

# CNN to Mitigate Atmospheric Turbulence Effect on Shack-Hartmann Wavefront Sensing: A Case Study on The Magdalena Ridge Observatory Interferometer

Siavash Norouzi<sup>1,2</sup>, James J. D. Luis<sup>3</sup>, Ramyaa Ramyaa<sup>1</sup>, John S. Young<sup>3</sup>, Eugene B. Seneta<sup>3</sup>, Morteza Darvish Morshedi Hosseini<sup>4</sup>, E. Robert Ligon<sup>5</sup>

<sup>1</sup>Department of Computer Science, New Mexico Institute of Mining and Technology; Socorro, NM USA

<sup>2</sup>Magdalena Ridge Observatory, New Mexico Institute of Mining and Technology; Socorro, NM USA

<sup>3</sup>Cavendish Laboratory, University of Cambridge; Cambridge, UK

<sup>4</sup>Department of Electrical and Computer Engineering, Binghamton University; Binghamton, NY USA

<sup>5</sup>CHARA Array; Mt. Wilson, CA USA

## Abstract

The Magdalena Ridge Observatory Interferometer (MROI) utilizes Shack-Hartmann Wavefront Sensing (SH-WFS) for the back-end stability of its beam relay systems in a unique design. The SH-WFS, however, is sensitive to atmospheric turbulence scintillation which can drastically affect its precision in calculating the position of the beam profile it sees. A large number of images are needed to counteract the turbulence effect. Here we use deep learning as an alternative to long averaging cycles. A CNN was trained to map from a number of initial images of a series of star frames to the average image of the entire series at different positions of the beam profile. Under typical seeing conditions expected at MROI, the results showed that the network can map 10 input frames to the average of 100 within the permissible error margin of 0.1 pixels and furnish proper generalization to beam position movements not seen during training. The network can also perform better than the averaging technique when both techniques operate on small numbers of input frames such as 10 or 20.

## Introduction

Stellar interferometers are facilities that consist of multiple telescopes. They recover the angular resolution that a large single dish can achieve, but without the engineering complexity and cost. The Magdalena Ridge Observatory Interferometer (MROI) will have 10 telescopes when completed. Light will be collected by these telescopes and relayed over hundreds of meters in vacuum pipes which are part of the Beam Relay System (BRS). Beam angle in the BRS for any of the 10 beam lines cannot drift by more than 1  $\mu$ rad at any point during the night. Figure 1 shows a schematic of a single beam relay line that conveys light from a telescope to beam combiners for science observations. As illustrated in the figure, starlight exits the telescope in a parallel beam and is reflected by relay mirrors M4 and M5 towards a delay line that imparts a variable optical path up to 380 m. Subsequently, the beam is reduced in diameter by a beam compressor before being directed towards a set of beam combiners, which interfere together light from all the 10 telescopes for science measurements. The Fast Tip-Tilt (FTT) system is integrated to perform closed-loop correction of atmospheric tip-tilt (angle of arrival) fluctua-

tions on the light collected by the telescope [1].

The Automated Alignment System (AAS) is intended to counteract any optomechanical drifts in the beam lines with minimal human intervention [2]. The term alignment in the BRS is fully defined by two parameters: angle (tilt) and position (shear). The principal source of optomechanical drifts are the two flat mirrors M4 and M5 in (as depicted in Figure 1) that direct each beam from the telescope to a laboratory because they are exposed to outdoor temperature swings. The alignment of these two mirrors is queried by the Back-End Alignment System for Shear and Tilt (BEASST) which features a Shack-Hartmann Wavefront Sensor (SH-WFS). A SH-WFS consists of an array of microlenses (lenslets) with the same focal length which generate focused spots ultimately landing on a 2D camera. The beam profile underfills the sensor. This allows us to measure beam position because some lenslets are illuminated while others are unilluminated. Our algorithm reduces the image of the grid of focal spots to an image of the beam profile where each pixel is the summed brightness of a microlens [3]. The image is first segmented based on the location and size of the microlenses, then all pixels within each segment are summed. The reduced segments are restitched into a 2D image and the centroid is computed as an estimator for beam position. The offset of the beam from the optimal position is known

