



Experience with Artificial Neural Networks Applied in Multi-object Adaptive Optics

Sergio Luis Suárez Gómez^{1,2} , Carlos González-Gutiérrez^{2,3}, Enrique Díez Alonso^{2,3}, Jesús Daniel Santos^{1,2},
María Luisa Sánchez Rodríguez^{1,2}, Tim Morris⁴, James Osborn⁴, Alastair Basden⁴, Laura Bonavera^{1,2},
Joaquín González-Nuevo González^{1,2}, and Francisco Javier de Cos Juez^{2,3}

¹Department of Mathematics, University of Oviedo, Oviedo, Spain

²Instituto Universitario de Ciencias y Tecnologías Espaciales de Asturias (ICTEA), Oviedo, Spain

³Department of Exploitation and Exploration of Mines, University of Oviedo, Oviedo, Spain

⁴Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK

Received 2018 May 31; accepted 2019 May 1; published 2019 September 16

Abstract

The use of artificial Intelligence techniques has become widespread in many fields of science, due to their ability to learn from real data and adjust to complex models with ease. These techniques have landed in the field of adaptive optics, and are being used to correct distortions caused by atmospheric turbulence in astronomical images obtained by ground-based telescopes. Advances for multi-object adaptive optics are considered here, focusing particularly on artificial neural networks, which have shown great performance and robustness when compared with other artificial intelligence techniques. The use of artificial neural networks has evolved to the extent of the creation of a reconstruction technique that is capable of estimating the wavefront of light after being deformed by the atmosphere. Based on this idea, different solutions have been proposed in recent years, including the use of new types of artificial neural networks. The results of techniques based on artificial neural networks have led to further applications in the field of adaptive optics, which are included in here, such as the development of new techniques for solar observation or their application in novel types of sensors.

Key words: instrumentation: adaptive optics – methods: data analysis – techniques: image processing

Online material: color figures

1. Introduction

Adaptive optics (AO) has become an essential tool for improving the quality of images obtained by ground-based telescopes (Roddier 1999). The main goal of AO systems consists of being able to measure light aberrations produced by the rapidly changing turbulent atmosphere to correct the images produced by a telescope in real time. Atmospheric turbulence is first measured using sensors such as the widely used Shack-Hartmann wavefront sensor (SH-WFS; Platt & Shack 2001), which estimates the slopes or tilts in the deformed wavefront. The compensation is then performed with deformable mirrors (DM; Freeman & Pearson 1982), following calculations provided by reconstruction algorithms. These three elements can be combined in different configurations, such as the far more common single-conjugated AO (SCAO), followed by multi-conjugated AO (MCAO), multi-object adaptive optics (MOAO), laser-

tomography AO (LTAO), and ground-layer AO (GLAO; Tyson 2010).

Atmospheric turbulence is responsible for the aberrations present in the observation of a scientific object. It is a random phenomenon that distorts the light which goes through it. Models like Kolmogorov's model are used to represent atmospheric turbulence (Zilberman et al. 2008). It relies on parameters such as Fried's coherence length (r_0), which represents the intensity of the turbulence; the r_0 has a physical interpretation, the diameter of a telescope without atmospheric turbulence, that would have the same resolution as in our case with the atmosphere. A typical value of the r_0 parameter in a normal observation day is around 12–15 cm, while if the value is greater (17 cm or more), it indicates that the turbulence is very low. For values of r_0 smaller than 10 cm, we will have very poor observation conditions.

The atmosphere contains layers of moving air masses at different altitudes. Each layer has its own relative strength in the combined turbulence, depending on the turbulent profile (Dutton 1995). The altitude of the layers implies differences due to angular changes in the observed objects.

The search for qualitative improvements in the capability of observing the universe is translated in an increase in the size and



Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

resources of the instrumentation required. The newest and near-future telescopes are increasing in size and, consequently, in the amount of data retrieved (Rigaut 2002). This implies a big challenge for techniques such as AO, as the processing of large amounts of data slows the calculations of the corrections, which should really be performed in real time (Dipper et al. 2013).

Artificial intelligence (AI) techniques in AO are needed not only to improve the corrections provided by the established reconstruction algorithms, like the Learn and Apply (L + A; Vidal et al. 2010), based in a tomographic algorithm, or the Least Squares (Ellerbroek 1994) ones, but also to manage the large amounts of data involved. AI techniques are widely known in recent times as powerful tools in the handling of big data and the mathematical modelling of physical systems (Russell & Norvig 2016). This is mainly due to the flexibility of such techniques, which rely on the ability of AI to successfully learn directly from the data of the real problem, allowing AI to be applied in many different research fields, such as speech or image recognition (Graves et al. 2013; Krizhevsky et al. 2012).

This paper reviews the state of the art in AI, and in particular artificial neural networks (ANNs) related to multi-object adaptive optics (MOAO), one of the configurations implemented in telescopes (Gendron et al. 2011; Lardi re et al. 2014). Moreover, new work developed on the subject is discussed, including the authors' own experience in this field.

Some AO systems, like CANARY, allow the on-sky validation of the reconstruction techniques, which is why this work is focused on the MOAO approach. Although it is only one of many possible AO configurations, the application of AI in other configurations is also expected to provide interesting results.

The evolution of the machine learning use presented here responds to the necessity of improving the time and quality of the reconstruction. Newer and more complex paradigms show their performance, as the AO applications considered are more general and closer to real situations, or even implemented in actual telescopes.

The review is organized as follows. Section 2 deals with the earliest research on the topic, and Section 3 presents the approaches that led to the modern AI techniques. Section 4 focuses on the performance of the most successful technique in MOAO reconstruction based on AI, and Section 5 describes the work dealing with future implementations in more complex instrumentation. Section 6 presents some insights into transversal works developed with these techniques. Finally, Section 7 summarizes the conclusions of all the information compiled in this review.

2. Previous AO Work

In 1990, a multilayer perceptron (MLP) was used to determine the variation in pathlength and wavefront tilt

between elements of multiple-telescope arrays due to atmospheric turbulence (Angel et al. 1990). Specifically, the adjustments were carried out by the six 1.8-m mirrors of the Multiple Mirror Telescope (MMT) based in Tucson, Arizona, operating at infrared wavelength ($2.2 \mu\text{m}$). The MLP with a single layer of hidden nodes was fed only by a pair of simultaneous in-focus and out-of-focus images of a reference star formed at the combined focus of all the array elements. The resulting corrections were able to recover the diffraction-limited performance with a resolution of 0.06 arcsec.

Later, the previously mentioned research team was able to show that implementation of the neural network on a coprocessor, known as transputer array, achieved the required real-time performance (less than 10ms), thanks to the parallel structure of both the neural network and the coprocessor. Nonetheless, this work lacked continuation, as this kind of coprocessors was outdated by the rise of other computing devices such as graphics processing units (GPU; Wizinowich et al. 1991).

In the same year, another MLP was successful in wavefront sensing on images in a net processing time of less than 10 ms (Wizinowich et al. 1991). The adaptive stabilization of the mean phase errors between two mirrors, which led to stable fringes with 0.1 arcsec resolution, was demonstrated. The total time employed in the correction of the images was between 10 and 30ms. This time includes the readout time of the detector, the time to perform the calculations of the net and the time to move the adaptive mirrors.

It was then proved that an ANN (a 300 input node MLP connected to a single layer of 36 nodes with sigmoid activation functions and a six nodes output layer) could predict the shape of a distorted wavefront based on data of its shape in the immediate past (Lloyd-Hart 1994). In spite of this, it was also found that when the conditions in which the training had taken place slightly varied, the neural network was not as effective as it was when the conditions were similar. It was expected to be solved by wider training or an increase in the input nodes.

