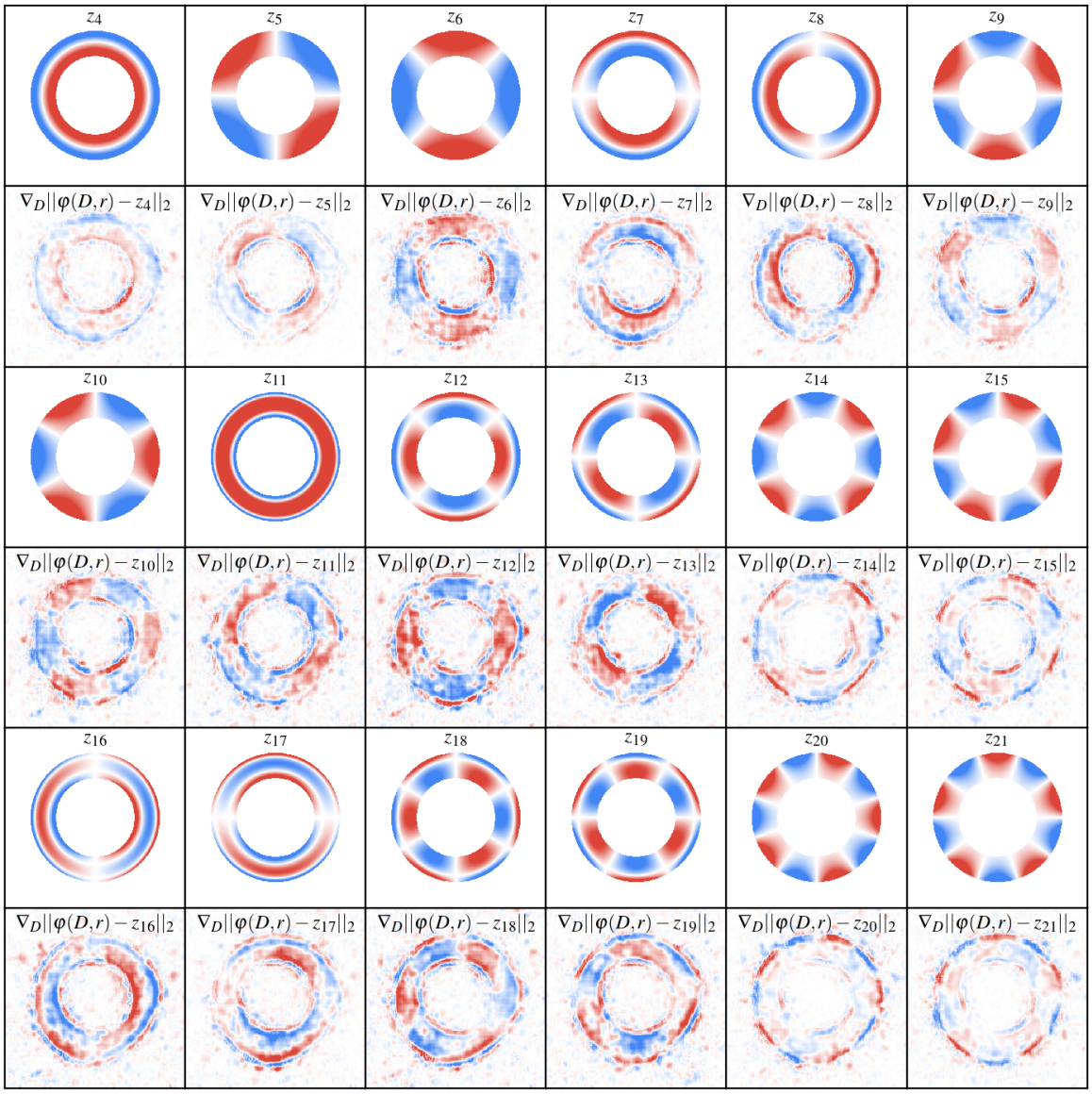

Figure 5. Foo.

[19] Joshua Meyers. batoid. https://github.com/jmeyers314/batoid, 2020. Accessed: 2020-04-07. 1, 3

[20] Dennis A. Montera, Byron M. Welsh, Michael C. Roggemann, and Dennis W. Ruck. Prediction of wave-front sensor slope measurements with artificial neural networks. *Applied Optics*, 36(3):675–681, Jan. 1997. 2

[21] Yohei Nishizaki, Matias Valdivia, Ryoichi Horisaki, Katsuhisa Kitaguchi, Mamoru Saito, Jun Tanida, and Esteban Vera. Deep learning wavefront sensing. *Optics Express*, 27(1):240, Jan. 2019. 2

[22] Scott W. Paine and James R. Fienup. Machine learning for improved image-based wavefront sensing. *Optics Letters*, 43(6):1235, Mar. 2018. 2

[23] B. T. P. Rowe, M. Jarvis, R. Mandelbaum, G. M. Bernstein, J. Bosch, M. Simet, J. E. Meyers, T. Kacprzak, R. Nakajima, J. Zuntz, H. Miyatake, J. P. Dietrich, R. Armstrong, P. Melchior, and M. S. S. Gill. GALSIM: The modular galaxy image simulation toolkit. *Astronomy and Computing*, 10:121–150, Apr. 2015. 1, 3

[24] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vision*, 115(3):211252, Dec. 2015. 2

