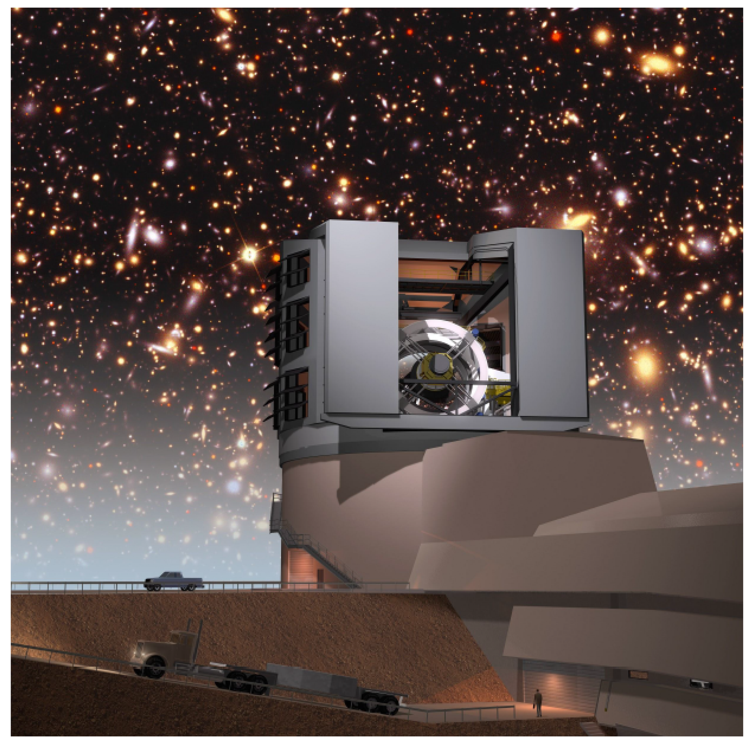
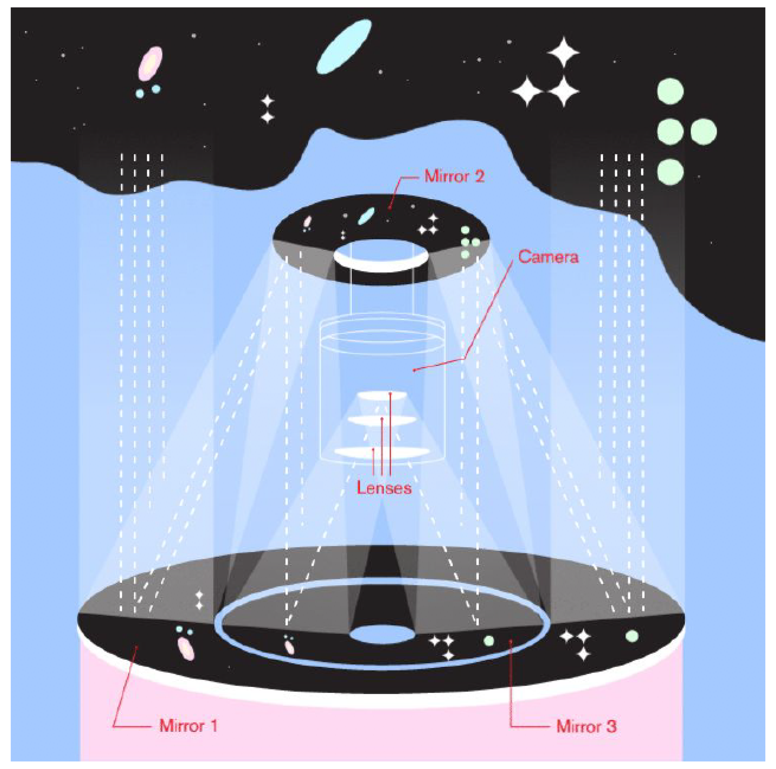
Active Optical Control with Machine Learning: A Proof of Concept for the Vera C. Rubin Observatory

Jun E. Yin, Daniel J. Eisenstein, Douglas P. Finkbeiner, Christopher W. Stubbs, Yue Wang