In 1991, an MLP using a back-propagation algorithm with a gradient descent technique was applied to in- and out-of-focus images of a real star, Vega, to estimate the optical phase distortion (Sandler et al. 1991). These images were obtained with the 1.5 metre single-mirror telescope at the Starfire Optical Range of the Air Force Phillips Laboratory near Albuquerque, New Mexico. The results of the experiment were in complete accordance with phase reconstructions obtained at the same time with a conventional SH-WFS: for Zernikes modes between 4 and 7 the root-mean squared error (RMSE) is $\lambda/14$, being $\lambda = 0.85 \mu\text{m}$.

Subsequently, it was demonstrated that a back-propagation MLP trained on real astronomical data was able to make good predictions about atmospherically-distorted wavefronts, obtaining Strehl of 0.36 with the MLP against 0.18 from the lagging system (Jorgenson & Aitken 1992). Those results implied

improvements in the performance of an adaptive optics system; nevertheless, only the first three Zernike modes (Noll 1976) were taken into consideration.

Two years later, the respective performances of MLPs and linear predictors were compared. Real astronomical data from a wavefront sensor with two channels, one for visible and one for infrared, was used for that purpose, neural networks being slightly better than linear predictors in more poorly seeing data sets; MLPs showed three times more squared error (Jorgenson & Aitken 1994).

In 1995, a numerical model of a multichannel adaptive optics system regulated by a feed-forward MLP was built (Vdovin 1995). The neural network was trained to predict the vector of adaptive mirror signals from the measured intensity vectors of the input aberrations. High efficiency of control was demonstrated—an error of less than 20% in the predictions—but spatial spectrum was limited.

The use of an MLP with a nonlinear hyperbolic tangent was investigated in 1996 (Montera et al. 1996). The hidden layer had 80 nodes and a nonlinear sigmoid or a linear summation output layer, to reduce the SH-WFS slope measurement error, and to estimate parameters such as the Fried coherence length (r_0), and the variance of the SH-WFS slope measurement error. The MLP was trained by standard back-propagation techniques. Results from their MLPs were compared with a classical statistics-based method and both were found to be successful in estimating values of r_0 , the best performance depending on the number of frames used for estimating the r_0 . Statistic methods performed better in reducing SH-WFS slope measurement error, while neural networks gave better results in estimating the variance of the SH-WFS slope measurement error; less than 25%, compared with the other methods, which had less than 35%.

In another work, the use of neural networks in predicting future SH-WFS slope measurements was investigated. Contrary to what a single statistical solution can do, it was shown that different MLPs were able to perform under a broad range of seeing conditions -different slopes-, making it possible to enlarge the set of wind speeds under which adaptive optics can work (Montera et al. 1997).

In 1999, different topologies/configurations of MLPs trained with back propagation were compared, in terms of predictive power in presence of noise (McGuire et al. 1999). A linear network predictor consisting of three layers, with only one hidden layer, was found to have a lower training time, residual phase variance, and higher tolerance to noise than the nonlinear neural network predictors. It was called linear because the output layer and the hidden layer had linear transfer functions, while the others had linear transfer functions for the output layer and sigmoidal transfer function for the hidden ones. The wavefront error was improved to $0.7rad^2$ while the error of the uncorrected wavefront was $1.71rad^2$ and managed to “improve the image resolution by a factor of three.”

Delving into their research, linear predictors and data were compared for a 2-m adaptive optics telescope simulated for the purpose. However, it was found that both linear networks and back-propagation training of nonlinear multilayer neural networks were quite slow, getting stuck on plateaus or in local minima, while recursive least squares training and adaptive natural gradient learning for linear predictors were found to be two orders of magnitude faster and convergent with global minimum error (around $0.01 rad^2$ in 1000 frames of 5ms each) (McGuire et al. 2000).

In 2004, a system was described (Chundi et al. 2004) for feature extraction based on a discrete cosine transform (DCT), well known at that time for its fundamental role in JPEG image standard compression, with the aim of obtaining a computationally-efficient neural network: a dimension reduction factor of 150 was achieved. The performance of a conventional MLP was compared with that of a radial basis function neural network (RBFNN). The RBFNN uses a set of Gaussian functions that cover the input space fully and its output consists of a linear combination of those functions. Both approaches were satisfactory candidates for the estimation of wavefront parameters, obtaining as they did a Strehl ratio of almost 95%.

In 2014, the problem of the treatment of images from SH-WFS in extremely large telescopes was addressed (Mello et al. 2014). An MLP, in the presence of turbulence, using back propagation for computing the centroid of elongated spots, was superior to existing techniques (reducing the average pixel error by more than 50%) and was a viable and noise-resistant technique for use in SH-WFSs. Although the use of noise in training gave better results, those results did not improve much when the noise level was lower than the training level.

A back-propagation MLP to sensorless adaptive optics was applied in 2017 (Wang et al. 2018). It was found that the MLP method greatly improved the system’s real-time capacity and also the Strehl Ratio, which measures the quality of optical image formation based on the peak intensity from the aberrated images. This method achieves a Strehl ratio of 0.70, whereas the other methods compared gave just 0.64.

3. Preliminary Approximations to MOAO with AI

Modern lines of research on this topic began with the comparison of AI techniques to model the control of actuators in a deformable mirror for open-loop adaptive optics (Guzmán et al. 2010). In this case, the mirror to be modeled had 97 electrostrictive actuators and the multivariate adaptive regression splines (MARS) proposed by Friedman (Friedman 1991; Sekulic & Kowalski 1992) turned out to be the optimal models.

The DM surface measurements were taken with an interferometer that works with monochromatic light at 633 nm, displaying the surface of the mirror in a 512×512 pixel-phase map. These measurements were used for the training; MARS models split the space of predictors, the

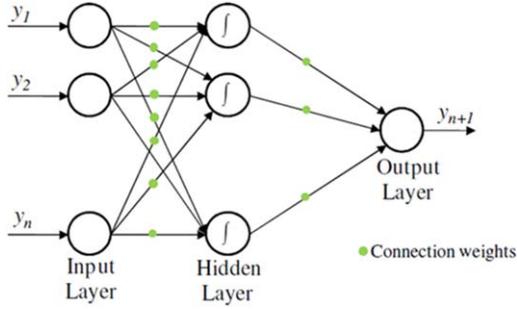


Figure 1. Example of a multilayer perceptron network topology. Adapted from Gomez et al. (2017). CC BY 4.0

(A color version of this figure is available in the online journal.)

possible intervals of values that the function can take, into several regions, which can have overlap between them. The splines functions fitted into the defined sections of the space of predictors. Relationships between variables were modelled with the piecewise polynomials, or splines.

These relationships can be expressed as Equation (1):

$$y' = f_M(x) = c_0 + \sum_{m=1}^M c_m B_m(x). \quad (1)$$

In this formulation, y' is the prediction of the MARS model for the dependent variable, c_0 is a constant, $B_m(x)$ is the m th single spline function (or a product of two or more spline basis functions), and c_m is its correspondent basis function coefficient. The basic functions to be included in the model are determined with the generalized cross-validation criteria (Friedman & Roosen 1995), which corresponds to the mean squared residual error divided by a penalty, based on the model complexity.

As the position of the actuators is that predicted by MARS in this problem, the use of these models implies an improvement in the deformable mirrors latency with respect to other iterative models, with reasonable errors.

