An approach using deep learning for tomographic reconstruction in solar observation

S.L. Suárez-Gómez[a], C. González-Gutiérrez*[a], F. Sánchez-Lasheras[a], A.G. Basden[b], I. Montilla[c], M. Reyes[c], M. Collados-Vera[c], F.J. de Cos Juez[a]

[a]Mining Exploitation and Prospecting Department, University of Oviedo, 33004, Oviedo, Spain; [b]Dept. of Physics, Centre for Advanced Instrumentation, University of Durham, South Road, Durham DH1 3LE, UK; [c]Instituto de Astrofísica de Canarias, San Cristobal de La Laguna, Spain.

ABSTRACT

Solar observations with large telescopes have to face numerous challenges for its implementation. Adapting the current reconstructor systems of adaptive optics, which were developed for night observations to remove the atmospheric turbulence, is one of them. Neural networks were proved as a great solution in different fields, as image processing or in tomographic reconstruction. In this paper, possible uses of neural networks in solar adaptive optics will be explored, providing some early results of their application to different fields of adaptive optics.

Keywords: Adaptive Optics, Neural Networks, Solar

1. INTRODUCTION

10-meter class telescopes have been widely used during last decades to observe different phenomena that happens millions of kilometers away from our planet. Traditionally, these telescopes were used during night observations, and have been growing until achieving sizes of several tens of meters, like the European Extremely Large Telescope (E-ELT) [1]. However, the observation of the sun, especially in the optic spectrum, have gained a lot of interest during the last years [2]. The development of the European Solar Telescope (EST) [3], a 4 meters solar telescope has created new challenges regarding not only the engineering, but the control systems to handle this new generation of solar telescopes [4].

Adaptive Optics (AO) has been widely used in night observations, and it also provides great improvements when observing the sun with large telescopes [5]. The use of AO, implies the necessity of a reconstruction system to compensate the aberrations introduced by the atmosphere [6]. Machine learning techniques as SOM [7] or MARS [8] have been used in nocturnal observations with success [9], but the use of neural networks [10] has been proved to be a better solution [11].

Artificial Neural Networks (ANNs) have been successfully applied to several fields during last years, like image classification, speech recognition or language processing [12]–[15]. Their success on nocturnal observations [16], makes them great candidates to be applied also in solar observation. In the present paper, different proposals about the use of ANNs in solar telescopes will be detailed, such as their use in the calculation of the slopes, or their utilization as tomographic reconstructor.

Section 2, consists in a short introduction about solar adaptive optics and how a system can be simulated. In Section 3, ANNs will be detailed, and examples of how they work will be given. Section 4 is an explanation of how solar AO and neural networks could be combined. The last section will be the conclusions of the work and future lines.

*gonzalezgcarlos@uniovi.com;

2. SOLAR ADAPTIVE OPTICS

2.1 Adaptive Optics Systems

AO is one of the most common techniques to compensate the influence of the turbulence in stellar observations, during both day and night. It helps to improve the quality of the image received by ground-based telescopes through the use of real-time control systems. Typically, these systems are composed by a deformable mirror, a Wavefront Sensor (WFS) and a control system that includes a wavefront reconstructor [17]. A summary of the structure of an AO system can be observed in Figure 1.
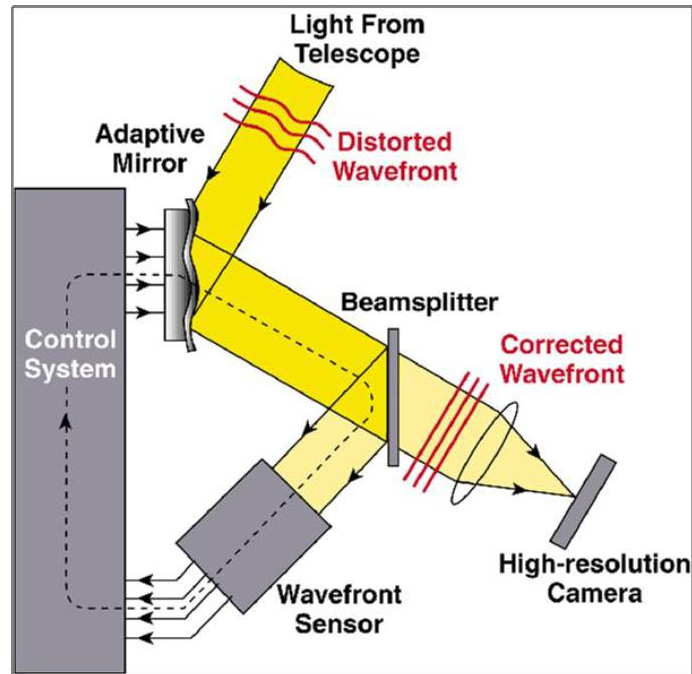

