

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

Deep learning-based focal plane wavefront sensing for classical and coronagraphic imaging

Quesnel, Maxime, Orban de Xivry, Gilles, Louppe, Gilles, Absil, Olivier

Maxime Quesnel, Gilles Orban de Xivry, Gilles Louppe, Olivier Absil, "Deep learning-based focal plane wavefront sensing for classical and coronagraphic imaging," Proc. SPIE 11448, Adaptive Optics Systems VII, 114481G (13 December 2020); doi: 10.1117/12.2562456

SPIE.

Event: SPIE Astronomical Telescopes + Instrumentation, 2020, Online Only

Deep learning-based focal plane wavefront sensing for classical and coronagraphic imaging

Maxime Quesnel^{a,b}, Gilles Orban de Xivry^b, Gilles Louppe^a, and Olivier Absil^b

^aMontefiore Institute of Electrical Engineering and Computer Science, University of Liège

^bSpace sciences, Technologies and Astrophysics Research (STAR) Institute, University of Liège

ABSTRACT

High-contrast imaging instruments are today primarily limited by non-common path aberrations appearing between the wavefront sensor of the adaptive optics system and the science camera. Early attempts at using artificial neural networks for focal-plane wavefront sensing showed some successful results but today's higher computational power and deep architectures promise increased performance, flexibility and robustness that have yet to be exploited. We implement two convolutional neural networks (CNN) to estimate wavefront errors from simulated point-spread functions in both low and high aberration regimes. We then extend our CNN model by a mixture density network (MDN) and show that it can assess the ambiguity on the phase sign by predicting each Zernike coefficient as a probability distribution. Our method is also applied with the Vector Vortex coronagraph (VVC), comparing the phase retrieval performance with classical imaging. Finally, preliminary results indicate that the VVC combined with polarized light can lift the sign ambiguity.

Keywords: Machine learning, convolutional neural networks, focal plane wavefront sensing, phase retrieval, vector vortex coronagraphs, high contrast imaging

1. INTRODUCTION

Exoplanet imaging faces several challenges, principally due to the high contrasts and small angular separations between exoplanets and their parent star. Coronagraphy and, for ground-based observations, adaptive optics (AO) offer effective instrumental solutions to these limitations. Nonetheless, residual wavefront errors (WFE) and in particular non-common path aberrations (NCPAs) between the science camera and the AO wavefront sensor are a key limitation. Slowly varying NCPAs produce quasi-static speckles¹ in the focal plane that are very difficult to calibrate or remove with post-processing techniques due to their dynamic. Additionally, these speckles share a very similar signature to that of planetary companions. Eliminating NCPAs is of prime importance for high-contrast imaging today.²

One approach to solve this issue is focal-plane wavefront sensing (FPWFS), which measures the NCPAs from focal plane images without requiring any additional optical element. An important hindrance to FPWFS is the so-called twin-image problem,³⁻⁵ i.e. the non-unique response in the focal-plane for certain phase modes. This corresponds to a sign ambiguity regarding Zernike modes of even radial order when using centro-symmetric pupils. Because knowing the sign of only one even mode is enough to lift the ambiguity, a standard solution is to incorporate phase diversity.^{3,6} This is achieved by introducing a known even mode aberration (typically a defocus) and subsequently by using two images, with and without phase diversity.

Iterative algorithms such as the Gerchberg-Saxton algorithm^{7,8} have been developed for FPWFS. Although they can predict the phase with good accuracy, they face some drawbacks: they need a sensible starting guess while they do not offer the guarantee to converge to a global minimum and their iterative nature make them not very efficient for real-time corrections. On the other hand, machine learning techniques have already been implemented for numerous applications with impressive success. Some properties of deep neural networks that can be advantageous for phase retrieval are their non-linear nature, real-time prediction capabilities and promising robustness. Convolutional neural networks (CNN) are particularly efficient at modeling the spatial information

Further author information: (Send correspondence to Maxime Quesnel)
E-mail: maxime.quesnel@uliege.be

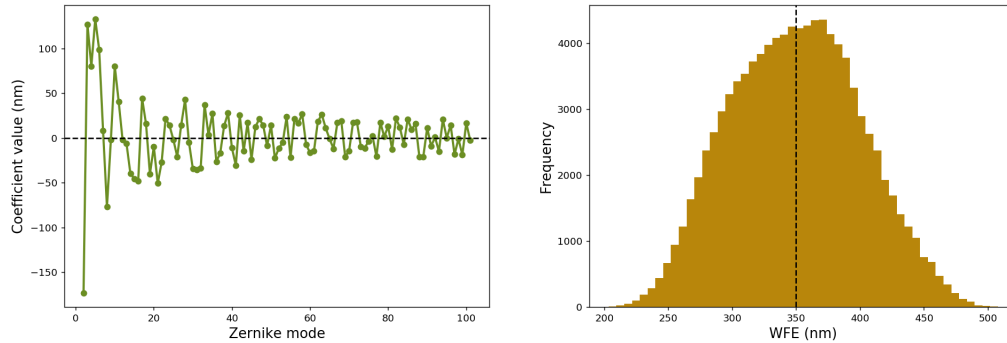


Figure 1: **Left:** Example of a generated set of Zernike coefficients. **Right:** Example of a RMS WFE distribution in a dataset composed of 10^5 samples. The median RMS WFE considered here is equal to 350 nm.

in images to infer real world quantities. Deep learning-based techniques have already been implemented for wavefront sensing as a supervised regression task with favourable results. Neural networks were first applied to estimate phase aberrations inside multiple-telescope arrays⁹ before being implemented in various projects, including phase retrieval inside the Hubble Space telescope.¹⁰ More recently, CNNs have also been used for FPWFS, targeting for example segment co-phasing¹¹ or atmospheric compensation.^{12,13}

In this paper we present our CNN-based method for focal plane wavefront sensing in high-contrast imaging. We most notably propose to implement mixture density models to predict the sign ambiguity with probability distributions and we test our method with the Vector Vortex coronagraph (VVC). In Sec. 2, we first give a detailed description of our data generation procedure and deep learning approach. We present in Sec. 3 a performance comparison between the U-Net and ResNet-50 architecture as a baseline, before evaluating the mixture density models and studying the impact of simulated post-VVC PSFs on the performance. The effect of polarization with vortex phase masks regarding the sign ambiguity is also examined. Finally, we conclude and discuss about the next steps ahead in Sec. 4. While the present work describes specific aspects related to our method such as the implementation with the VVC, a thorough study in the case of classical imaging is addressed by Orban de Xivry et al.¹⁴ (in preparation).

2. METHOD

In this section we describe the building components of our simulations and the deep learning approach considered to retrieve phase aberrations from point-spread functions.