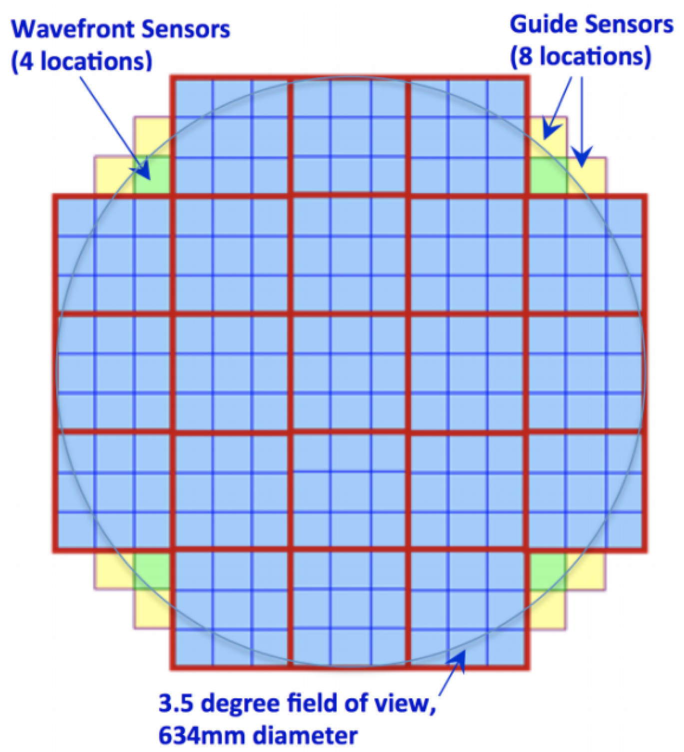
Vera C. Rubin Observatory (Rubin)

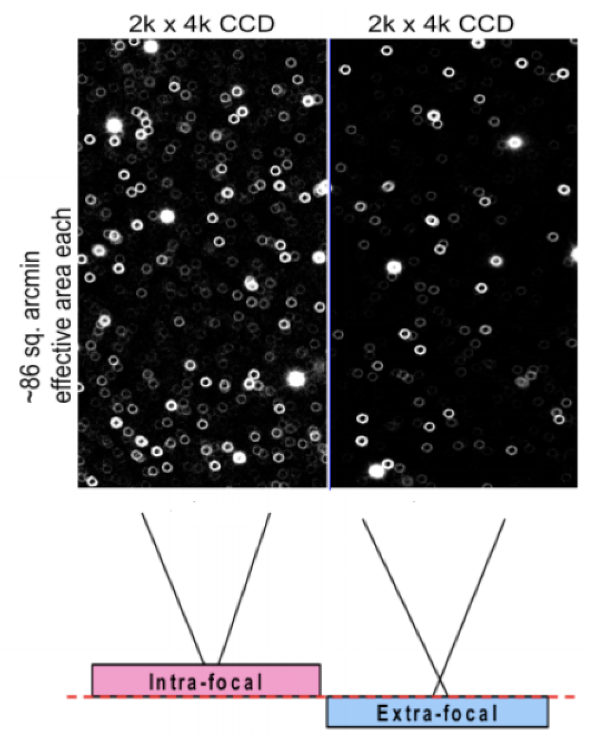




Optical design of Rubin

Active Optics Systems (AOS)