Figure 1: Adaptive Optics System

The Shack-Hartmann Wave-Front Sensor (SH-WFS) [18] is frequently used in AO systems. It is composed by a grid of lenses with the same focal length. The sensor has the ability to characterize the aberration of an incoming wavefront by dividing it into discrete sub-pupils. Once is divided, the deviation from the focal spot can be measured, and a vector of tilts which characterizes the aberration could be created.

For the adaptive mirror, it is common to use a deformable mirror that corrects phase aberrations. This can be achieved by changing the shape of the mirror with the correct wavefront shape. Usually, the mirror surface is deformed by several actuators located at the back of a thin reflecting surface.

In the case of the reconstruction and control system, it is possible to employ several different techniques. The most commonly used reconstructor is based on a matrix vector multiplication operation, with the control matrix being defined by either least squares [19] or minimum variance techniques [20]. During the last years, novelty methods have been applied with success, like Learn and Apply [21], or the use of Artificial Neural Network as tomographic reconstructor [11].

2.2 Difference between night and day observations

In its basic design, solar AO systems are quite similar to night-time AO systems. However, there are some significant differences between night and day observations. The atmospheric turbulence is the cause of the different aberrations produced in ground-based telescopes. In day time observations, due to heating of the ground produced by direct sunlight,

the ground-layer of the turbulence is much stronger than night observations [22]. Also, this heating make that Fried parameters fluctuates significantly on short time scales (seconds) and often drops to values of just a few centimeters [17].

Another big challenge in day observations, is the development of suitable wavefront sensors. Traditional WFS are not adequate since the Sun itself is used as reference, so a correlation SH-WFS is used instead. Laser guide stars are not useful when observing the sun, since it will be needed to be extremely bright lasers, making it a very expensive option.

2.3 Durham AO Simulator Platform (DASP)

DASP is an Open-Source simulator for adaptive optics. It was developed at the University of Durham and is available for download in [23]. It performs simulations from night and solar observations, providing information regarding the different WFS, centroiding, slopes, deformable mirror shapes, etc.