2.1 Synthetic data generation

2.1.1 Dataset generation procedure

For focal-plane wavefront sensing in a deep learning framework we construct a labelled dataset composed of point-spread functions and phase maps. In classical imaging, i.e. with a direct optical propagation between pupil and focal planes, the PSF can be expressed from the phase as

$$PSF(x, y) = |\mathcal{F}[A(x, y) e^{i\Theta(x, y)}]|^2, \quad (1)$$

where $A(x, y)$ is the system's aperture function and $\Theta(x, y)$ the phase. One can note that due of the squared complex modulus, only intensities are encoded in the PSF and the phase information is therefore lost. We simulate the phase aberrations using Zernike polynomials which make up an orthonormal basis:

$$\Theta(x, y) = \sum_{i=0}^N c_i Z_i(x, y), \quad (2)$$

where Θ is the complete phase map, Z_i the Zernike polynomials, c_i the corresponding coefficients, and N the number of modes considered. The sets of Zernike coefficients are first randomly generated within the range $[-1, 1]$

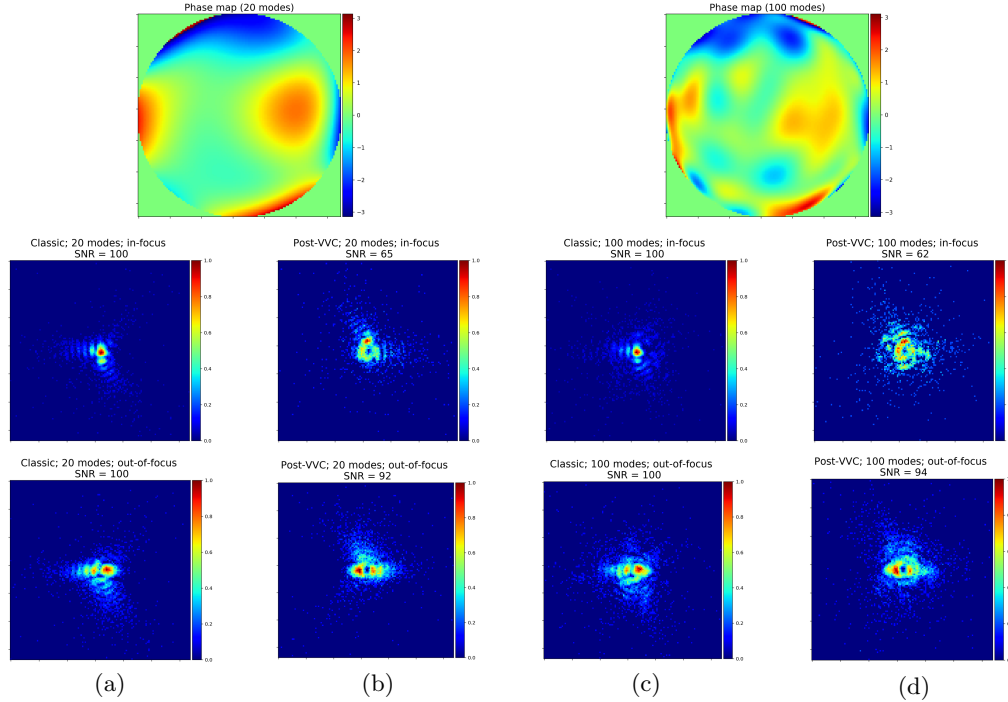


Figure 2: Example of phase maps and pre-processed PSFs with a median RMS WFE of 350 nm and an additional defocus of 550 nm. A fixed stellar flux of 10^4 photons (SNR = 100) is considered. **Top:** Phase maps with 20 (left) and 100 modes (right). **Middle:** Corresponding normalized in-focus PSFs. **Bottom:** Normalized out-of-focus PSFs. (a) Classical imaging PSFs; 20 modes. (b) Post-VVC PSFs; 20 modes. (c) Classical imaging PSFs; 100 modes. (d) Post-VVC PSFs; 100 modes.

and each coefficient is then divided by its corresponding radial order to approximate a $1/f^2$ power spectral density profile, typically encountered with good quality optics.¹⁵ The entire set of coefficients for each phase map is re-scaled according to a given median root-mean-square (RMS) wavefront error \overline{WFE} :

$$\hat{c}_k = \frac{c_k}{\overline{WFE} \times \text{median} \left(\sqrt{\sum_i c_{k,i}^2} \right)}, \quad (3)$$

where \hat{c}_k is the normalized set of coefficients for the phase map k (see an example Fig. 1; left). Our input wavefront error is therefore a distribution centered around \overline{WFE} (Fig. 1; right), and a higher \overline{WFE} leads to a larger distribution. We then construct the phase maps using Eq. (2) and we follow the Noll convention¹⁶ regarding the ordering of the Zernike polynomials, starting from the tip mode. We simulate a circular entrance aperture, i.e. without considering the central obstruction and structure support due to a secondary mirror.

The PSFs are computed using the PROPER¹⁷ optical propagation package. Our data is generated in the K band at $\lambda = 2200$ nm, we consider an aperture diameter of 10 m, a pixel scale of $0.25 \lambda/D/\text{pix} \simeq 11.43$ mas/pix, and a field-of-view of $32.3 \lambda/D \simeq 1.47''$. The phase maps and final PSFs are cropped to contain 129×129 pixels. Each sample from the dataset is composed of in-focus and also out-of-focus PSFs to lift the sign ambiguity of the phase. A defocus of $\lambda/4 = 550$ nm is introduced in the pupil plane to obtain the out-of-focus PSFs. Fig. 2 illustrates phase maps and PSFs generated with a median RMS WFE of 350 nm (1 rad) distributed over 20 and 100 modes.