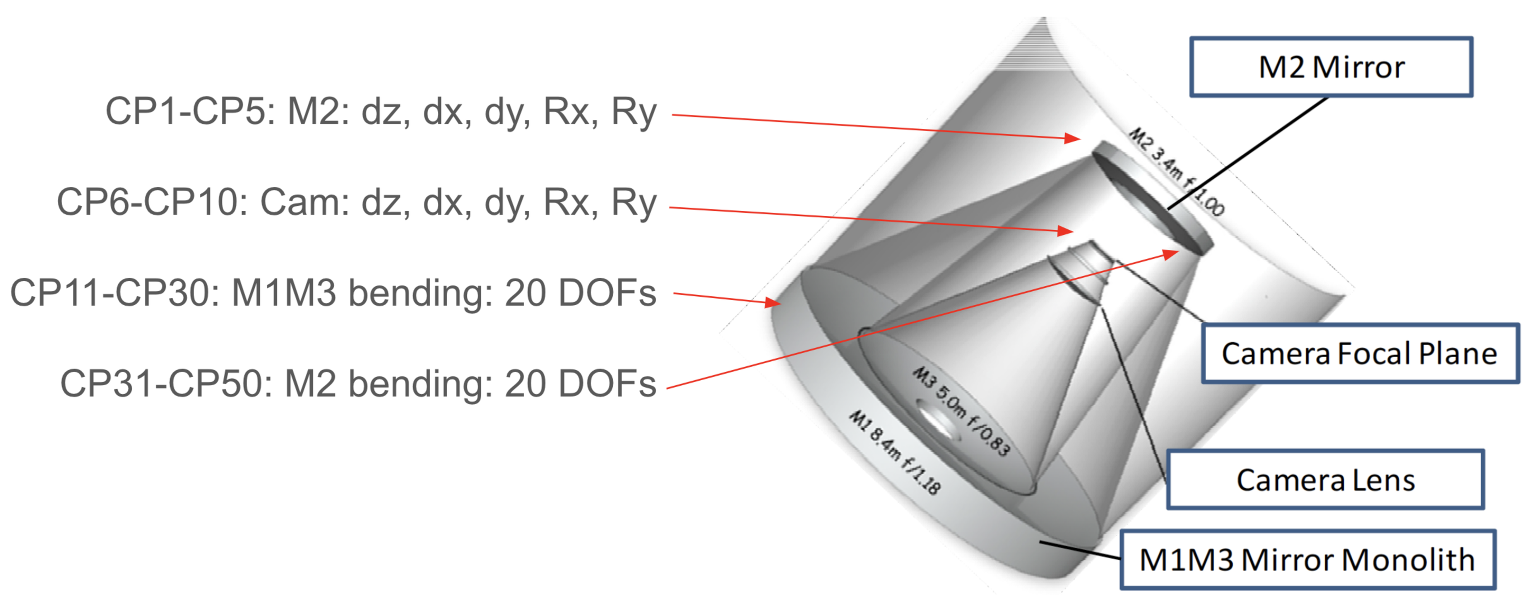




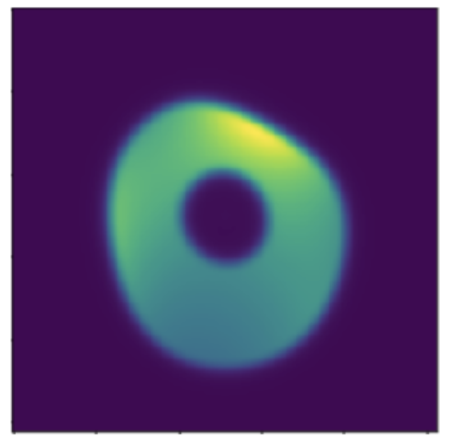
Reprinted and adapted with permission from Xin et al. (2015) The Optical Society.

The Vera C. Rubin Observatory focal plane and the schematic operation of the split wavefront sensors

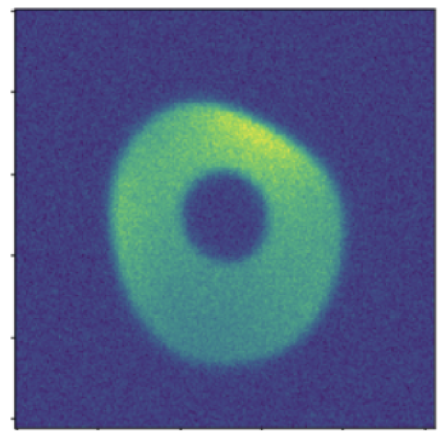
Control Parameters and Wavefront Generation