The simulator has a console terminal and a graphical interface to configure the adaptive optics system. It allows the modification of the different parameters that conform the turbulence profile (number and height of the layers, Fried parameter, wind speed, etc.), and the intensity of the different reference stars. To check the quality of the simulation, DASP also provides images in real-time from the performed simulations as it is shown in Figure 2(b), so researchers can check that simulations are being implemented correctly.
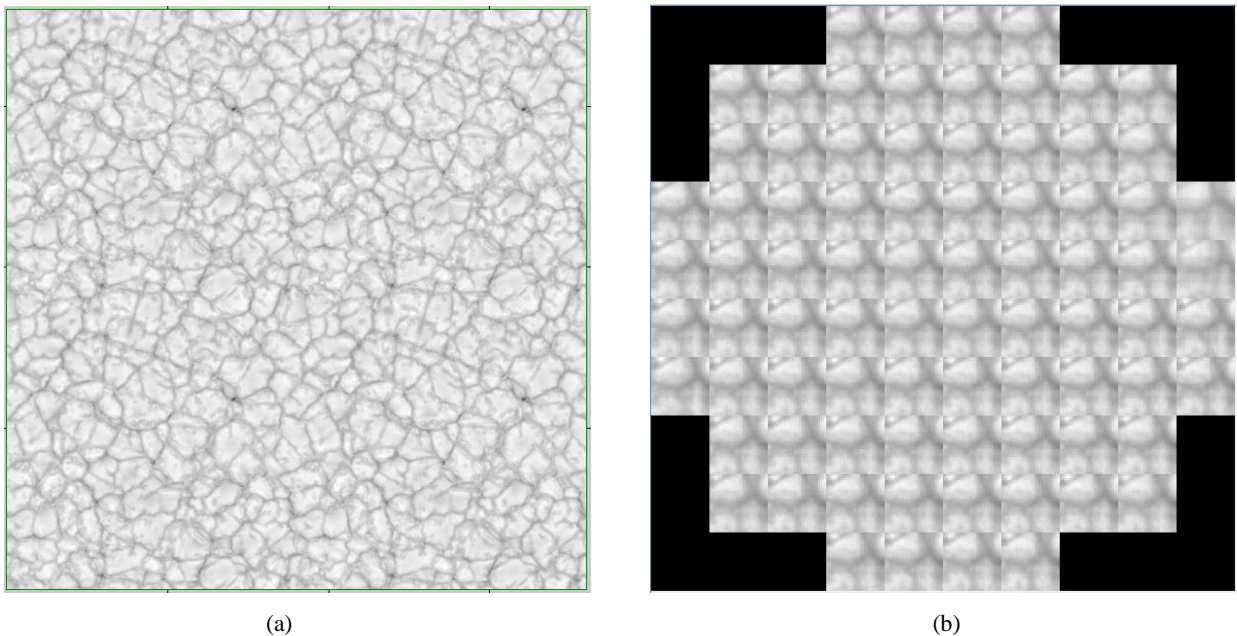

(a)   (b)

Figure 2: (a) Picture from a portion of the sun. (b) Solar observation in DASP

3.   CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks (CNNs) have been widely used processing images during the last years [24]. This method of Artificial Intelligence, has the ability to extract information from images, after a process of training the neural network with large sets of images.

3.1 Architecture of Convolutional Neural Networks

The composition of CNNs is structured in layers. These layers are composed by neurons, which perform an operation over the input. There are different types of layers, which are usually divided in two sections, as it is shown in Figure 3. In the feature learning side, the most common type is the convolutional layer, which is composed by different filters that are convoluted along the full entrance. After this operation, an activation function like the Rectified Linear Unit (ReLU) [25] is applied, to help during the training process [26]. The last step in a typical group of layers, is pooling, which select the most significant value (maximum, median or average), in a certain group of pixels. By doing this, it is possible to reduce

the size of an image, and extract the most significant features. Nesting several of these blocks provide a better extraction of relevant information.
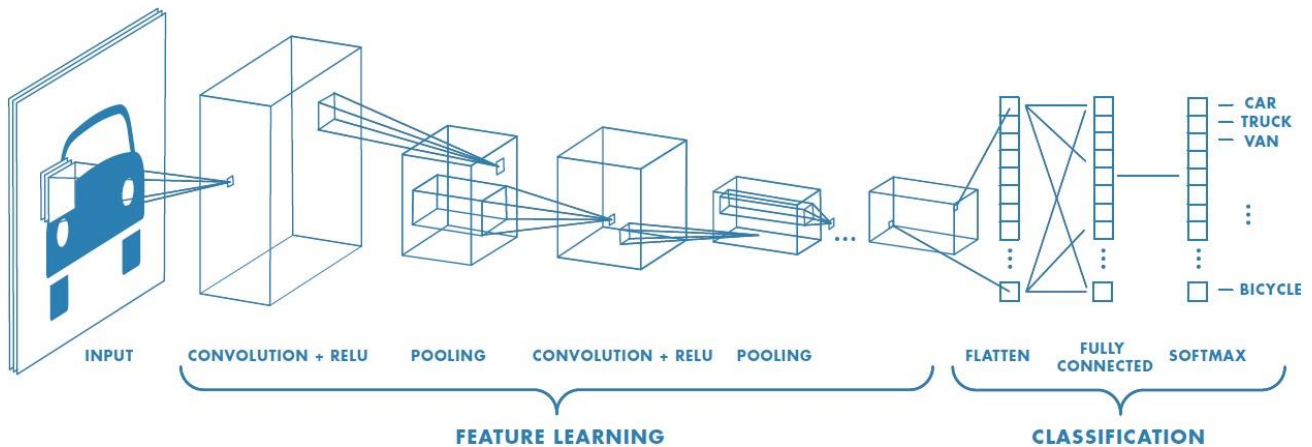