2.1.2 Implementation of the Vector Vortex coronagraph

The Vector Vortex coronagraph¹⁸ (VVC) is an optical device designed to suppress the direct starlight in order to reach higher contrasts in the star's surroundings. It uses a transparent phase mask at the focal plane that diffracts the on-axis light outside the pupil. This light is then blocked by a downstream Lyot stop. The VVC generates

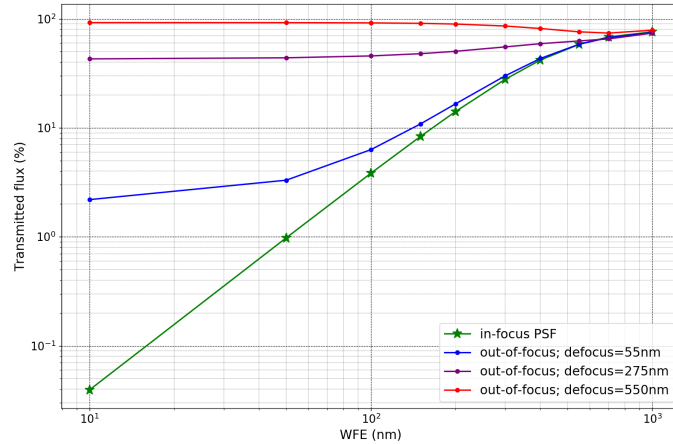


Figure 3: Transmitted flux after the vortex mask and the downstream Lyot stop as a function of the level of wavefront aberration, for different additional defocus. For each case a single data sample generated with 20 Zernike modes is used.

two conjugated phase ramps $\exp(i \pm l_p \theta)$, with l_p the topological charge and the θ the azimuthal coordinate, to both circular polarization of the incoming light. Among the advantages of vortex phase masks compared to other coronagraphs are their achromaticity and small inner working angle,¹⁹ and VVCs have been integrated with success in several high-contrast imaging instruments since 2012 (VLT/NACO and VISIR, Keck/NIRC2, LBT/LMIRCam).²⁰ For a flat wavefront and a circular pupil, the VVC perfectly cancel the on-axis PSF. If the wavefront contains phase aberrations however, some light will leak. Fig. 3 illustrates the influence of the incoming level of wavefront error on the transmitted flux ratio. One can notice that regardless of the additional defocus, the transmitted flux tends to converge towards a similar fraction for very high input wavefront errors. Since the VVC removes the Airy pattern, we can expect the speckle field to be more prominent in the focal plane. Considering a fixed integration time, the signal-to-noise ratio (SNR) will be lower compared to classical imaging. These two aspects have an impact for focal-plane wavefront sensing compared to a case without coronagraph (see Sec. 3.3.1 for some comparison results).

Our post-VVC PSFs are generated using the HEEPS²¹ optical propagation package which incorporates PROPER.¹⁷ We consider a topological charge $l_p = 2$. In order to simulate non-polarized light, two PSFs are generated with each opposite topological charge l_p before being added together. We also perform some experiments with only one sign of the phase screw ($l_p = +2$) to investigate whether selecting one circular polarization could lift the sign ambiguity (Sec. 3.3.2). The diffracted light outside the geometric pupil is blocked by a Lyot stop undersized to 98% of the pupil. To generate the out-of-focus PSFs the defocus is introduced in the entrance pupil plane instead of the Lyot plane. This is equivalent to injecting the defocus with the deformable mirror instead of shifting the detector out of focus. Hence, a defocus upstream of the vortex mask will produce a greater leakage of the on-axis PSF, producing higher SNR and thus more precise measurements. Examples of post-VVC PSF pairs can be found in Fig. 2.

2.1.3 Data pre-processing

In our simulations we only consider photon noise, which means the SNR is defined as $\text{SNR} = \sqrt{N_{ph}}$ with N_{ph} the number of photons. A SNR of 100 is chosen for most of our experiments. We also take into account the extinction factor induced by the VVC when comparing with classical imaging (Sec. 3.3.1). Then we apply a square-root stretching operation to the PSFs to help the CNN identify the speckle distribution. Finally, we normalize the PSFs with a min-max scaling to obtain flux in the range [0,1], which ensures we feed to the CNN same-scale quantities. This has been shown to provide better stability and performance during the training of deep neural networks.²²

2.2 Deep learning models

2.2.1 Convolutional Neural Networks

We build models whose goal is to map phase aberrations y from PSFs x , i.e. to approximate a non-linear function f such that $y \approx f(x)$. In order to do so, we have implemented two deep neural network architectures: ResNet-50²³ and U-Net.²⁴ Two other architectures, VGG-16²⁵ and Inception-v3,²⁶ have also been tested but produced poorer results on our datasets and are thus not presented in this study.

On the one hand, ResNet-50 is a residual neural network composed of 50 layers that uses skip connections by summation, improving the gradient's flow during backpropagation, as well as batch normalization layers. The architecture therefore overcomes the vanishing gradient problem²⁷ and the degradation inherent to very deep networks.²³ In our case ResNet-50 predicts the Zernike coefficients from PSFs. On the other hand, U-Net is an architecture developed for image segmentation that is composed of a downsampling path followed by upsamplings of the feature maps, producing outputs at the same dimension as the input. It also uses skip connections by concatenating feature maps from previous layers in order to retain relevant information. The softmax layer at the end of the architecture was removed to perform a regression task instead of segmentation and we infer directly phase maps from PSFs. Finally, with ResNet-50 the initialized weights used are pre-trained with the ImageNet dataset while U-Net is trained from scratch.

2.2.2 Mixture Density Layer

A Mixture density network²⁸ (MDN) predicts weighted sums of probability distributions instead of single quantities, allowing to estimate uncertainties in the data. We decide to use a mixture of Gaussian distributions for our models. We only need to modify the last layer of ResNet-50 to produce three parameters per Gaussian m and Zernike mode z : the mean $\mu_{z,k}(x)$, the standard deviation $\sigma_{z,k}(x)$ and the weight $\alpha_{z,k}(x)$ given to each Gaussian. From these outputs, we reconstruct the likelihood of a coefficient y_z given an input x as

$$\begin{aligned} p(y_z|x) &= \sum_k \alpha_{z,k}(x) \mathcal{N}(y_z|\mu_{z,k}(x), \sigma_{z,k}^2(x)) \\ &= \sum_{k=1}^N \frac{\alpha_{z,k}(x)}{\sqrt{2\pi} \sigma_{z,k}(x)} \exp\left(-\frac{(y_z - \mu_{z,k}(x))^2}{2\sigma_{z,k}^2(x)}\right). \end{aligned} \quad (4)$$

For our study we use two Gaussian distributions ($N = 2$), which means that we predict six targets for each Zernike mode instead of a single one. We also set $\sum_k \alpha_{z,k}(x) = 1$, with $\alpha_{z,k}(x) \in [0, 1]$. This architecture allows us to measure prediction uncertainties and to assess the sign ambiguity observed with even Zernike modes. We expect to predict one Gaussian close to the true value of the Zernike coefficient and another one around the opposite value when there is sign degeneracy, i.e. only for even modes. Such method could for instance be exploited to measure the capability of diversity techniques on lifting the sign ambiguity. In the following sections we call ResNet-MDN the ResNet-50 architecture combined with the mixture density layer.