Later the same year, the first models of ANNs were implemented to offer an alternative to the models used for controlling the DM shape (Guzmán et al. 2010). These ANN models were a feed-forward MLP, which is the one of the most-used types of ANN. The architecture of an MLP consists of neurons arranged in layers: the input layer has as many neurons as the number of input variables; in the same way, the output layer has the same number of neurons as the number of desired output variables. The desired number of hidden layers is placed between the input and output layer, with a selected number of neurons for each one. The neurons of each layer are connected with all the neurons in the next layers. The connections are determined by connection weights and the information flows through the network as is shown in Figure 1.

Source nodes in the input layers supply the input vector (y^{l-1}_i the input value from each neuron i), which constitutes the input signal for the neurons in the hidden layer. Each neuron computes

a weighted sum of its inputs from the previous layer, and the weighted sum is locally transformed by an activation or transfer function (f , the activation function of the neuron). The neuron then sends the result (Z^l_j is the output value of neuron j) to the neurons of the next layer. In this study, the area of the DM pupil was modeled to use surface measurements (positions of the actuator from the DM), corresponding to the actuator matrix, the matrix that represents the DM. The elements in the matrix are the position values corresponding to each singular actuator. The output was the predicted actuator position values. The computation performed by each neuron is expressed as follows (Equation (2)):

$$Z^l_j = f\left(\sum_{i=0}^n w^l_{ji} y^{l-1}_i\right), \quad i = 0, \dots, n; j = 1, \dots, m, \quad (2)$$

with w^l_{ji} the weight from neuron i in layer $l-1$ to neuron j in layer l . Once the signal has moved forward to all the layers of the network, the signal given by the last one, the output layer, constitutes the response of the network to the given input. To fit the characteristics of the function embedded in the data, a learning algorithm is needed to train the network. The learning is thus performed with iterations; the learning process performed once over the whole set of training samples is called an epoch.

As for the characteristics of the system, two different topologies were presented, with the aim of estimating the computational cost of the training: a network (ANNb from Figure 2) with a topology of 30 neurons in the input layer, 40 neurons in the hidden layer, and 30 neurons as output and a smaller one (ANNs from Figure 2) of 12 neurons in the input layer, 16 neurons in the hidden layer, and 12 as output. Sigmoid activation functions were applied in the hidden and output layers. To match the amount of data from the DM actuator matrix, preprocessing techniques were applied. The networks are proposed to clone or replicate the models on the physical characteristics and the isotropic behavior of DMs. The results are shown in Figure 2.

The training process of the two MLPs used the error back-propagation algorithm. The weights were updated according to gradient descent (without batch) with momentum in offline training.

The best results are obtained by the MARS model, which has the lowest error. However, MARS is also the model with the most outliers or inconsistent residuals, with a standard deviation of 2.2%. By comparison, the ANNb model has a standard deviation of 1%. The authors conclude that in large exposures, lower standard deviations may be preferred.

4. MOAO Reconstruction with Artificial Neural Networks

4.1. Simulation

These first attempts to improve MOAO systems suggested another perspective of the problem: the need not only to model

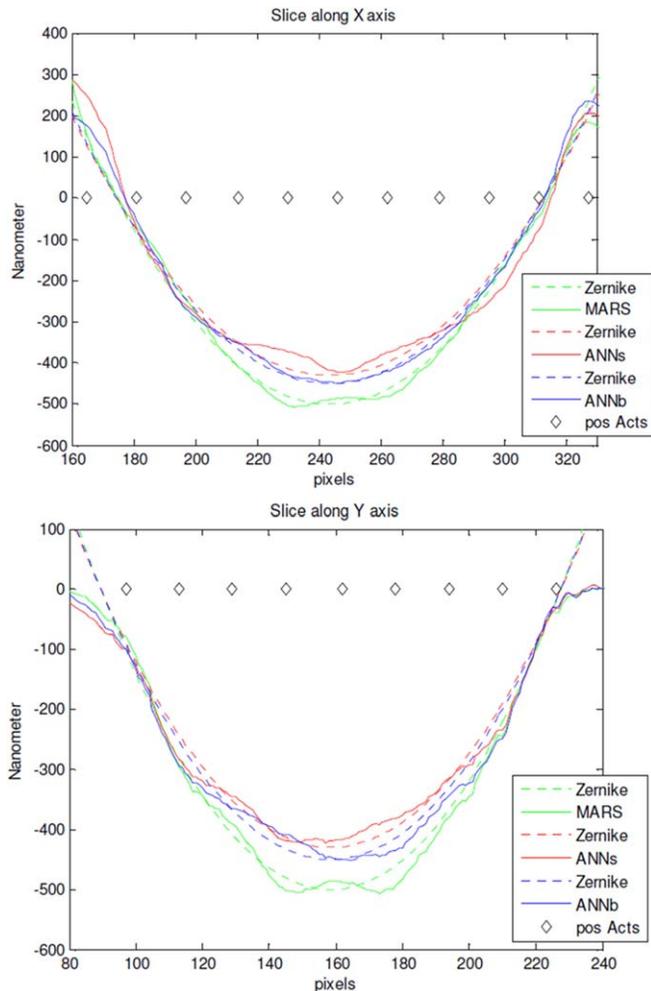


Figure 2. Slice in the X axis for a Zernike focus term (up). Slice in the Y axis for a Zernike focus term (down). Zernike estimations (dashed lines) and the correspondent estimation with MARS (green line), ANNs (red line), and ANNb (blue line) for test data. Adapted from Guzmán et al. (2010). ©2010 Optical Society of America. CC BY 3.0

(A color version of this figure is available in the online journal.)

the DM response, but also the effect of the atmospheric distortions on the upcoming light.

A tomographic reconstructor using an MLP with feed-forward propagation, via the information obtained from SH-WFS (Osborn et al. 2011), was the earliest model for patterning the response to the phase aberrations from the science target in a MOAO configuration.

In this article, the Complex Atmospheric Reconstructor based on Machine Learning (CARMEN) was presented. It was trained to correct any turbulent profile that it might be exposed to. To attain this goal, the neural network was trained with a large number of independent turbulence profiles. The MLP performs the reconstruction using the off-axis WFS slopes as

input and the desired on-axis target Zernike coefficients as output.

To set up the architecture of the network, comparisons with several combinations of training and architectures were performed, by changing the number of hidden layers, the number of neurons for each layer, the activation function, the number of epochs, and the number of turbulence layers to which the network is exposed during the training procedure. This leads to the determination of the optimal architecture for an AO system, which consisted of a single hidden layer with the same number of neurons as number of inputs. The topology determined here provided the basis for all the improvements and adjustments in later developments, where ANNs were used as AO reconstruction techniques.

The application of CARMEN used SH-WFS data from a new turbulent profile, giving as a response the estimation of on-axis Zernike coefficients. At this stage, the number of Zernike coefficients was limited to a maximum of 27 Zernike coefficients for the 27 first Zernike models (predicting up to 6th order Zernikes, instead of a value of the order of the number of actuators in the DM).

Once CARMEN had been set as a reconstruction technique, the first tests were performed using simulations in scenarios similar to those already presented, being trained with SH-WFS measurement of a MOAO system and estimating the Zernikes of the simulated image with aberrations (de Cos Juez et al. 2012). This is also expanded in Guzman et al. (2012).