Figure 3: Example of an image detector using convolutional neural network

After the convolutional stage of the neural network, the different features are flattened and used as input to a Multi-Layer Perceptron (MLP). A MLP is composed of several fully connected layers, where all the neurons from one layer, are connected to all the neurons in the previous and the next layer. There is also an output layer that provides the expected output, which can be a category, real numbers, etc.

3.2 Training

One of the keys of neural networks is their capacity to learn and extract characteristics from sets of data. This process is known as training, and it allows the ANN not only to work with known data, but also to generalize with unknown data [27].

In a more practical definition, the process of training consists of calculating the optimal values of filters and weights of all neurons. This is an iterative process which requires sets of known data to compute these values. The first step is to randomly initialize the weights of all the neurons. Although it seems quite arbitrary, it has been proved that specific random initialization helps in the learning process [28]. Once the values are initialized, the network is fed with known sets of inputs and outputs. The result provided by the neural network is compared with the expected output, obtaining certain error. This error has to be minimized by using a gradient descent algorithm, and is going to be backpropagated through the different layers of the neural network [29]. After the propagation of the error, the values of the filters and the weights are updated. The previous steps are repeated until the complete dataset had trained the network. A summary of the training process it is shown in Figure 4
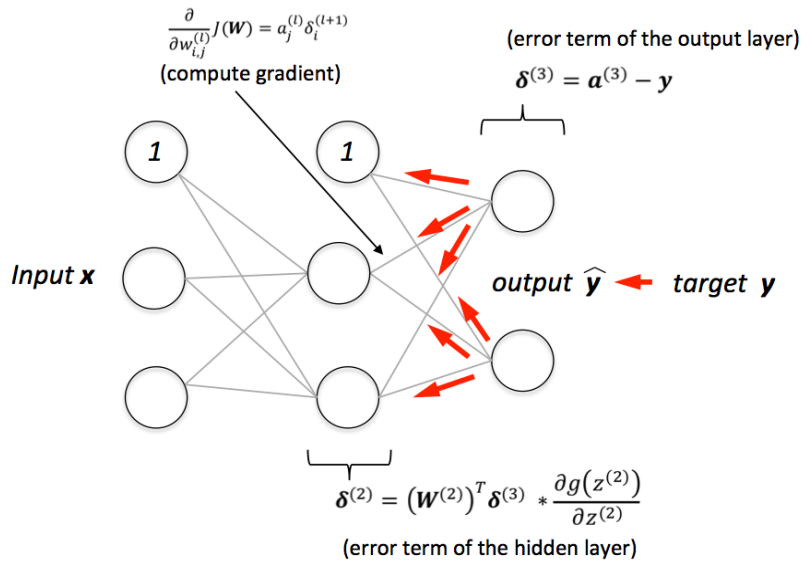
$$\frac{\partial}{\partial w_{i,j}^{(l)}} J(W) = a_j^{(l)} \delta_i^{(l+1)}$$
(compute gradient)