2.2.3 Training procedure

To train our model we need to define a loss function that an optimizer will minimize by successive updates of the neural network's weights, allowing to fit the model to the data. In the standard case, i.e without the mixture density layer, our loss function is the root-mean-square error of the phase residuals

$$\mathcal{L}_{cnn}(\Theta, \hat{\Theta}(x; \phi)) = \sqrt{\frac{1}{N} \sum_{i,j} (\Theta_{i,j} - \hat{\Theta}_{i,j}(x; \phi))^2}, \quad (5)$$

where $\Theta_{i,j}$ and $\hat{\Theta}_{i,j}(x; \phi)$ are the true and estimated phase respectively, ϕ represents the neural network weights, x corresponds to the input PSF(s) and N is the total number of pixels in each phase map. On the other hand,

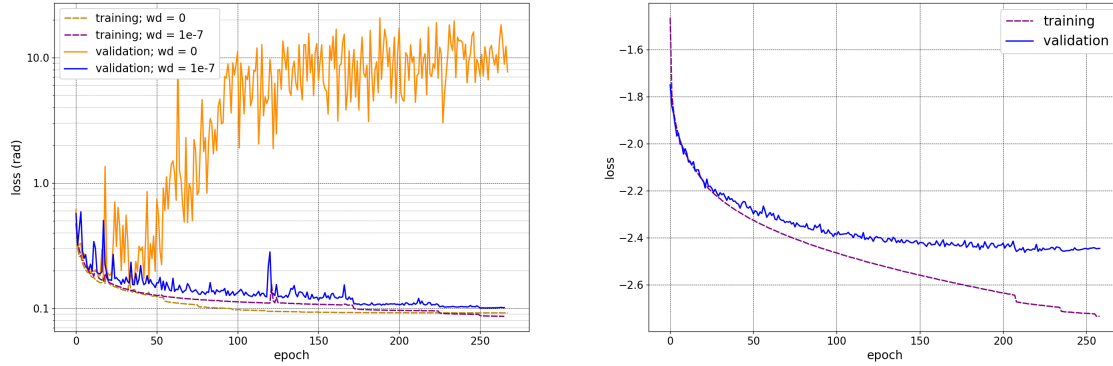


Figure 4: Illustration of training and validation losses obtained through the epochs, with phase diversity and on a dataset comprising 100 modes, a median input RMS WFE of 350 nm and classical imaging PSFs. **Left:** Using ResNet-50 with and without weight decay ("wd"). **Right:** Using ResNet-MDN (without weight regularization).

with ResNet-MDN architecture we want to maximize the likelihood $p(y_z|x)$ of the true Zernike coefficients y_z defined in Eq. (4). This amounts to minimizing

$$\mathcal{L}_{mdn}(y_z, x; \phi) = -\log p_\phi(y_z|x). \quad (6)$$

We also add a small offset $\epsilon = 0.005$ rad to the predicted $\sigma_{z,m}(x)$ to improve the stability of the minimization. During training the loss is averaged over the elements of the current batch before being fed to the back-propagation algorithm.

A dataset composed of 10^5 PSF pairs is randomly split into training (90%) and validation (10%) sets. A batch of 64 data samples is fed to the neural network at each iteration and we use the Adam optimizer,²⁹ which is a well known flavour of stochastic gradient descent. We set a learning rate of 10^{-3} for the CNN while a learning rate of 10^{-5} is preferred for ResNet-MDN to avoid training instabilities which we have observed with its loss function. We then decrease the learning rate by a factor of 2 each time the validation loss stagnates over 20 epochs, resulting in immediate drops of the loss, as it can be noticed in Fig. 4. Only the model defined at the epoch with the lowest validation loss is kept for evaluation.

Since we have observed ResNet-50 to be particularly prone to overfitting, we apply weight regularization with this architecture. Weight decay adds a penalty to the loss function and we found a factor of 10^{-7} to yield the best performance. More generally, DNNs trained with Adam and/or batch normalization have a specific response to weight decay, and multiple discussions can be found about this topic in the literature.³⁰⁻³² An example of the impact of regularization in our case is shown Fig. 4 (left). When training with only in-focus PSFs, even if overfitting can be overcome with a higher weight decay, performance do not improve in this case.

3. EXPERIMENTS AND RESULTS

In this section, performance are assessed with both classical and coronagraphic imaging and we also illustrate how mixture density models can exhibit prediction uncertainties. To evaluate our models, we use test datasets composed of 1000 samples generated the same way as the training and validation sets, except for a different random seed on the Zernike coefficients.

3.1 Comparison between CNN architectures

We give an overview of the performance of the ResNet-50 and U-Net architectures with our approach. A more detailed analysis is done by Orban de Xivry et al.¹⁴ (in preparation).

Here the models are trained with datasets containing classical imaging PSFs (without the VVC). Models trained with and without phase diversity are also compared. We make predictions on the test dataset and compute the root-mean-square error (RMSE) on the phase residuals, defined the same way as the loss function metric in Eq. (5). The performance are illustrated in Fig. 5 for a median input RMSE of 350 nm (1 rad)

Table 1: Comparison of RMS WFE on the phase residuals (in nm) averaged over 10^3 test samples. The models were trained on 10^5 PSF pairs with phase diversity and a SNR of 100.

Number of modes	20		100	
	70	350	70	350
U-Net	8.7	13.6	20.2	56.9
ResNet	8.8	12.5	19.9	38.9
ResNet-MDN	9.2	15.0	20.7	71.8

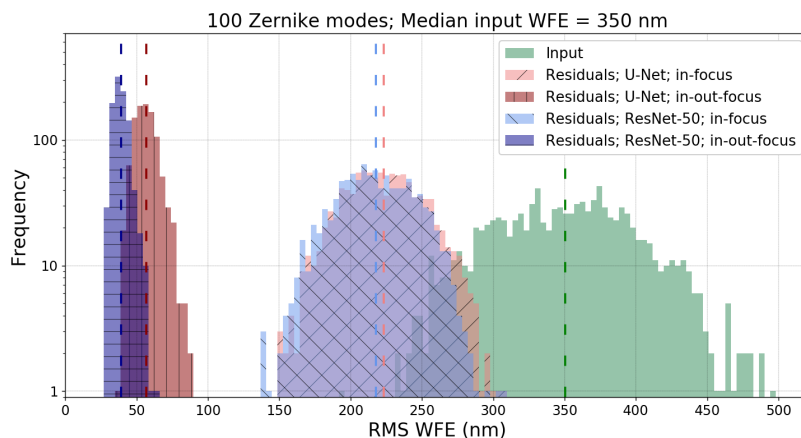


Figure 5: Empirical distributions of RMS WFE on the test dataset for models trained over 100 modes and with a median WFE of 350 nm (green). Performance with the ResNet-50 and U-Net architectures are compared, as are the cases with (in-out-focus; darker colors) and without (in-focus; lighter) phase diversity.

distributed over 100 Zernike modes and with a SNR of 100. The histograms represent the RMSEs computed from each test sample. Using diversity clearly reduces the error on the phase residuals, due to having two inputs instead of a single one (the effective SNR increases by a factor $\sqrt{2}$), but most importantly because the predictions are especially improved on the even modes, since the sign ambiguity is overcome. ResNet-50 is able to learn 100 modes slightly better than U-Net but, as shown in Tab. 1, we obtain almost identical RMSEs between the two architectures when trained with 20 modes or an input WFE of 70 nm. In the following experiments, we use exclusively ResNet-50 and ResNet-MDN.