A snapshot of an atmospheric profile is principally defined by several parameters: its r_0 , the number of turbulence layers that appear in the atmosphere, the altitude of these layers and their relative strength compared with other layers. As was established in the studies described above, CARMEN was trained using a single turbulent layer and a fixed r_0 of 12 cm, as it provides a good balance between performance and amount of training data. The height of the turbulent layer ranges from 0 m to 15500 m with 100 m step for training data, for a total of 155 different altitudes. At each altitude, 1000 samples of phase screens were randomly generated; this led to the creation of 155,000 training samples. The network was trained with Monte Carlo simulation data from Durham AO real-time controller (DARC; Basden et al. 2010), for different training scenarios. The simplest network is comprised of an MLP of just one hidden layer, containing the same number of neurons as the input (allowing full mapping), and a back-propagation training algorithm with a sigmoid activation function and a value of learning rate of 0.01 (Osborn et al. 2012).

Test data was also simulated from DARC. The first test, named atm1, had an r_0 of 16 cm with four layers of altitude 0, 4000, 10,000, and 15,500 meters, with relative strengths of 0.65, 0.15, 0.1, and 0.1 respectively. The second test, named atm2, had an r_0 of 12 cm with four layers of altitude 0, 2500, 4000, and 13,500 meters, with relative strengths of 0.45, 0.15, 0.3, and 0.1, respectively. The last test, named atm3, had an r_0

Table 1

Reconstruction Errors for Test Data, Consisting of Three Simulated Atmospheres of Different Turbulent Intensities from the Weakest (atm1) to the Strongest (atm3)

Test Name	Reconstructor	Strehl Ratio	FWHM (arcsec)	E50d (arcsec)	WFE (nm)	Improvement in WFE (%)
atm1	Uncorrected	0.048	0.319	0.482	644	...
	LS	0.296	0.099	0.299	293	54.50
	L + A	0.402	0.089	0.293	251	61.02
	CARMEN	0.462	0.088	0.279	231	64.13
atm2	Uncorrected	0.025	0.458	0.633	817	...
	LS	0.230	0.100	0.443	322	60.58
	L + A	0.300	0.091	0.436	289	64.62
	CARMEN	0.370	0.088	0.393	262	67.93
atm3	Uncorrected	0.012	0.684	0.912	1088	...
	LS	0.068	0.143	0.690	454	58.27
	L + A	0.100	0.104	0.688	409	62.40
	CARMEN	0.125	0.101	0.660	387	64.43

Note. All of the tests were simulated with four layers at different altitudes. metrics considered were WFE in nm, PSF strehl ratio, FWHM in arcseconds, and E50d in arcseconds. The error metrics were calculated comparing CARMEN performance with other reconstructors. reproduced from Osborn et al. (2012).

of 8.5 cm with four layers of altitude 0, 6500, 10,000, and 15,500 meters, with relative strengths of 0.80, 0.05, 0.1, and 0.05, respectively. These tests were chosen to compare fairly with other reconstructors, with different simulated phase maps. The selected conditions match those atmospheric profiles in La Palma measured in good, average, and poor conditions (Osborn et al. 2012).

The training was performed with BP algorithm with gradient descend (without batch) and in offline training. All the results and experiments realized from now on, when considering simulated data, is done in single runs. Comparisons between different standardized reconstruction techniques used in MOAO are performed. In particular, techniques such as Least Squares (Ellerbroek 1994) and Learn and Apply (Vidal et al. 2010) were used. Techniques like Linear Quadratic Gaussian (LQG; Paschall & Anderson 1993; Petit et al. 2008; Sivo et al. 2014) were discarded because CARMEN was not a predictive technique, so it was considered as being incomparable with such techniques.

To assess the results given by the reconstructed wavefront, the values listed in Table 1 correspond to the error metrics. The first column presents the name of each case, i.e., the simulated atmospheres (from atm1, the less turbulent, to atm3, the most turbulent) and the second column lists the different reconstructors used in each case. The PSF Strehl ratio, the azimuthally-averaged PSF FWHM, the diameter of 50% encircled energy (E50d) in the H band (1650 nm), and the wavefront error (WFE) are shown in the columns from three to six, respectively.

Table 1 shows the most representative results in the comparison between different reconstruction techniques. Moreover, several tests were also performed considering different turbulent profiles and metrics, and the generally promising results encourage further development and improvement of the

reconstructor, to be much more competitive with the already existing MOAO reconstructors; in particular, CARMEN seems more robust to changes in turbulence strength.

4.2. On-sky

The encouraging results obtained by CARMEN when applied with simulation data lead to the implementation of the reconstructor to deal with a real telescope situation; see Osborn et al. (2014).

The development of their reconstructor technique (CARMEN) was based on previously acquired knowledge in topology and training specificities. Aiming to be implemented on real telescopes, the offline training was required for suiting the target. For the same reason, a calibration with an AO bench was also performed. These adjustments to the original training were performed with the CANARY calibration unit, used to generate additional training data.

CANARY is a flexible AO demonstration bench at the 4.2-m William Herschel Telescope (La Palma) (Gendron et al. 2011). The bench has a modular design that allows for the testing and validation of early developments and concepts in the field of AO, and in the wider field of astronomical instrumentation.

To be as accurate as possible for all the prototypes that will be tested with CANARY, the bench makes it possible to simulate the atmosphere and the telescope calibration unit. In particular, CANARY contains a truth sensor (TS), an additional on-axis SH-WFS that made the calibration of CARMEN possible, being then improved to reconstruct the on-axis slopes, regardless of the atmospheric turbulence profile, using the off-axis slopes from the guide star SH-WFSs as inputs. In particular, these slopes measured from the SH-WFS are known as centroids, the focused points from the light that reaches each sub-aperture of the sensor. The coordinates of each centroid

were used, then, as the input and output values for the calibrated CARMEN. The bench validation enabled the adequate preparation for the subsequent implementation on-sky.

The CARMEN reconstruction technique, with the previous calibration on bench, was finally implemented on a real telescope on the nights of 2013 July 22 and 24. The calibration was performed on the first night.

CANARY was operated by switching between the L + A and CARMEN tomographic reconstruction techniques to prevent bias in the results by using different reconstructors at different times during changeable conditions. Data acquired corresponded to 36 exposures that were made with each technique, recording the Strehl ratio from the CANARY science camera in the *H* band.

The real data obtained from the telescope implementation can be found in Osborn et al. (2014), as can the recorded Strehl and the Zernikes from the recovered phases. It is compared with the performance of ground-layer adaptive optics (GLAO; Tokovinin 2004), which is a particularly useful technique when turbulence is concentrated in a turbulence layer at a very low altitude, known as a ground layer.

The results given by CARMEN in reconstructed Strehl and Zernike variance were very close, but slightly lower than L + A (the results from L + A were approximately 5% better than those of CARMEN). However, the greatest advantage of CARMEN is that it is more stable when considering altitudes of varying turbulence, according to in-lab results shown in Table 1.

5. Development at ELT Scales

The development of large telescopes, in particular the future European Extremely Large Telescope (E-ELT), the Giant Magellan Telescope (GMT) and the Thirty Meter Telescope (TMT), highlights the need for great computational ability to process enormous amounts of data (Ramsay et al. 2014).

On larger scales, such as the E-ELT, the use of SCAO and MCAO configurations is expected. In the baseline configuration it has six laser guide stars (LGS) and three natural guide stars (NGS). Each LGS has a Shack-Hartmann (SH) wavefront sensor (WFS) of 84×84 subapertures that operated at 500 Hz, used for measurement of slopes. It was possible for the images obtained from the SH to be used as input for the training of CNNs. If the original configuration of CARMEN is maintained (with the same number of neurons in the hidden layer as in the input), it will have around 127k input neurons, 127k hidden neurons and 14k output neurons. This will lead to around 18 billion operations to get an output. If the frequency of the inputs is 500 Hz, around 9 trillion operations per second will be needed to run CARMEN in the E-ELT.