(error term of the output layer)
$$\delta^{(3)} = a^{(3)} - y$$

Input x

output $\widehat{y}$ ← **target** y

$$\delta^{(2)} = \left(W^{(2)}\right)^T \delta^{(3)} * \frac{\partial g\left(z^{(2)}\right)}{\partial z^{(2)}}$$
(error term of the hidden layer)

Figure 4: Neural network training

3.3 Neural networks in AO

Regarding night observations, there are already some experiments that mixed neural networks and AO. One of the most successful is the Complex Atmospheric Reconstructor based on Machine lEarNing (CARMEN) [30]. This reconstructor uses neural network to create a tomographic profile of the atmospheric turbulence. It uses a MLP trained with a single turbulence layer and fixed r0, but with variable height. This neural network has obtained great results during simulation compared to other reconstructors based on least squares methods or covariance matrixes [31]. CARMEN has also been tested on-sky, on a Nasmyth platform of the 4.2 m William Herschel Telescope, one of the Isaac Newton Group of Telescopes of the *Observatorio del Roque de los Muchachos*, La Palma, Canary Islands, Spain. During the tests, the reconstructor showed great performance, similar to the obtained during the simulation stage [16].

Another interesting idea of applying neural networks in adaptive optics, have been shown with the Adaptive Optics Lucky Imager (AOLI) [32]. This instrument uses a Tomographic Pupil Image Wave-Front Sensor (TPI-WFS) instead of a traditional SH-WFS, and provides two pictures, re-imaged at 2 mm from the pupil, as it can be seen in Figure 5.
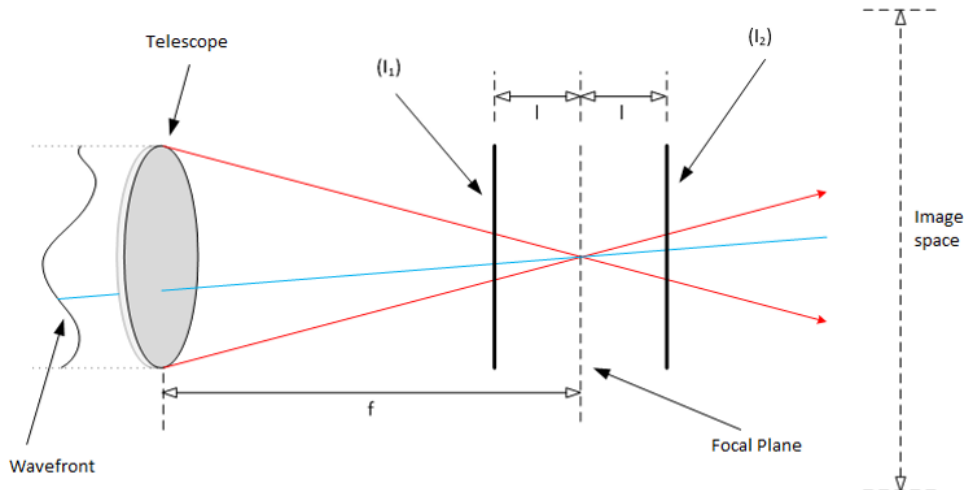
Figure 5: Tomographic Pupil Image Wave-Front Sensor

These two images, are used as inputs to the neural network, along with the polynomial Zernikes as outputs, which allows to reconstruct the aberrated wavefront. Using thousands of simulated images with variable r0, a CNN is trained. This neural network obtained a much better reconstruction than the traditional methods, especially for high values of r0 [33].

Also, the use of convolutional neural networks has a great potential solving the anisoplanatism issue with the correlation SH-WFS. Working on extended and low-contrast objects the Field Of View (FOV) has to be large enough to contain structure for the correlation algorithm to work robustly (~8-10 arcsec). With such FOV, the anisoplanatism affects the measurements of the correlating SH-WFS decreasing the sensitivity to wavefront distortions introduced at large heights above the telescope aperture [4]. CNNs have the potential to perform reconstruction using as reference images with smaller FOVs, alleviating this effect and increasing the performance.