3.2 Predicted uncertainties with mixture density models

From mixture model predictions, the RMS WFE can be assessed by computing the weighted mean $\hat{\mu}_z^* = \hat{\alpha}_{z,1}\hat{\mu}_{z,1} + \hat{\alpha}_{z,2}\hat{\mu}_{z,2}$ as the predicted Zernike coefficient for each mode z before obtaining the RMSE of the residual phase with Eqs. (2) and (5). We observe a degradation in performance with the mixture density models in terms of RMSE compared to the standard CNN models (Tab. 1). This is especially true for 100 Zernike modes and an input WFE of 350 nm, and we have observed that training on 5×10^5 images compensate for the difference in performance. We thus present the following results from models trained with this dataset size.

As illustrated with reconstructed distributions from the predicted parameters in Fig. 6, ResNet-MDN is well capable of recognizing the sign ambiguity when fed only with in-focus PSFs, by estimating two Gaussian distributions centered near the true coefficient and its opposite value. This works particularly well for low order Zernike modes, while the estimations degrade for higher order modes: the predicted absolute weight difference $\Delta\hat{\alpha}_z = |\hat{\alpha}_{z,1} - \hat{\alpha}_{z,2}|$ starts to be equal to 1 for even modes from $z > 20$, which means only one Gaussian is predicted instead of two (Fig. 7; left). In these cases the predicted $\hat{\mu}_z$ of the single Gaussian is approximately equal to zero, which is what the standard CNN predicts for even mode when facing the twin-image problem. More training data helps improve the predictions for high-order even modes. On the other hand, with phase diversity the MDN predicts a single Gaussian for every Zernike mode as expected.

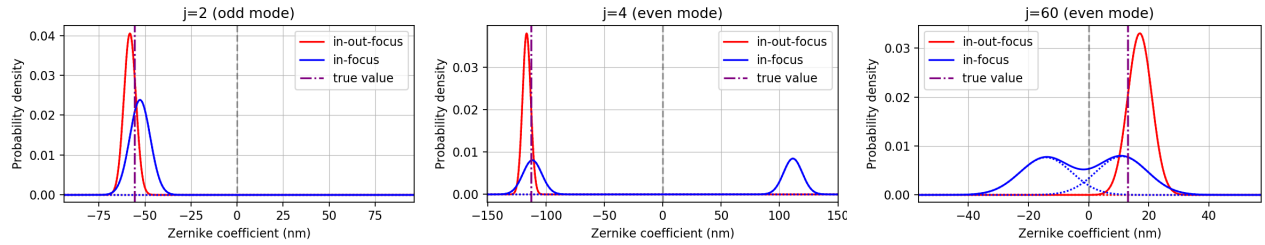


Figure 6: Examples of predicted distributions with a mixture model for an odd mode (tip; left), a low order even mode (defocus; middle) and a higher order even mode (right).

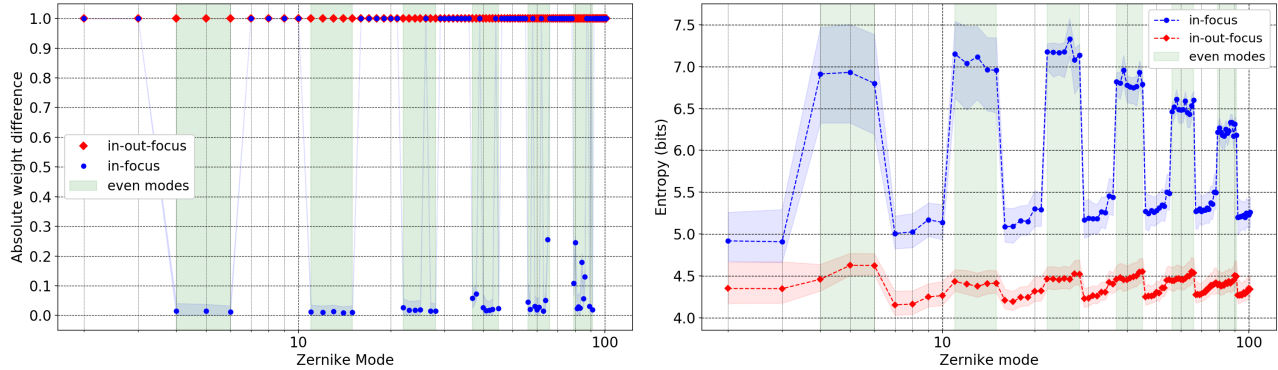


Figure 7: **Left:** Absolute difference of the predicted weights. A value of 1 indicates only a single Gaussian is predicted while 0 shows that two Gaussians with equal weights are obtained. **Right:** The entropy of the predicted Gaussian mixture for each Zernike mode. In both plots the median values are represented and the shaded regions extend to the 5-95th percentiles.

The metric we use here to measure the sign ambiguity is the entropy³³

$$H = - \sum_i p(y_i|x) \log_2 p(y_i|x), \quad (7)$$

where y_i are Zernike coefficient values (x-axis of Fig. 6). The advantage of this metric is that it depends both on the number of Gaussians and their widths, i.e. it considers the total uncertainty in the prediction. We show the entropy computed for each Zernike mode in Fig. 7 (right). The entropy is consistently higher for even modes compared to odd modes without phase diversity, which showcases the twin-image problem. One can also notice that for the high order modes for which ResNet-MDN is not able to predict two Gaussians (Fig. 7; left), the entropy is still important because the predicted $\hat{\sigma}_z$ is sufficiently large in these cases. Regarding odd Zernike modes, the entropy is always higher when trained only with in-focus PSFs since it lacks the additional out-of-focus PSFs that naturally improves the performance.

3.3 Influence of the Vector Vortex coronagraph

3.3.1 Performance comparison with classical imaging

In this section we study the performance using post-coronagraphic PSFs, i.e. a dataset generated with the VVC described in Sec. 2.1.2, comparing the residual errors obtained with models trained on classical imaging PSFs. In particular, we study three types of models: trained using classical PSFs and a fixed SNR of 100, using post-VVC PSFs at the same SNR of 100 (for both in and out-of-focus PSFs) and using post-VVC PSFs at a fixed stellar flux, resulting in lower SNRs in the focal plane because of the flux suppression caused by the vortex mask.