CARMEN was developed with the statistical software R, providing a slow performance both in training and recall, as

was shown in Osborn et al. (2014). To be ready for the challenge provided by large telescopes in the near future, all efforts were focused on improving the algorithms dealing with the huge amounts of data that will be collected by the sensors of such telescopes in future.

The implementation of GPUs allowed for a high level of parallelization of convolutional neural networks (Krizhevsky 2014) in different applications, like voice and image recognition and also in AO (Marichal-Hernández et al. 2005). Given the recent success of neural networks, different frameworks were designed to speed up the training and recall of ANNs, allowing scientists to reach alternative solutions in their research.

In an early phase, different frameworks were tested, exploring their various capabilities and their suitability for use with CARMEN. The first option was Torch (González-Gutiérrez et al. 2016), a framework developed in Lua with GPU support, which provides fast execution and the possibility to improve and complete the framework by importing external modules. After this initial study, CARMEN was also developed in Theano, which is a Python library developed at the University of Montreal and uses other popular Python libraries such as numpy and Scipy. It provides symbolic differentiation and GPU support, which make it very popular among researchers (Al-Rfou et al. 2016). Theano was around 1.5 times faster in training than Torch (Suárez Gómez et al. 2016).

This first approach to the different frameworks concluded with a detailed comparison of the performance of both previous frameworks, while also adding Caffe and CUDA. Caffe is a deep learning framework developed by the Berkeley Vision and Learning Center. It has GPU support and provides several interfaces for Matlab, Python, and command line execution. Caffe is easier to use than Torch and Theano, as it only requires the edition of a text file with the parameters. CUDA is a native GPU language, which enables developers to create an ad hoc solution for each neural network. However, creating each neural network requires a bigger effort than in the case of the frameworks.

In González-Gutiérrez et al. (2017), data from different-sized telescope sensors was compared for both training and recall times. The code developed in CUDA is the best solution in all scenarios, being at least around 1.5 times faster than others for training and twice as fast in recall than the others.

However, although the obtained results are an improvement on the initial development of CARMEN, these implementations are not easily scalable, because they are limited to the use of a single GPU needing to be improved before being used in a real telescope. To this end, a new implementation was developed which allows the use of as many GPUs as the system has available.

Consequently, new studies for evaluating the performance of the different frameworks were required. Torch provides the “*DataParallelTable*” module, which allows for an easy multi-GPU implementation. In the case of Caffe, the regular interface

already provides easy multi-GPU support by simply splitting the input. In the case of CUDA, new codes were developed to split in different GPUs, as it is not supported by the original implementation. Theano was abandoned by its developers, so it was omitted from this new comparison (Lamblin 2017) and TensorFlow, which has achieved significant growth in recent years, was included instead (Abadi et al. 2016). TensorFlow is a new framework developed in Python with multi-GPU support created by Google. One of the key advantages of this framework is the use of graphs, which make it easier to deploy, debug, and test the different neural networks.

The comparison was performed with single runs and simulated data of different available configurations of CANARY (González-Gutiérrez et al. 2018) and, consequently by comparing networks based on CARMEN with topologies of the correspondent sizes; CANARY B1 is designed with four natural guide stars (NGS), its SH-WFS had 7×7 subapertures, with 36 of them activated regarding the telescope pupil and secondary obscuration. CANARY C2 is designed for MOAO applied with four laser guide star (LGS), and SH-WFS of 14×14 subapertures that correspond to 144 active ones. DRAGON is the largest AO system considered, which is still under development and consequently, the most challenging situation for a CARMEN-like network, having four LGS and four NGS, where each sensor has 704 operative subapertures.

This study led to two main conclusions. First, the frameworks considered are not suitable for speeding up the reconstructor when using more than one GPU. Second, the use of several GPUs with CUDA is not optimal with small size sensors, as both the training and recall times increase with the number of GPUs (up to four times slower). However, for DRAGON (Basden et al. 2016), the largest system used in the comparison, CUDA is able to speed up the training process when the number of GPUs increases, as can be seen in Figure 3.

In the case of recall, two different situations needed to be studied. There are two possible scenarios concerning how the reconstructor would be implemented in a real telescope. In the ideal case, the reconstructor would read the sensor information directly from the random access memory (RAM) of the control system. This would allow for a higher level of integration and a faster output and recall of CARMEN. However, as this scenario cannot always be assumed, it is necessary to study the situation where the information has to be read directly from the hard drive. In Figure 4, both cases are analyzed again for DRAGON, the biggest system currently available. Results show that increasing the number of GPUs provides a significant improvement (up to twice as fast) in recall, which is expected to keep accelerating when the size of the sensor and number of GPUs increase. In Figure 4, an average of the recall of 10,000 samples was used, with a standard deviation below 1%.

These results indicate that in the case of extremely large telescopes, with a large number of SH-WFS and subapertures, the use of GPUs is a suitable solution for managing that amount

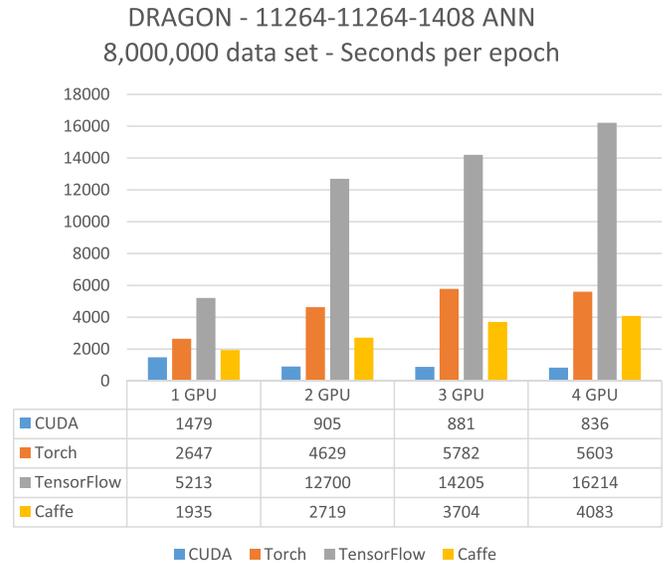


Figure 3. Training times for DRAGON (8,000,000 samples), comparing different frameworks and numbers of GPUs used for training. Adapted from González-Gutiérrez et al. (2018). CC BY 4.0.

(A color version of this figure is available in the online journal.)

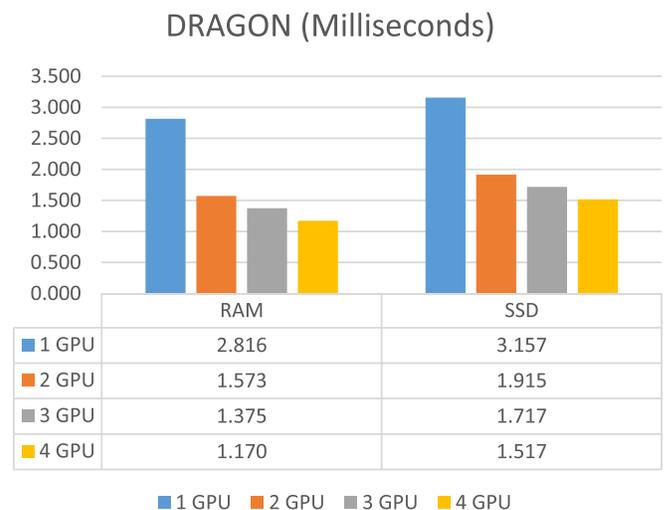


Figure 4. Average recall times (10,000 samples have been used), from DRAGON in CUDA when the number of GPUs changes. Loading data from solid-state drive (SSD) and directly from random access memory (RAM). Adapted from González-Gutiérrez et al. (2018). CC BY 4.0.