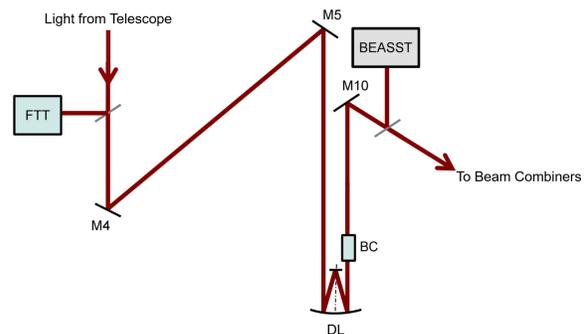
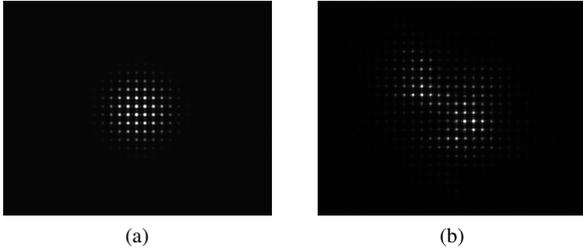


Figure 1. Schematic of a single beam train of Magdalena Ridge Observatory Interferometer



**Figure 2.** Microlens array illuminated under a) undisturbed and b) atmospherically disturbed light sources

as beam shear. As another output of this process, the software calculates the beam tilt, but it is outside the scope of this paper.

This method of position finding is, however, rather sensitive to intensity modulation across the beam profile [4]. Thus, the calculated shear is prone to error due to atmospheric turbulence scintillation. Figure 2 shows the microlenses illuminated in absence (a) and presence (b) of atmospheric turbulence. It should be noted that while the profile shown in Figure 2(a) remains consistent over time, the profile in Figure 2(b) changes randomly from frame to frame due to turbulence. Due to the atmospherically disturbed modulation of light intensity in a starlight frame, the BEASST software would take tens of seconds to average out the atmospheric perturbations to acquire the precision within the shear error margin of  $\pm 0.1$  pixels (microlenses). Beam misalignments will be routinely counteracted every 10 min or so with the aid of a stable reference light source in place of atmospherically-disturbed starlight. However, the shear of this reference beam must be calibrated against the shear of the stellar beam once per hour. This procedure is only allocated two min of operational time, including re-pointing of the telescope from a science target to a bright star (20 sec) and image acquisition (60 sec to 100 sec). The aim of this research work is to utilize deep learning in a way that the BEASST software would require less frames for shear calculation but have comparable output to the conventional averaging technique. As a result, the amount of time needed for measuring shear on starlight would be reduced. Figure 3 illustrates our intended outcome in the pipeline which replaces the averaging technique with CNN.

### Related Works

Deep learning techniques have offered promising solutions to many image processing problems where accuracy and efficiency are essential. Several deep convolutional network-based solutions have been proposed in the literature for similar problems involving light aberrations in SH-WFS applications. Suárez Gómez *et al.* [5] used a deep convolutional network and successfully predicted the centroids on a SH-WFS for wavefront reconstruction. Deep convolutional networks have shown the ability to effectively model optical processes irrespective of the complexity of the underlying physics [6, 7, 8, 9, 10, 11]. U-Net, which is primarily used for image segmentation [12], was extended by Swanson *et al.* [13] for SH-WFS wavefront reconstruction. In their modified architecture, the X and Y slopes of the microlenses of SH-WFS were input separately into the network. DuBose *et al.* [11] extended this work by adding a third encoder arm that uses the total intensity of each subaperture. Zhang *et al.* [14] developed a feed-forward Denoising Convolu-

tional Neural Network (DnCNN) which can be trained to perform Gaussian denoising with unknown noise level (i.e., blind Gaussian denoising). One single DnCNN exhibits high effectiveness in general denoising tasks such as Gaussian denoising, single image super-resolution and JPEG image deblocking and it can be efficiently implemented on GPU. Li *et al.* [15] combined DnCNN with the Phase Diversity (PD) algorithm to improve wavefront sensing accuracy. They report significant accuracy enhancement, their composite PD algorithm with DnCNN showed up to 82.35% better accuracy compared with the traditional PD algorithm.