4. PRELIMINARY RESULTS

After introducing different concepts regarding neural networks and solar adaptive optics, the proposed experiment will be detailed. The use of neural networks for solar observations in this work has two different purposes; first, the usage of solar images obtained from different WFS to estimate the slopes of the different sub-apertures in a SH-WFS. Also, to achieve the reconstruction of the turbulence profile, and compensate the aberration introduced by the atmosphere modifying a deformable mirror.

Both experiments share similar techniques regarding the training process. The starting point is a picture of a portion of the sun, as it is shown in Figure 2(a). This image is split in portions of 144x144 pixels and those pieces are used as inputs for the simulator. With DASP, different turbulences profiles are created, using one single turbulence layer with variable height, as it was done in CARMEN [30], and applied to the different portions of the image. After this process, it is possible to obtain inputs images as in Figure 2(b), which will feed the input of the neural network.

Thousands of images are generated, which are used to train the CNN. Also, to validate the result, a different set is created, using different turbulence profiles, which the goal to avoid overfitting and overtraining the neural network to the training set [34].

4.1 Slopes

To calculate the slopes of the different SH-WFS subapertures, DASP provide the perfect slopes from each subaperture. Using the input from Figure 2(b) and the perfect slopes as outputs, it is possible to train a neural network. Once is trained the CNN can predict the slopes using only the information presented in the image.

Different experiments have been conducted, obtaining and error of about 25% compared to the perfect value, when the outputs are normalized between [-1, 1]. Since the neural network, DASP and other techniques are still in early stages of their development, it has not been possible to compare different methods to compute the slopes with the neural network.

4.2 Deformable Mirror

In this scenario, the goal is to completely remove the atmospheric turbulence by using a tomographic reconstructor and a deformable mirror. The neural network is trained using the same input than the previous case but using the values of the actuators from the deformable mirror as output.

In the validation process, an absolute error of about 20% compared to the ideal value was obtained. This error is obtained when the values of the actuators are normalized between [-1, 1], and compared directly. In the future could be interesting to compute the error in optical terms, like the wavefront error, to compare the differences between the ideal and the predicted wavefront.

5. CONCLUSIONS AND FUTURE LINES

This paper has given some insights into solar AO and how CNNs works. Some examples of how AO have been combined with neural networks have been provided, showing how the application of neural networks to solar observations is a viable and promising option for their use in the future solar telescopes, especially in the future EST. Also, some results have shown, proving that there are different fields in solar AO where CNNs can be applied successfully.

However, it is important to remark that the project is in a very early stage, and it is needed to test the different experiments in a much deeper way. But there are lots of different ideas to improve the performance of CNNs that could be applied to solar AO, like the use of recurrent neural networks [35], on-line training [36], [37] or even apply the notion of classification when computing the outputs. For last, it is interesting to keep in mind that in astronomy, not only AO could be benefited for the usage of neural networks since, for instance, the detection of exoplanets [38], [39] has already done.

REFERENCES

[1]   S. K. Ramsay, M. M. Casali, J. C. González, and N. Hubin, "The E-ELT instrument roadmap: a status report," p. 91471Z, 2014.

[2]   M. Langlois, G. Moretto, K. Richards, S. Hegwer, and T. Rimmele, "Solar Multi - Conjugate Adaptive Optics at the Dunn Solar Telescope : Preliminary Results."

[3]   M. Collados *et al.*, "European Solar Telescope: Progress status," *Astron. Nachrichten*, vol. 331, no. 6, pp. 615–619, Jun. 2010.

[4]   I. Montilla, J. Marino, A. Asensio Ramos, M. Collados, L. Montoya, and M. Tallon, "Solar adaptive optics: specificities, lessons learned, and open alternatives," *Proc. SPIE*, vol. 9909, p. 99091H, 2016.