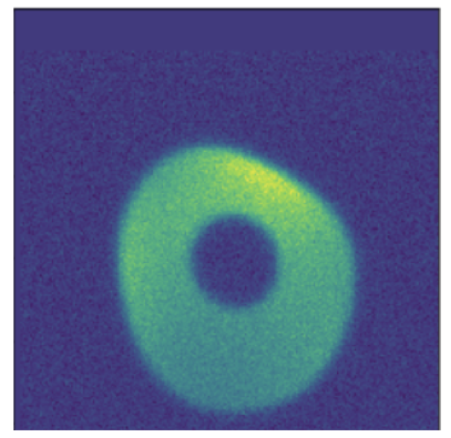
Donut generation



No noise and no shift



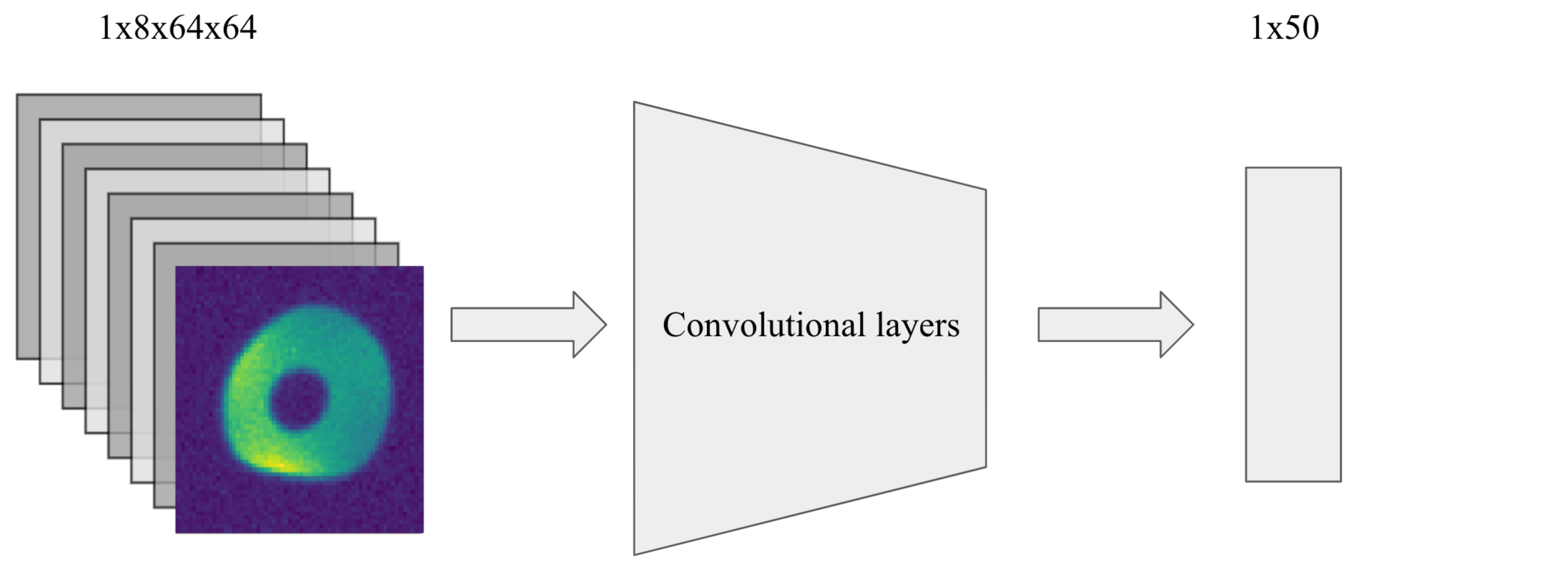
With noise but no shift



With noise and shift

Donut images were produced using the *makedonut* code provided by A. Roodman.