As mentioned earlier, the MROI uses SH-WFS in a unique fashion to maintain the stability of its beam relay system. Subsequently, our purpose of measuring shear is different from any other work reported in the literature. However, the works mentioned earlier enhanced our incentive by indicating that SH-WFS data can be learned by deep neural networks despite the presence of light intensity modulation. Our objective is to acquire a denoised image from a series of images disturbed by atmospheric noise, which is modeled as Additive Gaussian Noise (AGN). We used the DnCNN which is a denoising network designed for AGN. The simulation results presented in this paper confirm the effectiveness of this proposed approach.

### Proposed Approach

As illustrated in Figure 4, the DnCNN consists of a convolution layer with a Rectified Linear Unit (ReLU), a variable number  $d$  of convolution layers each with batch normalization and ReLU, and a single convolution layer. The problem to solve was defined as a translation from a number of initial frames of a batch of starlight frames into a single image that contains the average of all of the frames in that batch. The initial frames, whose quantity  $n$  was to be determined by the network's performance, were used as the input and the average frame as the output or the ground truth in the training of the network. Each of the initial input frames and also the average frame were introduced as tensors to the network. The input consisted of a tensor of  $n \times 32 \times 32$  dimensions, with  $n$  indicating the number of input frames. When fed to the network, the channels increased to 64 and dimensions reduced to  $16 \times 16$ , both remaining constant throughout the entire network. The network's output was a tensor with one channel (i.e. grayscale), as in camera output, with dimensions upscaled to  $32 \times 32$ . A kernel size of 3, padding of 1 and batch size of 32 were used in training of the network. The optimizer was Adam [16] with the learning rate of  $2 \times 10^{-4}$ . The mean square error was used as the loss function. In terms of the network's depth  $d$ , two values 17 and 20 have been mentioned in the work by Zhang *et al.* [14]. We will show later that these two depth values cannot furnish proper generalization and instead the depth of five layers is optimal for our data. Separate comprehensive analyses of our findings are presented in the subsequent sections.

### Train Dataset

In order to generate training, a simulation code was used to model atmospheric turbulence on the frames seen by the SH-WFS. The training dataset consisted of extreme shear scenarios in X and Y directions as well as zero shear (i.e. centroid of the pupil being at the frame's center). To accommodate all possible shear values occurring in the BRS, the simulation was set up to generate pupil frames sheared at  $-60 \mu\text{m}$  and  $60 \mu\text{m}$  in X and Y

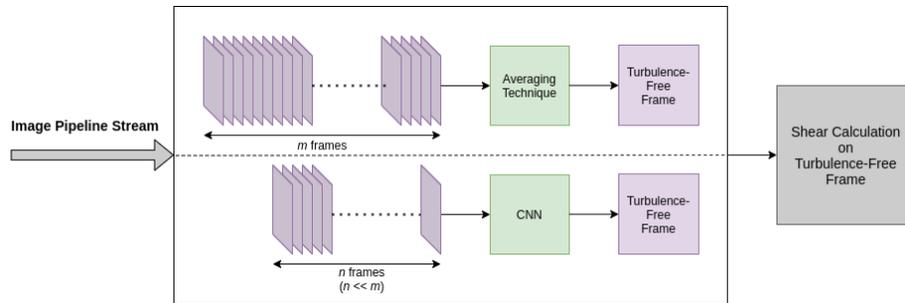


Figure 3. Schema of the placement of CNN in our pipeline to replace averaging technique