(A color version of this figure is available in the online journal.)

of data. Moreover, increasing the number of GPUs can speed up both the training and the recall of the reconstructor.

6. Current Challenges

Different studies based on the MOAO reconstructors developed with ANNs have been carried out concerning

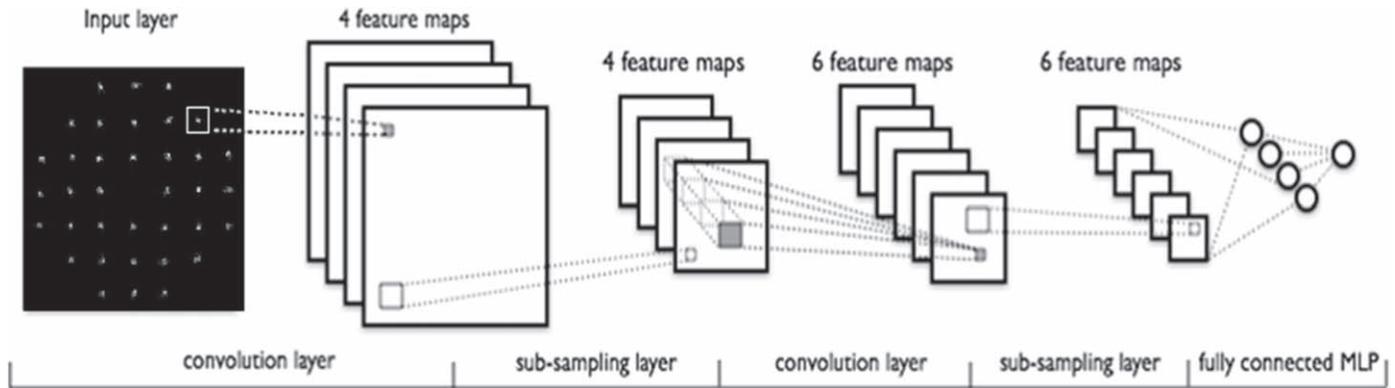


Figure 5. Example of the topology of a convolutional neural network. Images can be used as input, followed by sequences of convolution and subsampling layers, until a multilayer perceptron is reached. Reprinted by permissions from Springer Nature: Springer Nature eBook (Suárez Gómez et al. 2018). © 2018.

different ways of processing data, resulting in the first steps in the elaboration of alternative models.

The next evolution of CARMEN involves the inclusion of all the information from the SH-WFS rather than simply working with centroids. To achieve this goal, the full image of the sensor is needed as input: this can be accomplished by using convolutional neural networks (CNN; Krizhevsky et al. 2012). This subtype of ANN makes it possible to use images as input and apply several filters to those images to extract their main features and reconstruct the wavefront, as illustrated in Figure 5.

The CNN input used the full image of three SH-WFS, therefore with three channels. It was composed of four convolutional layers, each with four kernels of 5×5 size, followed by ReLU as the activation function. Pooling was applied, with sizes of 2×2 for the first two convolutional layers, and 4×4 sizes for the last two convolutional layers. This resulted in 768 images of 2×2 size that were connected, as input, to the fully connected layers; these had 3072 neurons as input. The hidden layer was set with 216 neurons and 72 for the output layer, the centroids of the SH-WFS of the scientific object.

The training was performed with a single turbulent layer, simulating data that varies the altitude and the turbulence strength by means of the parameter r_0 . Each of the combinations of heights (from 0 to 15500 meters, with steps of 100 meters each) and r_0 values (from 8 to 20 cm, with steps of 1 cm each) was sampled 100 times, there being a total of 202800 samples. To test the performance of the reconstructor, three different data sets were simulated. Each of them was composed of two layers with equal strength, with the first layer at 0 meters and the second with fixed altitude: 5000 m for the first, 10,000 m for the second and 15,000 m for the last one. For each test set, the r_0 values varied from 5 to 20 cm with 1 cm step, and were sampled 1000 times, there being a total of 16,000 samples in each set. These sets were simulated based on the three

extreme test cases (de Cos Juez et al. 2012; Osborn et al. 2012). By using these tests, it is possible to observe how the reconstructor behaves when changing two of the main components of a turbulence profile: its r_0 and the altitude of the layers. It may be possible to increase the altitudes for the second layer, but according to previous studies, these three situations should provide enough variability for the purposes of testing. Also, using several values of r_0 makes it possible to test how the reconstructor behaves when the strength of the turbulence changes. For the MLP training, centroids from these samples are used as input; for the convolutional, the images of SH of the same samples are used as input.

The convolutional approach performs better for single runs, in terms of normalized error (as a percentage of error), it being an alternative to the original CARMEN implementation, as shown in Figure 6. The principal advantage of this technique is that it allows us to obtain more information than just using centroids coordinates, providing better results than the MLP version of CARMEN (Suárez Gómez et al. 2018). However, this new reconstructor needs to be checked further, with optical measurements and for bigger systems in real telescopes, although the results shown are quite promising.

The use of the complete image from a sensor opens up new possibilities concerning the tomographic reconstruction of atmospheric profiles in different scenarios. Some recent studies have shown how it is possible to use CNN to compute the slopes in an SH-WFS for solar observation, or how to use that image to shape the DM to compensate the turbulence (Suárez Gómez et al. 2017a).

At this point, it is interesting to note the higher complexity of the CNN when compared with that of the previous MLPs. This convolutional approach implies an increment on the dimension due to the parameters to tune (number of filters, size of filters, size of pooling, number of layers, etc.) compared with the MLP. However, it is necessary to compare how good the reconstruction is from both methods, to decide if such effort in

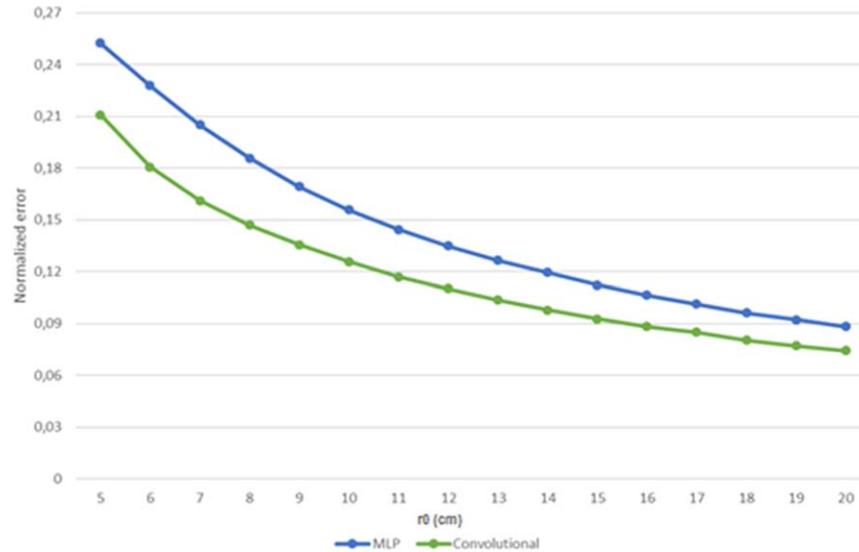


Figure 6. CARMEN MLP and Convolutional CARMEN performance in terms of normalized error for the test data with turbulence layer at 5000 meters. Reprinted by permissions from Springer Nature: Springer Nature eBook (Suárez Gómez et al. 2018). © 2018.

(A color version of this figure is available in the online journal.)

fitting a higher number of parameters is worthwhile. This work should be done in the near future, along with the training and recall speed of both ANNs.