The results are displayed Fig. 8. At low input RMS WFE, the performance for each compared cases reach a plateau set by the photon noise limit.¹⁴ On the other hand, at high RMS WFE the performance degrades and this deviation from the photon noise limit relates to an additional model error (for classical imaging, see

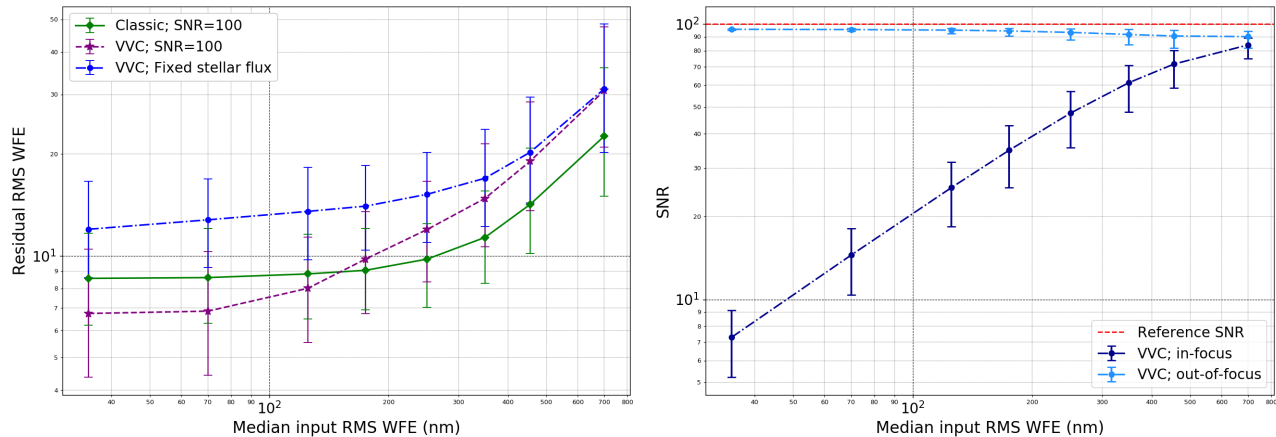


Figure 8: **Left:** Comparison of residual RMS WFE between models trained on post-VVC and classical imaging PSFs for different input WFE, each reported point corresponding to a specific model. **Right:** SNR of the in and out-of-focus PSFs for each model with the VVC, considering a fixed stellar flux (blue curve in the left plot). In both plots the median values are represented and the error bars correspond to the 5-95th percentiles.

Orban de Xivry et al., in preparation¹⁴). Models trained with post-VVC PSFs at a fixed stellar flux (blue curve) give higher residual RMS WFE for any input WFE which is essentially due to their lower SNR in the focal plane (see Fig. 8; right). Finally, models trained on post-VVC PSFs at a fixed SNR of 100 (purple curve) give better performance in the low aberration regime (≤ 125 nm) compared to classical PSFs at the same SNR, but poorer performance at higher aberration levels. Since the vortex mask removes the Airy disk, we would expect the aberrations to be better revealed in the focal plane, potentially improving learning the mapping between the phase and PSFs. This could explain the gain in performance obtained with the VVC at low levels of aberrations but the stronger degradation in performance with higher input RMS WFE is not understood at present. This trend has also been observed with models trained on other SNRs and level of additional defocus. However, the differences of performance are relatively small and mostly within the error bars.

3.3.2 Impact of polarized light on the sign ambiguity

We investigate whether polarized light could help to lift the sign ambiguity. For that we use a dataset composed of post-VVC PSFs simulated with only one circular polarization, as described in Sec. 2.1.2. Fig. 9 (top plots) shows the entropy of the predicted mixture of Gaussians from models trained at low (70 nm) and high (350 nm) aberration regimes, considering 20 Zernike modes and a SNR of 100. The figure demonstrates that the ambiguity is lifted in both aberration levels, which can be observed with an entropy that tends to be flattened along the radial orders (even compared to odd modes). Even though the prediction uncertainties can be higher with polarization regarding the odd modes, the resulting RMSEs on the residual phase maps are nonetheless lower (Fig. 9; bottom plots). The ambiguity appears to be lifted at all trained SNR levels.

Using phase diversity with additional out-of-focus PSFs clearly improves the performance, especially at higher SNR. It is likely that applying both circular polarizations as separated input to the CNN will reduce this gap in performance and an analysis in this regard is currently on-going. The main disadvantage of using defocused PSFs is that their measurements decrease the science duty cycle, while with the polarization approach all PSFs could be exploited. Depending on the scientific objective, polarized light can be beneficial,^{34,35} and an optical polarizing system could potentially be added after the VVC.³⁶

4. CONCLUSIONS

In this work we have developed a deep learning-based method to predict wavefront aberrations typical of NCPAs from simulated PSFs, notably using the ResNet-50 architecture. We have tested our approach on both classical and post-coronagraphic PSFs with vector vortex phase masks. Models trained with both types of PSFs yield similar performance at the same level of photon noise, albeit with slight differences. We have also introduced