[5]   T. R. Rimmele, "Recent advances in solar adaptive optics," vol. 5490, pp. 34–47, Oct. 2004.

[6]   J. M. Beckers, "ADAPTIVE OPTICS FOR ASTRONOMY: Principles, Perfomance, and Applications," *Annu. Rev. Astron. Astrophys.*, vol. 31, pp. 13–62, 1993.

[7]   J. De Andrés, F. Sánchez-Lasheras, P. Lorca, and F. J. de Cos Juez, "A Hybrid Device of Self Organizing Maps (SOM) and Multivariate Adaptive Regression Splines (MARS) for the forecasting of firms' bankruptcy," *Account. Manag. Inf. Syst.*, vol. 10, no. 3, pp. 351–374, 2011.

[8]   J. R. Alonso Fernández, C. Díaz Muñiz, P. J. García Nieto, F. J. de Cos Juez, F. Sánchez Lasheras, and M. N. Roqueñí, "Forecasting the cyanotoxins presence in fresh waters: A new model based on genetic algorithms combined with the MARS technique," *Ecol. Eng.*, vol. 53, pp. 68–78, 2013.

[9]   D. Guzmán, F. J. de Cos Juez, F. Sánchez Lasheras, R. Myers, and L. Young, "Deformable mirror model for open-loop adaptive optics using multivariate adaptive regression splines," *Opt. Express*, vol. 18, no. 7, pp. 6492–6505, 2010.

[10]   C. González-Gutiérrez *et al.*, "Comparative Study of Neural Network Frameworks for the Next Generation of Adaptive Optics Systems," *Sensors*, vol. 17, no. 6, p. 1263, Jun. 2017.

[11]   J. Osborn *et al.*, "First on-sky results of a neural network based tomographic reconstructor: Carmen on Canary," in *SPIE Astronomical Telescopes + Instrumentation*, 2014, p. 91484M.

[12]   P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. Lecun, "OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks," *arXiv Prepr.*, 2014.

[13]   J.-L. Casteleiro-Roca, J. Calvo-Rolle, J. Méndez Pérez, N. Roqueñí Gutiérrez, and F. de Cos Juez, "Hybrid Intelligent System to Perform Fault Detection on BIS Sensor During Surgeries," *Sensors*, vol. 17, no. 1, p. 179, Jan. 2017.

[14]   F. J. de Cos Juez, F. Sánchez Lasheras, P. J. García Nieto, and A. Álvarez-Arenal, "Non-linear numerical analysis of a double-threaded titanium alloy dental implant by FEM," *Appl. Math. Comput.*, vol. 206, no. 2, pp. 952–967, Dec. 2008.

[15]   I. Sutskever, O. Vinyals, and Q. V Le, "Sequence to Sequence Learning with Neural Networks."

[16]   J. Osborn *et al.*, "Open-loop tomography with artificial neural networks on CANARY: On-sky results," *Mon. Not. R. Astron. Soc.*, vol. 441, no. 3, pp. 2508–2514, 2014.

[17]   T. R. Rimmele and J. Marino, "Solar Adaptive Optics.," *Living Rev. Sol. Phys.*, vol. 8, p. 2, 2011.

[18]   B. C. Platt and R. Shack, "History and Principles of Shack-Hartmann Wavefront Sensing," *J. Refract. Surg.*, vol. 17, no. 5, pp. S573–S577, 2001.

[19]   B. L. Ellerbroek, "First-order performance evaluation of adaptive-optics systems for atmospheric-turbulence compensation in extended-field-of-view astronomical telescopes," *J. Opt. Soc. Am. A*, vol. 11, no. 2, p. 783, Feb. 1994.

[20]   M. C. Roggemann, "Optical performance of fully and partially compensated adaptive optics systems using least-squares and minimum variance phase reconstructors," *Comput. Electr. Eng.*, vol. 18, no. 6, pp. 451–466, Nov. 1992.