[25] D. G. Sandler, T. K. Barrett, D. A. Palmer, R. Q. Fugate, and

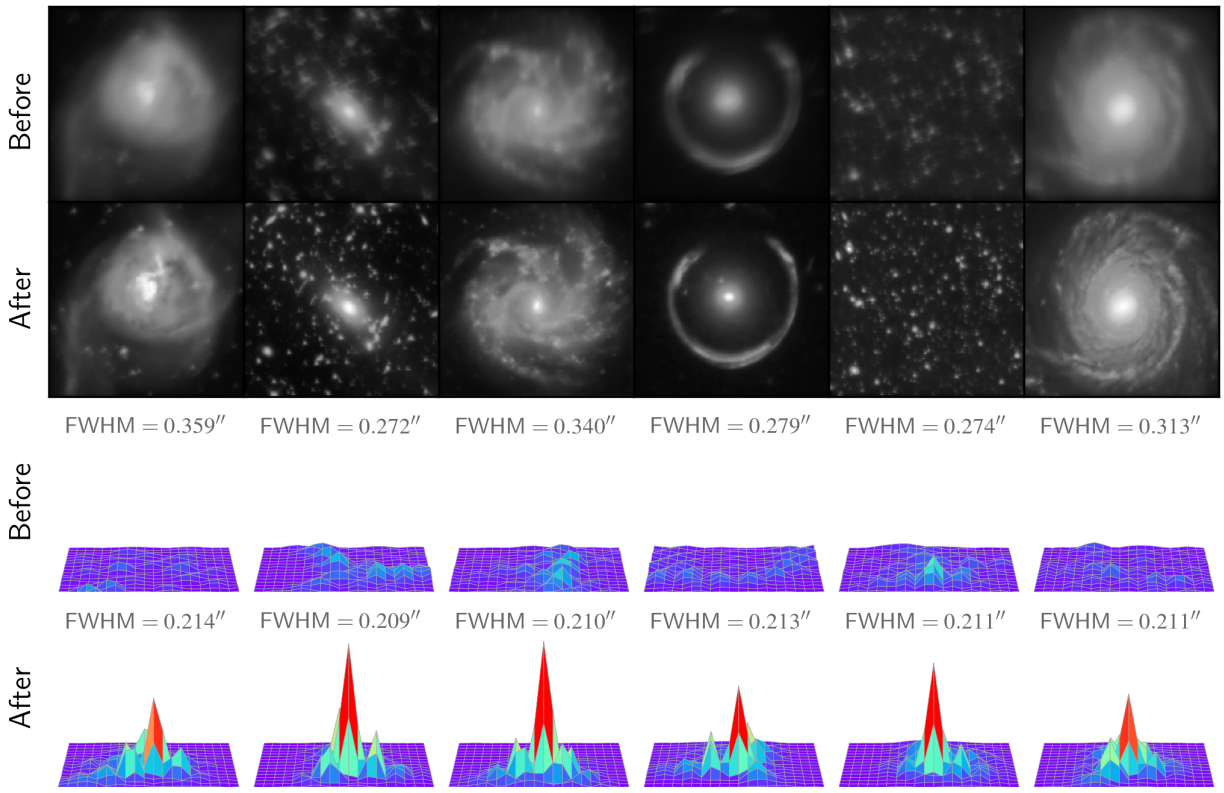

Figure 6. Example of a short caption, which should be centered.

W. J. Wild. Use of a neural network to control an adaptive optics system for an astronomical telescope. *Nature*, 351(6324):300–302, May 1991. 2

[26] Damien Sucher, Guillaume Butel, Guillaume Briche, Jean-François Blanc, Arnaud Liotard, Marc Bernot, Mikaël. Carlavan, Aurélien Suau, Nisrine Louh, Lauriane Galtier, Sebastien Guionie, Thierry Viard, Stéphanie Behar-Lafenetre, Fabrice Champandard, Jean-Bernard Ghibaudo, and Vincent Costes. Active optics for space telescopes. In *Astronomical Optics: Design, Manufacture, and Test of Space and Ground Systems II*, volume 11116 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 1111611, Sept. 2019. 1

[27] Christian Szegedy, Alexander Toshev, and Dumitru Erhan. Deep neural networks for object detection. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'13, page 25532561, Red Hook, NY, USA, 2013. Curran Associates Inc. 2

[28] David Thomas, Joshua Meyers, and Steven M. Kahn. Wide-field wavefront sensing with convolutional neural networks and ordinary least squares. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 11448 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 114484H, Dec. 2020. 2

[29] T. Viard, J. F. Blanc, C. Devilliers, F. Champandard, B. Bailly, F. Falzon, J. B. Ghibaudo, D. Sucher, G. Briche, V. Costes, and C. Du Jeu. Active optics for next generation of space observation instruments. In *International Conference on Space Optics — ICSO 2018*, volume 11180 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 1118008, July 2019. 1

[30] S. J. Weddell and R. Y. Webb. Reservoir Computing for Prediction of the Spatially-Variant Point Spread Function. *IEEE Journal of Selected Topics in Signal Processing*, 2(5):624–634, Nov. 2008. 2

[31] Pingwei Zhou, Dongxu Zhang, Guang Liu, and Changxiang Yan. Development of space active optics for a whiffletree supported mirror. *Applied Optics*, 58(21):5740, July 2019. 1