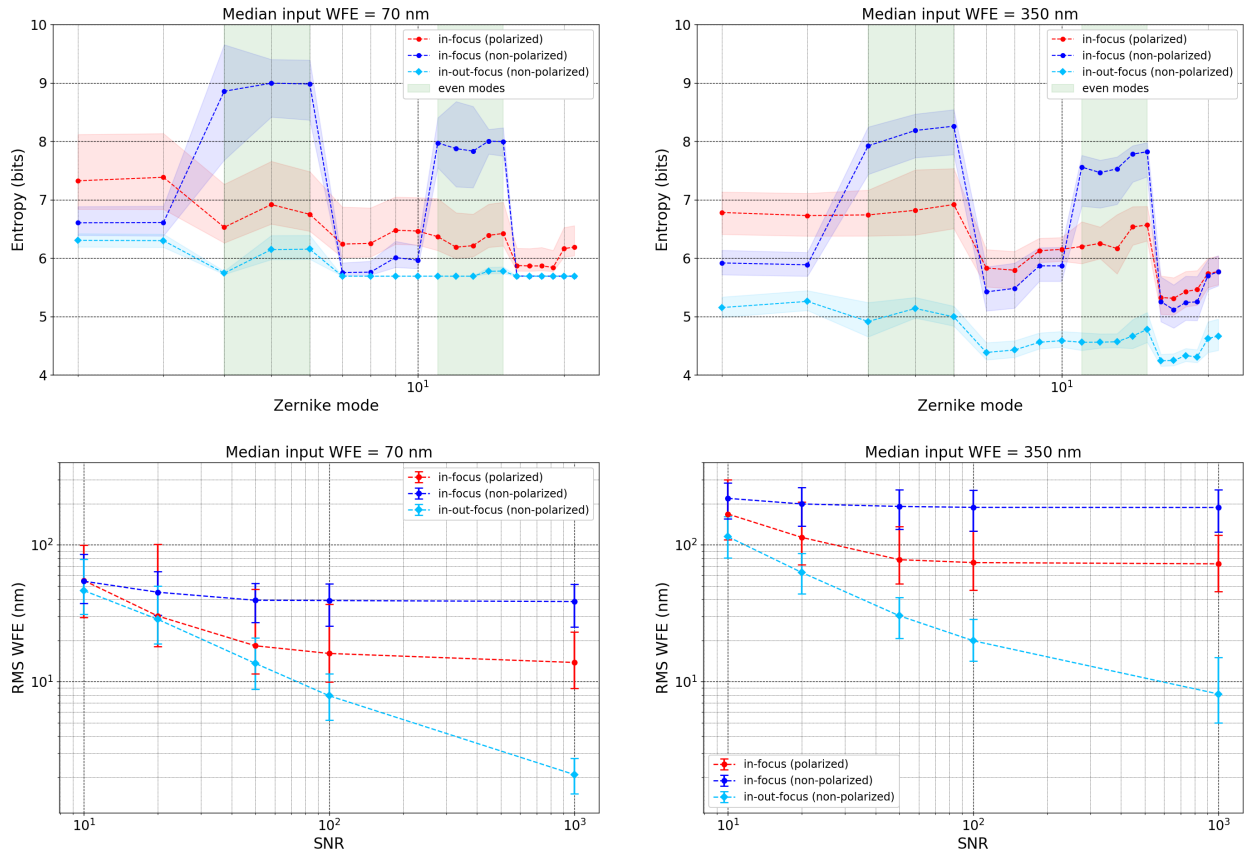


Figure 9: **Top:** Entropies on the predicted mixture of Gaussian distributions with phase diversity (cyan) and only in-focus PSFs with and without polarized light (red and blue respectively). **Bottom:** Residual phase RMSE for models trained at different levels of SNR. The datasets used are composed of post-VVC PSFs and a median input RMS WFE of 70 nm (**left**) and 350 nm (**right**), each distributed over 20 Zernike modes. In each case the median value is represented and the associated errors are the 5-95th percentiles.

mixture density models to predict the ambiguity on the phase sign that proved to work well in our setup, even though some limitations regarding higher order Zernike modes have been observed. Finally, a preliminary study indicates that the sign ambiguity can be lifted with polarized light when using the Vector Vortex coronagraph. The performance using both circular polarization as separated input images have yet to be assessed. A more detailed study of our method applied to classical imaging is presented by Orban de Xivry et al.¹⁴ (in preparation). In particular, it discusses the robustness of the models, how the performance relate to the theoretical limit and a comparison is made with the iterative Gerchberg-Saxton algorithm.

Future work will comprise the generation of more realistic datasets. For example, to consider the central obstruction in the entrance pupil can be crucial, especially with the VVC.³⁵ We are currently focusing on developing a CNN-based method to model real data using our simulations and in-lab data sets of much smaller size. Introducing other types of phase diversity techniques such as taking into account atmospheric turbulence residuals will also be studied. Finally, even though CNNs offer very promising results, applications to in-lab and on-sky observations with the possible use of WFS telemetry will most likely lead us to explore other machine learning techniques.

ACKNOWLEDGMENTS

For our study we have used the optical propagation package PROPER,¹⁷ HEEPS,²¹ and AOtools³⁷ to generate our datasets. We have trained our deep learning models using the library PyTorch³⁸ while Sacred³⁹ helped manage our experiments.

The research was supported by the the Wallonia-Brussels Federation (grant for Concerted Research Actions), and by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 819155).

REFERENCES

- [1] Macintosh, B., Poyneer, L., Sivaramakrishnan, A., and Marois, C., “Speckle lifetimes in high-contrast adaptive optics,” in [*Astronomical Adaptive Optics Systems and Applications II*], Tyson, R. K. and Lloyd-Hart, M., eds., **5903**, 170 – 177, International Society for Optics and Photonics, SPIE (2005).
- [2] Guyon, O., “Extreme adaptive optics,” *Annual Review of Astronomy and Astrophysics* **56**(1), 315–355 (2018).
- [3] Gonsalves, R. A., “Phase Retrieval And Diversity In Adaptive Optics,” *Optical Engineering* **21**(5), 829 – 832 (1982).
- [4] Guizar-Sicairos, M. and Fienup, J. R., “Understanding the twin-image problem in phase retrieval,” *J. Opt. Soc. Am. A* **29**, 2367–2375 (Nov 2012).
- [5] Bos, S. P., Doelman, D. S., Lozi, J., Guyon, O., Keller, C. U., Miller, K. L., Jovanovic, N., Martinache, F., and Snik, F., “Focal-plane wavefront sensing with the vector-apodizing phase plate,” *A&A* **632**, A48 (2019).
- [6] Vievard, S., Bos, S. P., Cassaing, F., Ceau, A., Guyon, O., Jovanovic, N., Keller, C., Lozi, J., Martinache, F., Mary, D., Montmerle-Bonnefois, A., Mugnier, L., &diaye, M., Norris, B., Sahoo, A., Sauvage, J.-F., Snik, F., Wilby, M. J., and Wong, A., “Overview of focal plane wavefront sensors to correct for the Low Wind Effect on SUBARU/SCEXAO,” in [*6th International Conference on Adaptive Optics for Extremely Large Telescopes, AO4ELT 2019*], (June 2019).
- [7] Gerchberg, R. W., “A practical algorithm for the determination of phase from image and diffraction plane pictures,” *Optik* **35**, 237–246 (1972).
- [8] Fienup, J., “Phase retrieval algorithms: a comparison,” *Applied optics* **21**, 2758–69 (08 1982).
- [9] Angel, J., Wizinowich, P., Lloyd-Hart, M., and Sandler, D., “Adaptive optics for array telescopes using neural-network techniques,” *Nature* **348**, 221–224 (1990).
- [10] Barrett, T. K. and Sandler, D. G., “Artificial neural network for the determination of hubble space telescope aberration from stellar images,” *Appl. Opt.* **32**, 1720–1727 (Apr 1993).
- [11] Paine, S. W. and Fienup, J. R., “Machine learning for improved image-based wavefront sensing,” *Opt. Lett.* **43**, 1235–1238 (Mar 2018).
- [12] Guo, H., Xu, Y., Li, Q., Du, S., He, D., Wang, Q., and Huang, Y., “Improved machine learning approach for wavefront sensing,” *Sensors* **19**, 3533 (08 2019).
- [13] Andersen, T., Owner-Petersen, M., and Enmark, A., “Image-based wavefront sensing for astronomy using neural networks,” *Journal of Astronomical Telescopes, Instruments, and Systems* **6** (July 2020).
- [14] Orban de Xivry, G., Quesnel, M., Vanberg, P.-O., Absil, O., and Louppe, G., “Focal plane wavefront sensing with convolutional neural networks,” (in prep., 2021).
- [15] Dohlen, K., Wildi, F. P., Puget, P., Mouillet, D., and Beuzit, J.-L., “SPHERE: Confronting in-lab performance with system analysis predictions,” in [*Second International Conference on Adaptive Optics for Extremely Large Telescopes. Online at <http://ao4elt2.lesia.obspm.fr> & <http://ao4elt2.lesia.obspm.fr/A>*], 75 (Sept. 2011).
- [16] Noll, R. J., “Zernike polynomials and atmospheric turbulence,” *Journal of the Optical Society of America (1917-1983)* **66**, 207–211 (Mar. 1976).
- [17] Krist, J. E., “PROPER: an optical propagation library for IDL,” in [*Optical Modeling and Performance Predictions III*], Kahan, M. A., ed., **6675**, 250 – 258, International Society for Optics and Photonics, SPIE (2007).