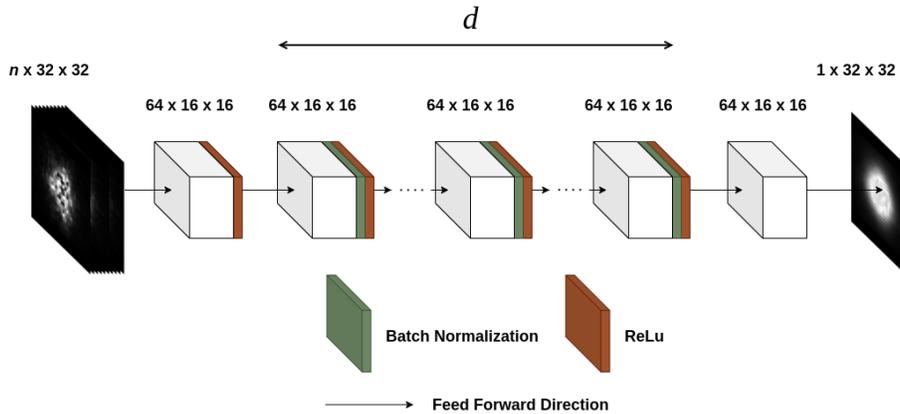


Figure 4. DnCNN architecture with depth  $d$  and  $n$  input frames

directions as well as zero shear. All the shear values in between the extremes were left up to the neural network to interpolate after training. The simulation modeled each iteration of sheared starlight at 16 sec of exposure time with 0.04 sec of subexposure, yielding 400 frames each containing a pupil with constant shear and random turbulence effect. The turbulence was modeled using the Python module MegaScreen developed by D. Buscher [17] with  $1.22 \mu\text{m}$  wavelength, the Fried parameter  $r_0 = 30 \text{ cm}$  and wind speed of 10 m/s. The ground truth for each iteration of shear was represented by the average of each batch of 400 frames. This iteration was repeated 1000 times for each of the aforementioned shear scenarios.

### Test Dataset

The testing dataset was simulated similarly but in order to save processing time, generate more data and be able to evaluate the networks on data with unseen exposure characteristics, the simulation modeled four sec of exposure time with 0.04 sec of subexposure, which yielded 100 frames. Each batch of 100 frames still had pupils with constant shears, but this time the simulation generated 25 batches to model shears of  $-60 \mu\text{m}$  to  $60 \mu\text{m}$  at  $30 \mu\text{m}$  intervals in X and Y directions. Each of the 25 batches was repeated 100 times to accommodate as much randomness as possible. Figure 5 shows a schematic of all the positions where the red circles denote the positions (or shear scenarios) the network saw during training. Figure 6 shows the shear X and shear Y values associated with each position number. The position number corresponds to increments in shear values in Shear X and Shear Y plots. The simulation incremented shear X and Y in sequential

steps, it increased shear Y values in five steps while keeping shear X constant and then incremented shear Y by one step, repeating the entire process for 25 steps. For each of the 25 position numbers, 100 batches (datapoints) each containing 100 frames with a constant shear value were generated. Each blue dot on the plots represents the shear value of an image that is equivalent to the average of all of the 10,000 frames corresponding to a position number. The Centroid (px) values are the distance between the centroid of the pupil and the image center in X and Y directions in pixels. It is observed that the range  $-60 \mu\text{m}$  to  $60 \mu\text{m}$  translates to approximately  $-0.2$  to  $0.2$  pixels in X and Y directions. It should be noted that the imperfect flatness occurring in Shear X, which might be due to unavoidable noise in centroiding, yields an error of at most 0.01 pixels which is an order of magnitude less than our goal of 0.1 pixels.

### Image Reduction Technique

Each of the simulated starlight frames was  $512 \times 640$  in shape which is identical to the output of the BEASST camera. The image processing pipeline overlays a square grid with 20 pixel spacing to represent the microlens array. The grid is aligned such that the spots are positioned close to the center of each box. The image is then segmented into the constituent boxes of the grid. For each box, the algorithm records the (X,Y) coordinates and the enclosed intensity (sum of all enclosed pixels). A new image is formed from the box coordinates and intensities, then saved to disk with dimensions of  $32 \times 24$ . Compared with the full  $640 \times 512$  images, these reduced frames saved memory (8 GB total GPU memory), allowed large training sets and accelerated the

training and testing processes. Prior to feeding the image to the neural network, each frame was placed at the center of a  $32 \times 32$  black canvas so that it became square in shape, conforming to a common practice in deep learning with images.