[21]   F. Vidal, E. Gendron, and G. Rousset, "Tomography approach for multi-object adaptive optics," *J. Opt. Soc. Am. A. Opt. Image Sci. Vis.*, vol. 27, no. 11, pp. A253–A264, Nov. 2010.

[22]   C. Turrado, M. López, F. Lasheras, B. Gómez, J. Rollé, and F. Juez, "Missing Data Imputation of Solar Radiation Data under Different Atmospheric Conditions," *Sensors*, vol. 14, no. 11, pp. 20382–20399, Oct. 2014.

[23]   A. G. Basden, "Durham AO Simulation Platform." [Online]. Available: https://gitlab.com/agb32/dasp. [Accessed: 14-Sep-2017].

[24]   Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[25]   V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, pp. 807–814.

[26]   A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[27]   M. Nielsen, "Neural Networks and Deep Learning," 2016. [Online]. Available: http://neuralnetworksanddeeplearning.com/index.html. [Accessed: 26-Aug-2016].

[28]   X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 2010, vol. 9, pp. 249–256.

[29]   J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61. pp. 85–117, 2015.

[30]   F. J. de Cos Juez, F. Sánchez Lasheras, N. Roqueñí, and J. Osborn, "An ANN-based smart tomographic reconstructor in a dynamic environment.," *Sensors (Basel).*, vol. 12, no. 7, pp. 8895–911, 2012.

[31]   J. Osborn *et al.*, "Using artificial neural networks for open-loop tomography," *Opt. Express*, vol. 20, no. 3, pp. 2420–2434, 2012.

[32]   S. Velasco *et al.*, "Adaptive Optics and Lucky Imager (AOLI): presentation and first light," in *Proceedings of the XI Scientific Meeting of the Spanish Astronomical Society*, 2014.

[33]   C. González-Gutiérrez, J. J. Fernández-Valdivia, S. L. Suárez-Gómez, J. M. Rodríguez-Ramos, L. F. Rodríguez-Ramos, and F. J. de Cos Juez, "New adaptive optics Tomographic Pupil Image reconstructor based on convolutional neural networks." Puerto de la Cruz, 2017.

[34]   I. V Tetko, D. J. Livingstone, and A. I. Luik, "Neural Network Studies. 1. Comparison of Overfitting and Overtraining," *J. Chem. Inf. Comput. Sci*, vol. 35, pp. 826–833, 1995.

[35]   S. Hochreiter, S. Hochreiter, J. Schmidhuber, and J. Schmidhuber, "Long short-term memory.," *Neural Comput.*, vol. 9, no. 8, pp. 1735–80, Nov. 1997.

[36]   S. L. Suárez Gómez, J. D. Santos Rodríguez, F. J. Iglesias Rodríguez, and F. J. de Cos Juez, "Analysis of the Temporal Structure Evolution of Physical Systems with the Self-Organising Tree Algorithm (SOTA): Application for Validating Neural Network Systems on Adaptive Optics Data before On-Sky Implementation," *Entropy*, vol. 19, no. 3, p. 103, Mar. 2017.

[37]   S. L. Suárez-Gómez, M. L. Sánchez, F. Blanco, J. Ayala, and F. J. De Cos Juez, "Successful sulfur recovery in low sulfurate compounds obtained from the zinc industry: Evaporation–condensation method," *J. Hazard. Mater.*, vol. 336, pp. 168–173, 2017.

[38]   K. A. Pearson, L. Palafox, and C. A. Griffith, "Searching for Exoplanets Using Artificial Intelligence," Jun. 2017.

[39]   P. Blanco-Rodríguez, L. Fernández-Serantes, A. Otero-Pazos, J. Calvo-Rolle, and F. de Cos Juez, "Radon Mitigation Approach in a Laboratory Measurement Room," *Sensors*, vol. 17, no. 5, p. 1090, May 2017.