Another approach using CNNs was also applied to a new curvature sensor instead of SH-WFS. This new sensor, called Tomographic Pupil Image (TPI-WFS; van Dam & Lane 2002) provides as input the defocused images of the phases to perform the reconstruction (Colodro-Conde et al. 2017). By using these images with a CNN, it is possible to obtain the Zernike polynomials of the distorted wavefront while substantially improving the quality of the reconstruction (Rodríguez Ramos et al. 2017). Different measures of the performance of the CNN over the original reconstructor are shown in (Sergio Luis, González-Gutiérrez, Alonso, Santos Rodríguez, Bonavera, et al. n.d.); a remarkable increase in terms of Strehl ratio is obtained, as is shown in Figure 7.

Another approach to MOAO reconstruction with neural networks is to study the temporal evolution of ANN (Suárez Gómez et al. 2017b). Simulated input and corrected data of the CARMEN reconstructor are treated together as time series to analyze the behavior of the reconstructor over time. One of the main conclusions is that although the correction is correlated with the evolution of time in the simulations, some information may be lost in the process. These results indicate that there is still room for improvement with other approaches, such as considering approaches that allow the use of more information for the reconstruction or recurrence patterns, to include the modelling of the temporal evolution.

The ANNs used in current research are not predictive and use no temporal information. However, the performance of the

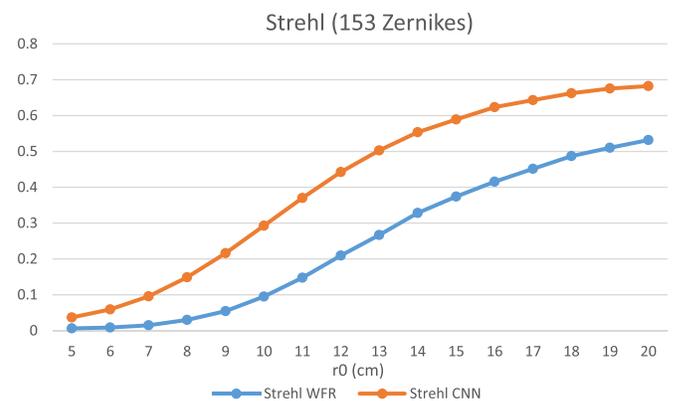


Figure 7. Strehl ratio of the difference between a recovered phase image and the reference one, measuring similarity for test data with the CNN reconstruction and WFR, the usual reconstruction for the TPI-WFS. Reprinted by permissions from Springer Nature: Springer Nature eBook (Suárez Gómez et al. 2018). © 2018.

(A color version of this figure is available in the online journal.)

reconstructor is expected to be enhanced by taking advantage of the temporal relation both in input and output (Suárez Gómez et al. 2017b). Future developments should use some architectures like a recurrent neural network such as long-short term memory (LSTM) and check if this new type of neural network provides any boost to its performance.

The evolution of the models allows us to take more information concerning the problems of AO into account; ANNs might be used to that end, for example regarding prediction, but to date there is no work in this direction.

7. Conclusions and Future Lines

Adaptive optics is a necessary technique for astronomical observation with ground-based telescopes. Moreover, with modern large telescopes a huge amount of data will need to be processed and the application of AI seems the optimal approach to solving this problem. AI applications in this field were already relatively common in the nineties and in the early years of the following decade. The studies that were carried out at that time suffered from limitations in the existent software. In this decade, many improvements have been made in terms of technology and computation, giving an increased importance to AI techniques in astronomy.

In particular, the most recent advances in the application of ANNs in MOAO began when remarkable results were obtained with both MARS models and MLP, when both were used in the modeling of DMs actuators. This, along with the possibility of accessing systems such as CANARY, led to the elaboration of competitive reconstruction techniques like CARMEN, the MLP reconstructor discussed here. Its basic characteristics and topology were determined. It was found that the MLPs were more reliable when used for long exposures, both in DM modelling and as a reconstructor. This was corroborated further when CARMEN was implemented on a real telescope

The obtained results allowed for further research, improving the application of these techniques to the increase of data in large telescopes like the E-ELT, where the usage of GPUs for calculations became an essential tool.

The success attained by MLPs in MOAO is currently leading to improvements with other approaches, such as using convolutional neural networks to deal with the complete image of the SH-WFS. This approximation provides more information to the ANN and has greater potential to improve the reconstruction of the aberrated wavefront. Furthermore, using this technique opens up new possibilities in daytime observations, as the challenges provided by an extended object observed with an SH-WFS are still far from being solved. The flexibility of ANNs suggest that they could be also used for new types of sensors like the TPI-WFS and opens up new possibilities for their use in different sensors, such as pyramidal or curvature; also, the approach of using them as tomographic reconstructors suggests that interesting results could be obtained if they were applied in other AO configurations apart from MOAO.

Machine-learning techniques are a great step forward in AO corrections. Many of these show improvements in speed and performance when compared to more classical reconstruction techniques. The evolution of the machine-learning techniques used meet the increasing necessity of searching for the real-time computation and the quality of the corrections. As an example, MARS techniques made corrections on Zernike coefficients, but the MLP methods made it possible to calculate corrections directly on centroids. As research evolved, tools

like CNNs allowed us to take images directly from the WFS, avoiding the usage of centroiding algorithms. In current applications like solar AO, more complex paradigms of neural networks became indispensable, making it possible to go from WFS images to DM corrections. Also, the research for implementations in GPUs speed up the calculations, making them competitive in regards to other reconstruction techniques.

By following the evolution in the use of machine-learning techniques in AO, the path to follow in current and future research is clearly defined, with more complex paradigms that permit image processing and make it possible to take into account other characteristics of the new challenges in AO, such as recurrent information or the lack of training samples.

The authors acknowledge Spanish ministry projects MINECO AYA2017-89121-P, MINECO AYA2015-65887-P and MINECO RYC-2013-13256.

ORCID iDs

Sergio Luis Suárez Gómez  <https://orcid.org/0000-0001-5260-4381>

References

- Abadi, M., Agarwal, A., Barham, P., et al. 2016, arXiv:1603.04467
- Al-Rfou, R., Alain, G., Almahairi, A., et al. 2016, arXiv:1605.02688
- Angel, J. R. P., Wizinowich, P., Lloyd-Hart, M., & Sandler, D. 1990, *Natur*, **348**, 221
- Basden, A., Geng, D., Myers, R., & Younger, E. 2010, *AOpt*, **49**, 6354
- Basden, A. G., Atkinson, D., Bharmal, N. A., et al. 2016, *MNRAS*, **459**, 1350
- Chundi, G. S., Lloyd-Hart, M., & Sundareshan, M. K. 2004, in IEEE Joint Conference on Neural Networks (Piscataway, NJ: IEEE), 2117
- Colodro-Conde, C., Velasco, S., Fernández-Valdivia, J. J., et al. 2017, *MNRAS*, **467**, 2855
- de Cos Juez, F. J., Lasheras, F. S., Roqueñí, N., & Osborn, J. 2012, *Sensors*, **12**, 8895
- Dipper, N., Basden, A., Bitenc, U., et al. 2013, in Proc. of the Third AO4ELT Conf. (*Firenze, Italy*) ed. S. Esposito & F. Luca, **41**
- Dutton, J. A. 1995, Dynamics of Atmospheric Motion: (formerly the Ceaseless Wind) (New York: Dover)
- Ellerbroek, B. L. 1994, *JOSAA*, **11**, 783
- Freeman, R. H., & Pearson, J. E. 1982, *ApOpt*, **21**, 580
- Friedman, J. H. 1991, *The Annals of Statistics*, **19**, 1
- Friedman, J. H., & Roosen, C. B. 1995, *Statistical Methods in Medical Research*, **4**, 197
- Gendron, E., Vidal, F., Brangier, M., et al. 2011, *A&A*, **529**, L2
- González-Gutiérrez, C., Sanchez-Rodríguez, M. L., Calvo-Rolle, J. L., & de Cos Juez, F. J. 2018, *Complexity*, **2018**, 5348265
- González-Gutiérrez, C., Santos-Rodríguez, J. D., Fernández Díaz, R. A., et al. 2016, in Advances in Intelligent Systems and Computing, Vol. 527 (Cham: Springer)
- González-Gutiérrez, C., Santos Rodríguez, J. D., Martínez-Zarzuela, M., et al. 2017, *Sensors*, **17**, 1263
- Graves, A., Mohamed, A.-r., & Hinton, G. 2013, in IEEE International Conference on Acoustics, Speech and Signal Processing (Piscataway, NJ: IEEE), 6645
- Guzmán, D., de Cos Juez, F. J., Lasheras, F. S., Myers, R., & Young, L. 2010, *OExpr*, **18**, 6492
- Guzmán, D., de Cos Juez, F. J., Myers, R., Guesalaga, A., & Lasheras, F. S. 2010, *OExpr*, **18**, 21356
- Guzman, D., Mello, A. T., Osborn, J., et al. 2012, *Proc. SPIE*, **8447**, 844740
- Jorgenson, M. B., & Aitken, G. J. M. 1992, *Proc. SPIE*, **1706**, 113