### Evaluation of the Network

As the optimal number of initial frames  $n$  was unknown, it was determined based on the network's evaluation performance. As the output of the network is a single image representing the average of all starlight frames in a batch, the performance of the network was evaluated based on the distance between the centroid of the actual average frame and that of the network's output in X and Y directions. The required measurement accuracy of BEASST is 0.5% of the pupil diameter. The pupil diameter in an input star frame in its compressed form measures approximately 20 pixels. Thus, the maximum error translates to  $\pm 0.1$  pixels for the entire pupil.

### Experimental Results

The DnCNN with five layers was trained in five different configurations where the network architectures were the same but the number of input frames was varied as  $n = 10, 20, 40, 60$  and  $80$ . Each configuration was then tested on the 25 position numbers aforementioned. In order to mitigate the effect of noise on centroid calculation, the output image of each network was subject to a thresholding technique where 30% of the maximum pixel value of the image was subtracted from all of the pixels and any resulting negative pixel was set to zero.

#### Performance on the Average of 100 frames

Figure 7 shows the performance of the network configurations in form of the difference between the average of the centroids of 100 network output frames at each position number and that of the average image of 10,000 frames corresponding to that position number. Given the maximum acceptable error of 0.1 pixels, the DnCNN architecture exhibits evidently good performance with almost every configuration tested, with only one exception occurring with  $n = 10$  at position number 21 in shear Y where the error is approximately 0.13 pixels. Another case is with  $n = 20$  at position number 24 in shear Y where the error is roughly 0.105 pixels which is hardly significant considering the overall performance of the configuration. It can be concluded that the DnCNN architecture with five layers can generate the average

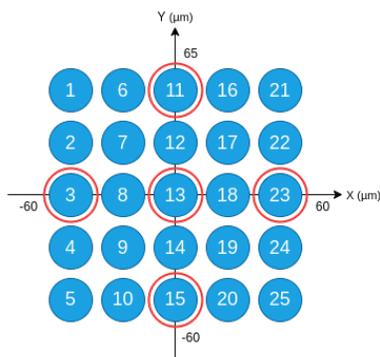


Figure 5. Position numbers used in training and testing

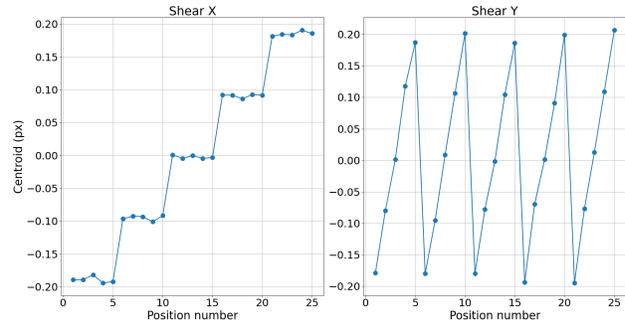


Figure 6. Position numbers and shear values

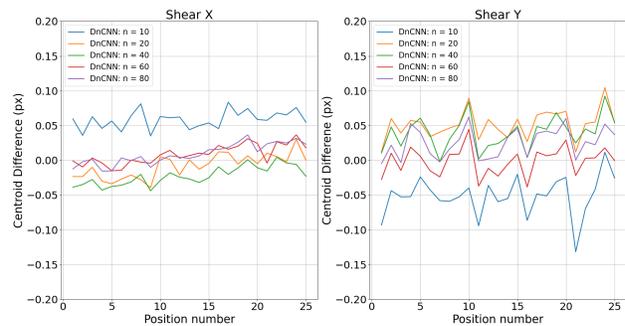


Figure 7. DnCNN Performance with  $d = 5$  layers and varying number of input frames ( $n$ )