Algorithm



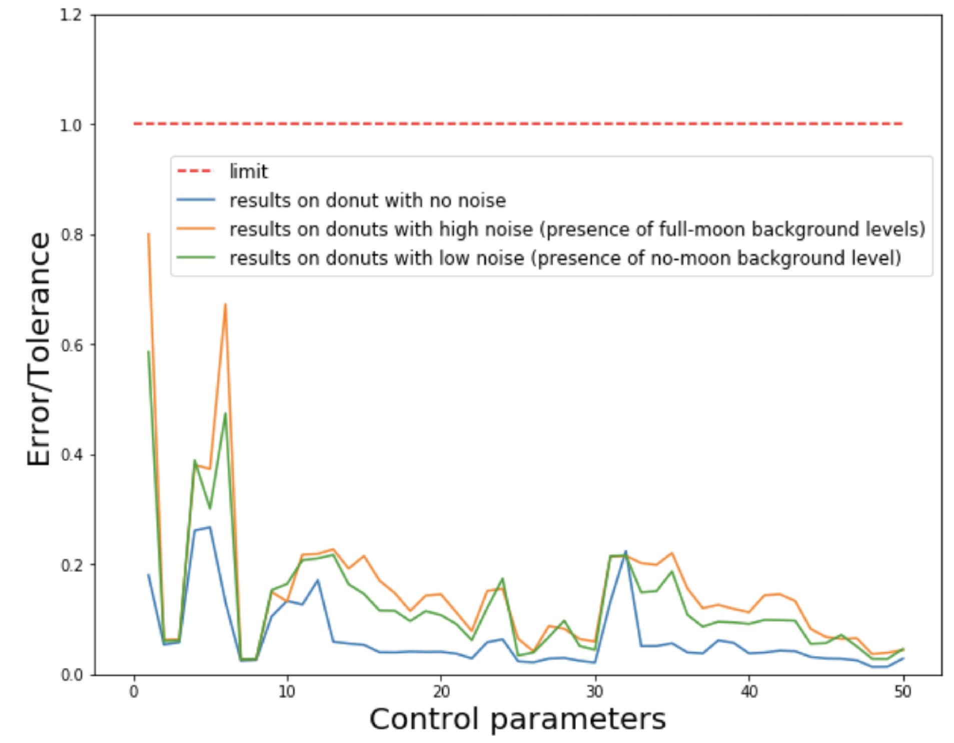
Loss function

$$L(\mathbf{y}, \mathbf{y}^*) = \sum_j \alpha_j L_2(y_j, y_j^*) + \beta f(\mathbf{y}^* - \mathbf{y})$$

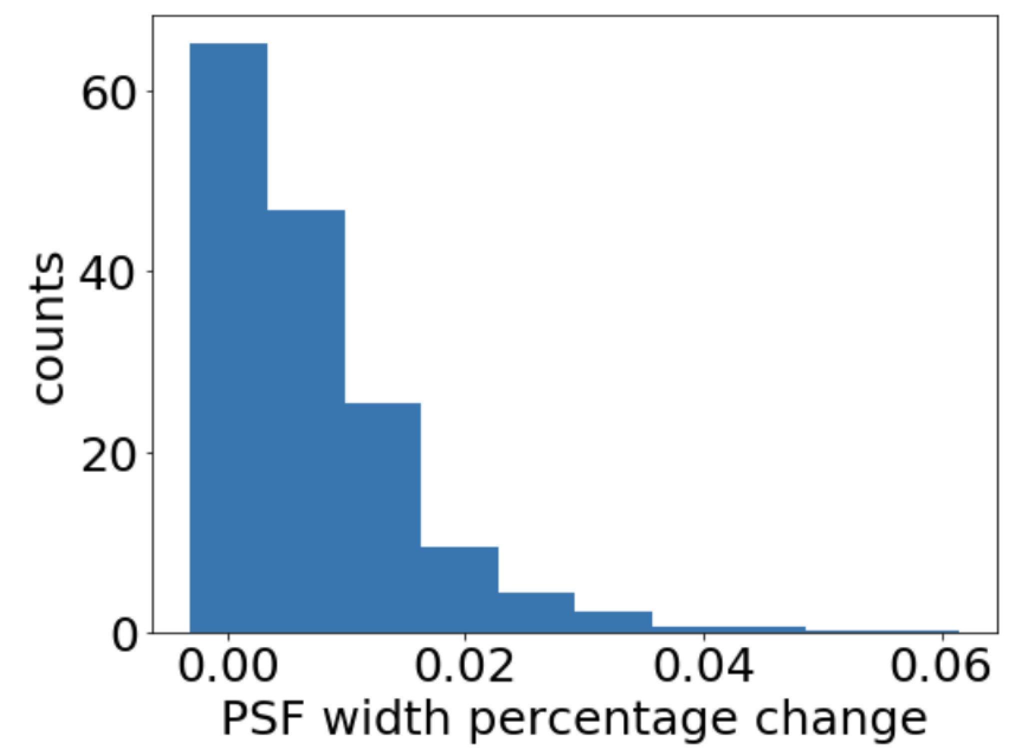
- scaled L2 loss;
- addition of a PSF term to the loss function;
- anti-aliasing pooling;
- self-attention

Donut image with all perturbations corrected in the optical system

Results



Prediction RMSE over error tolerance for the 50 control parameters. The RMSE of each CP output by the neural network is with in the tolerance.



Selecting the 10% worst cases based on PSF, we construct the PSF for typical seeing of ~0.65 arcsec, measure its FWHM, and find that the prediction error of the CPs makes only a small contribution to the FWHM.

Summary

- Using the scaled L2 loss and adding a PSF term to the loss function enhances performance substantially;
- Including anti-aliasing pooling, and augmenting the training data to include randomly shifted donuts, the resulting model performance is insensitive to image shift;
- Including self-attention modules in the CNN led to modest changes in performance;
- Significant up-front computational expense is rewarded with fast and accurate evaluation