- Jorgenson, M. B., & Aitken, G. J. M. 1994, in *Active and adaptive optics: ESO Conference and Workshop Proceedings*, ed. F. Merkle (Garching: ESO), 134
- Krizhevsky, A. 2014, arXiv:1404.5997
- Krizhevsky, A., Sutskever, L., & Hinton, G. E. 2012, in *Advances in Neural Information Processing Systems*, 1097 (<https://dl.acm.org/citation.cfm?id=2999257>)
- Lamblin, P. 2017, MILA and the Future of Theano
- Lardi re, O., Andersen, D., Blain, C., et al. 2014, *Proc. SPIE*, 9148, 91481G
- Lloyd-Hart, M. 1994, Application of Machine Learning to Real-Time Removal of Atmospheric Blurring of Astronomical Images (<https://pdfs.semanticscholar.org/87be/fc3f22cc32d3f7f06780b6d120e43dc224a.pdf>)
- Marichal-Hern andez, J. G., Rodr guez-Ramos, L. F., Rosa, F., & Rodr guez-Ramos, J. M. 2005, *ApOpt*, 44, 7587
- McGuire, P. C., Rhoadarmer, T. A., Avriel Coy, H., Angel, J. R. P., & Lloyd-Hart, M. 2000, *Proc. SPIE*, 4007, 682
- McGuire, P. C., Sandler, D. G., Lloyd-Hart, M., & Rhoadarmer, T. A. 1999, *LNP*, 552, 97
- Mello, A. T., Kanaan, A., Guzman, D., & Guesalaga, A. 2014, *MNRAS*, 440, 2781
- Montera, D. A., Welsh, B. M., Roggemann, M. C., & Ruck, D. W. 1996, *ApOpt*, 35, 4238
- Montera, D. A., Welsh, B. M., Roggemann, M. C., & Ruck, D. W. 1997, *ApOpt*, 36, 675
- Noll, R. J. 1976, *JOSA*, 66, 207
- Osborn, J., de Cos Juez, F. J., Guzman, D., et al. 2011, in *Proc. of the Adaptive Optics for Extremely Large Telescopes*, 2
- Osborn, J., de Cos Juez, F. J., Guzman, D., et al. 2012, *OExpr*, 20, 2420
- Osborn, J., Guzman, D., de Cos Juez, F. J., et al. 2014, *Proc. SPIE*, 9148, 91484M
- Osborn, J., Guzman, D., De Cos Juez, F. J., et al. 2014, *MNRAS*, 441, 2508
- Paschall, R. N., & Anderson, D. J. 1993, *ApOpt*, 32, 6347
- Petit, C., Conan, J.-M., Kulcs r, C., Raynaud, H.-F., & Fusco, T. 2008, *OExpr*, 16, 87
- Platt, B. C., & Shack, R. 2001, *Journal of Refractive Surgery*, 17, S573
- Ramsay, S. K., Casali, M. M., Gonz lez, J. C., & Hubin, N. 2014, *Proc. SPIE*, 9147, 91471Z
- Rigaut, F. 2002, in *Beyond conventional adaptive optics : a conference devoted to the development of adaptive optics for extremely large telescopes*, ed. E. Vernet et al. (Garching: ESO), 11
- Roddier, F. 1999, *Adaptive Optics in Astronomy* (New York: Cambridge Univ. Press)
- Rodr guez Ramos, L. F., et al. 2017, in *Proc. of the Adaptive Optics for Extremely Large Telescopes 5*. Instituto de Astrof sica de Canarias (IAC), Retrieved (<http://www.iac.es/congreso/AO4ELT5/pages/proceeding.php?id=40>).
- Russell, S. J., & Norvig, P. 2016, *Artificial Intelligence: A Modern Approach* (Malaysia: Pearson Education Limited)
- Sandler, D. G., Barrett, T. K., Palmer, D. A., Fugate, R. Q., & Wild, W. J. 1991, *Natur*, 351, 300
- Sekulic, S., & Kowalski, B. R. 1992, *J. Chemom.*, 6, 199
- Sivo, G., Kulcs r, C., Conan, J., et al. 2014, *OExpr*, 22, 23565
- Su rez G mez, S. L., Gonz lez-Guti rrez, C., D ez Alonso, E., et al. 2018, *LNCs*, 10870, 74
- Su rez-G mez, S. L., Gonz lez-Guti rrez, C., S nchez-Lasheras, F., et al. 2017a, in *Proc. of Adaptive Optics for Extremely Large Telescopes* (Tenerife, Canary Islands, Spain), 5, (doi:10.26698/AO4ELT5.0033)
- Su rez G mez, S. L., Gonz lez Guti rrez, C., Santos Rodr guez, J. D., et al. 2016, in *Proc. of the 16th International Conference on Intelligent Systems Design and Applications*, ed. A. M. Madureira et al. (Cham: Springer), 1051 (doi:10.1007/978-3-319-53480-0_103)
- Su rez G mez, S. L., Santos Rodr guez, J. D., Iglesias Rodr guez, F. J., & de Cos Juez, F. J. 2017b, *Entropy*, 19, 103
- Tokovinin, A. 2004, *PASP*, 116, 941
- Tyson, R. 2010, *Principles of Adaptive Optics* (Boca Raton, FL: CRC Press)
- van Dam, M. A., & Lane, R. G. 2002, *ApOpt*, 41, 5497
- Vdovin, G. V. 1995, *OptEn*, 34, 3249
- Vidal, F., Gendron, E., & Rousset, G. 2010, *JOSAA*, 27, A253
- Wang, Y., Xu, H., Li, D., et al. 2018, *NatSR*, 8, 1124
- Wizinowich, P., Loyd-Hart, M., & Angel, R. 1991, in *Transputing'91*, Vol. 1 (Amsterdam: IOS Press)
- Wizinowich, P. L., Lloyd-Hart, M., McLeod, B. A., et al. 1991, *Proc. SPIE*, 1542, 148
- Zilberman, A., Golbraikh, E., & Kopeika, N. S. 2008, *ApOpt*, 47, 6385