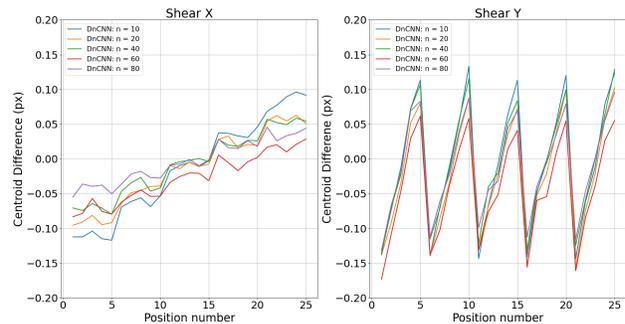
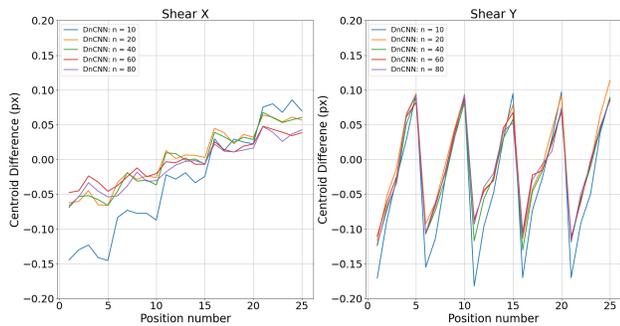


Figure 8. DnCNN Performance with  $d = 17$  layers and varying number of input frames ( $n$ )

of 100 frames in one data point with at least any number of input frames between 10 and 20 and its performance improves with higher  $n$  values as expected. As mentioned earlier, the DnCNN architectures with 17 and 20 layers as proposed by Zhang *et al.* were trained and tested with our data but did not provide generalization accuracy within our error margin. Figures 8 and 9 show the performance of these two architectures when trained and tested with the same varying  $n$  values. They both show an interesting trend of difference between the network and target output which increases in magnitude at extreme shear values both in Shear X and Shear Y plots. The trend is in form of a systematic error in both shear X and shear Y.

#### Performance on the Average of $n$ frames

Although the primary objective of utilizing DnCNN was to map from  $n$  number of frames to 100, it was decided to com-



**Figure 9.** DnCNN Performance with  $d = 20$  layers and varying number of input frames ( $n$ )

pare its performance with the average of  $n$  frames. A different series of analysis was carried out to evaluate the performance of the proposed approach with  $n$  number of input frames against the average of the same  $n$  frames. The performance was evaluated by comparing the distributions generated by the network and the averaging technique. Box plots were used to represent the distributions of the average images and the network outputs. The box plots in each position numbers were to show how each distribution is more concentrated around its median at different  $n$  values. The median for each distribution was adjusted according to the median of 100 frames such that both distributions will have medians roughly around zero. As shown in Figure 10, this analysis revealed that the network can produce more concentrated distributions than the averages only for low  $n$  values of 10 and 20 and it is gradually outperformed by the averaging technique starting from  $n = 40$ . This observation holds true for any position number regardless of whether seen or not seen by the network during training. Therefore, for the sake of brevity in this paper, the box plots of only the first position number are presented as this trend is very similar with any position number whether seen or not seen by the network in training. It should be noted the error in Y direction is larger than in X, which can be due to the fact that the simulation blows frozen turbulence in X direction.

## Conclusions

The DnCNN network was trained on simulated data for the problem of light intensity modulation facing the SH-WFS hardware serving a critical back-end stability for the light relay systems of the MROI. The results showed that the network learned the shear in turbulent-disturbed star frames and with 10 input frames, its performance was within the error margin of 0.1 pixels and improved with more input frames. The training was performed on only five position numbers but the network provided reasonable results for 25 evaluation positions. The network also showed to perform better than the image averaging technique for  $n = 10$  and  $n = 20$ . Our sensitivity analysis on the the network's depth indicated that the network architecture with the depth of five layers yields optimal results for our data.

This work used only one condition of the atmosphere (a single combination of  $r_0$ , wind speed and wind direction). In the near term, the network should be trained with a range of simulated atmospheric conditions. In the long term, the ultimate test will be to gather a library of real SH-WFS images under a range of conditions.

## Acknowledgments

We would like to thank Dr. Michelle J. Creech-Eakman and Dr. Van D. Romero at New Mexico Tech for providing financial support for this research. We would also like to extend our gratitude to Mr. Hamed Momeni at New Mexico Tech for sharing his invaluable expertise on the subject matter. Moreover, we acknowledge our system administrators Aedan Wells and Sylvain Jones at the department of computer science for their support on the GPU server.