- [18] Mawet, D., Riaud, P., Absil, O., and Surdej, a., “Annular groove phase mask coronagraph,” *The Astrophysical Journal* **633**, 1191 (11 2005).
- [19] Mawet, D., Serabyn, E., Liewer, K., Burruss, R., Hickey, J., and Shemo, D., “The vector vortex coronagraph: Laboratory results and first light at palomar observatory,” *The Astrophysical Journal* **709** (12 2009).
- [20] Absil, O., Mawet, D., Karlsson, M., Carlomagno, B., Christiaens, V., Defrère, D., Delacroix, C., Castella, B. F., Forsberg, P., Girard, J., González, C. A. G., Habraken, S., Hinz, P. M., Huby, E., Jolivet, A., Matthews, K., Milli, J., de Xivry, G. O., Pantin, E., Piron, P., Reggiani, M., Ruane, G. J., Serabyn, G., Surdej, J., Tristram, K. R. W., Catalán, E. V., Wertz, O., and Wizinowich, P., “Three years of harvest with the vector vortex coronagraph in the thermal infrared,” in [*Ground-based and Airborne Instrumentation for Astronomy VI*], Evans, C. J., Simard, L., and Takami, H., eds., **9908**, 182 – 195, International Society for Optics and Photonics, SPIE (2016).
- [21] Carlomagno, B., Delacroix, C., Absil, O., Cantalloube, F., Orban de Xivry, G., Pathak, P., Agocs, T., Bertram, T., Brand l, B., Burtscher, L., Feldt, M., Glauser, A., Hippler, S., Kenworthy, M., Stuik, R., and van Boekel, R., “METIS high-contrast imaging: design and expected performance,” *Journal of Astronomical Telescopes, Instruments, and Systems* **6**, 035005 (July 2020).
- [22] Bishop, C., [*Neural networks for pattern recognition*], Oxford University Press, USA (1995).
- [23] He, K., Zhang, X., Ren, S., and Sun, J., “Deep residual learning for image recognition,” (2015).
- [24] Ronneberger, O., Fischer, P., and Brox, T., “U-net: Convolutional networks for biomedical image segmentation,” (2015).
- [25] Simonyan, K. and Zisserman, A., “Very deep convolutional networks for large-scale image recognition,” *CoRR abs/1409.1556* (2015).
- [26] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z., “Rethinking the inception architecture for computer vision,” *CoRR abs/1512.00567* (2015).
- [27] Ioffe, S. and Szegedy, C., “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” *CoRR abs/1502.03167* (2015).
- [28] Bishop, C., “Mixture density networks,” (1994).
- [29] Kingma, D. P. and Ba, J., “Adam: A method for stochastic optimization,” (2017).
- [30] van Laarhoven, T., “L2 regularization versus batch and weight normalization,” (2017).
- [31] Zhang, G., Wang, C., Xu, B., and Grosse, R. B., “Three mechanisms of weight decay regularization,” *CoRR abs/1810.12281* (2018).
- [32] Loshchilov, I. and Hutter, F., “Decoupled weight decay regularization,” (2019).
- [33] Shannon, C. E., “A mathematical theory of communication,” *Bell System Technical Journal* **27**(4), 623–656 (1948).
- [34] Stam, D. M., De Rooij, W. A., Cornet, G., and Hovenier, J. W., “Integrating polarized light over a planetary disk applied to starlight reflected by extrasolar planets,” *A&A* **452**(2), 669–683 (2006).
- [35] Mawet, D., Pueyo, L., Moody, D., Krist, J., and Serabyn, E., “The Vector Vortex Coronagraph: sensitivity to central obscuration, low-order aberrations, chromaticism, and polarization,” in [*Modern Technologies in Space- and Ground-based Telescopes and Instrumentation*], Atad-Ettinger, E. and Lemke, D., eds., **7739**, 378 – 390, International Society for Optics and Photonics, SPIE (2010).
- [36] Riaud, P., Mawet, D., and Magette, A., “Nijboer-zernike phase retrieval for high contrast imaging. principle, on-sky demonstration with naco, and perspectives in vector vortex coronagraphy,” *ap* **545**, A150 (09 2012).
- [37] Townson, M. J., Farley, O. J. D., Orban de Xivry, G., Osborn, J., and Reeves, A. P., “Aotools: a python package for adaptive optics modelling and analysis,” *Optics Express* **27**, 31316 (Oct 2019).
- [38] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S., “Pytorch: An imperative style, high-performance deep learning library,” in [*Advances in Neural Information Processing Systems*], Wallach, H., Larochelle, H., Beygelzimer, A., dAlché-Buc, F., Fox, E., and Garnett, R., eds., **32**, 8026–8037, Curran Associates, Inc. (2019).
- [39] Klaus Greff, Aaron Klein, Martin Chovanec, Frank Hutter, and Jürgen Schmidhuber, “The Sacred Infrastructure for Computational Research,” in [*Proceedings of the 16th Python in Science Conference*], Katy Huff, David Lippa, Dillon Niederhut, and Pacer, M., eds., 49 – 56 (2017).