## References

- [1] J. Young, D. Buscher, M. Fisher, C. Haniff, A. Rea, E. B. Seneta, X. Sun, D. Wilson, A. Farris, A. Olivares, *et al.*, "The mroi fast tip-tilt correction and target acquisition system," in *Optical and Infrared Interferometry III*, vol. 8445, p. 84451V, International Society for Optics and Photonics, 2012.
- [2] J. J. Luis, R. Blasi, D. F. Buscher, A. Farris, R. Kelly, and R. Ligon, "Augmented design for a fully automated alignment system at the magdalena ridge observatory interferometer," in *Optical and Infrared Interferometry and Imaging VI*, vol. 10701, p. 107010M, International Society for Optics and Photonics, 2018.
- [3] J. J. Luis, D. F. Buscher, C. Dahl, J. F. Dooley, A. Farris, C. A. Haniff, E. R. Ligon, S. Norouzi, X. Sun, and J. S. Young, "Deployment of beam alignment hardware at the magdalena ridge observatory interferometer," in *Optical and Infrared Interferometry and Imaging VII*, vol. 11446, p. 1144628, International Society for Optics and Photonics, 2020.
- [4] V. Lakshminarayanan and A. Fleck, "Zernike polynomials: a guide," *Journal of Modern Optics*, vol. 58, no. 7, pp. 545–561, 2011.
- [5] S. L. S. Gómez, C. González-Gutiérrez, E. D. Alonso, J. D. S. Rodríguez, M. L. S. Rodríguez, J. C. Landeira, A. Basden, and J. Osborn, "Improving adaptive optics reconstructions with a deep learning approach," in *International Conference on Hybrid Artificial Intelligence Systems*, pp. 74–83, Springer, 2018.
- [6] S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," *Optica*, vol. 5, no. 7, pp. 803–813, 2018.
- [7] A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica*, vol. 4, no. 9, pp. 1117–1125, 2017.
- [8] Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light: Science & Applications*, vol. 7, no. 2, pp. 17141–17141, 2018.
- [9] S. W. Paine and J. R. Fienup, "Machine learning for improved image-based wavefront sensing," *Optics letters*, vol. 43, no. 6, pp. 1235–1238, 2018.
- [10] Y. Wu, Y. Luo, G. Chaudhari, Y. Rivenson, A. Calis, K. De Haan, and A. Ozcan, "Bright-field holography: cross-modality deep learning enables snapshot 3d imaging with bright-field contrast using a single hologram," *Light: Science & Applications*, vol. 8, no. 1, pp. 1–7, 2019.
- [11] T. B. DuBose, D. F. Gardner, and A. T. Watnik, "Intensity-

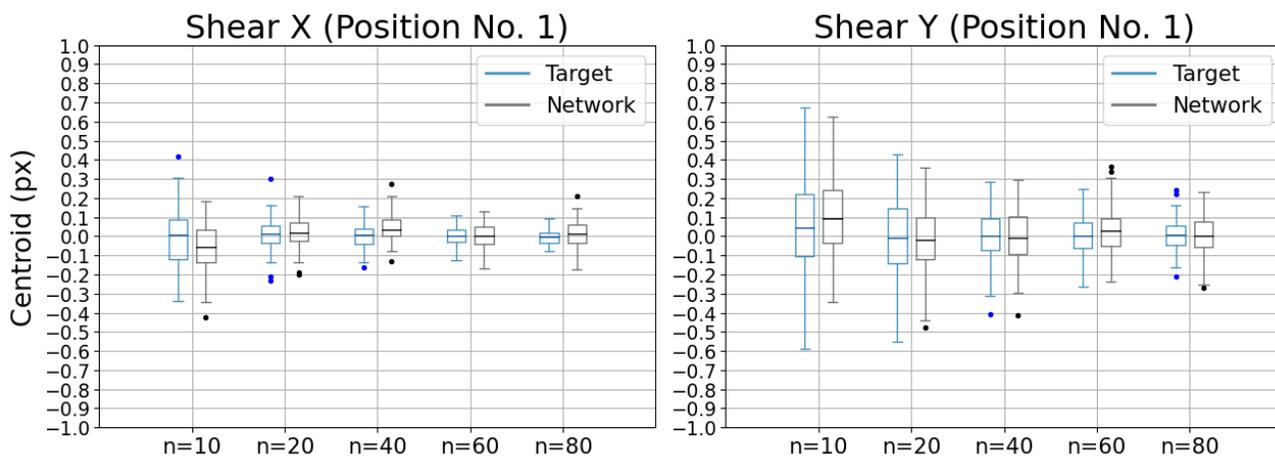


Figure 10. Box plots comparing performance of the proposed approach and averaging technique when both operate on  $n$  frames

enhanced deep network wavefront reconstruction in shack-hartmann sensors,” *Optics letters*, vol. 45, no. 7, pp. 1699–1702, 2020.

- [12] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.
- [13] R. Swanson, M. Lamb, C. Correia, S. Sivanandam, and K. Kutulakos, “Wavefront reconstruction and prediction with convolutional neural networks,” in *Adaptive Optics Systems VI*, vol. 10703, p. 107031F, International Society for Optics and Photonics, 2018.
- [14] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE transactions on image processing*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [15] D. Li, S. Xu, D. Wang, and D. Yan, “Phase diversity algorithm with high noise robust based on deep denoising convolutional neural network,” *Optics express*, vol. 27, no. 16, pp. 22846–22854, 2019.
- [16] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [17] D. F. Buscher, “Simulating large atmospheric phase screens using a woofer-tweeter algorithm,” *Optics express*, vol. 24, no. 20, pp. 23566–23571, 2016.

## Author Biography

Siavash Norouzi is a Master’s student of Computer Science at New Mexico Tech. He worked as a software engineer on the design and development of control software for the distributed systems of the MROI’s Automated Alignment System. His Master’s project revolved around the deep learning module of the BEASST software that he developed for the MROI. He holds another Master’s degree from New Mexico Tech with specialization in computational mechanics.

James Luis graduated from the University of St Andrews in 2013 with a MPhys degree in Physics. After a brief stint develop-

ing optical diagnostics for plasma physics experiments, he is now in the final stages of obtaining a PhD in Physics at the University of Cambridge, specialising in optical instrumentation for astronomy. His research tackles the problem of beam misalignment in long baseline optical interferometers such as the Magdalena Ridge Observatory Interferometer.

Ramyaa is an assistant professor of Computer Science at New Mexico Tech. She did her PhD at Indiana University, USA, Post doctoral research at Ludwig Maximilian University, Munich and has masters degrees from Carnegie Mellon University, USA and University of Georgia, USA. She was a fellow at Simons institute of Theory, UC Berkeley, USA. Her primary fields of research are Theory of Computation and Logic. Her secondary research area is Artificial Intelligence, focusing on machine learning.

John Young received a BA in Natural Sciences (1995) and a PhD in Astronomy (1999) from the University of Cambridge. He has subsequently worked on a range of projects in astronomical optical interferometry and novel instrumentation. He is an author on 22 refereed publications and over 100 conference papers.

Dr. Seneta received his BSc and MSc in Physics from Sydney University in 1988 and 1991 respectively. He was awarded a PhD in Biomedical Sciences from the University of Technology, Sydney in 1997. He then worked as an instrumentation scientist in Sydney University’s Astronomy Department until 2002, when he moved to the Cavendish Laboratory in the United Kingdom. There, he continues to develop instrumentation for optical interferometry, exoplanet detection and nanometer-accuracy etching.

Morteza Darvish Morshedi Hosseini is a PhD candidate in Electrical and Computer Engineering at Binghamton University. Morteza received his M.S. in Information Technology Engineering with a concentration on Information Security from University of Isfahan in 2016. His research areas of interest are centered around digital image processing.

Robert Ligon is currently a research scientist at the CHARA array on Mount Wilson. His interests include the development of optical instruments for astronomy.