

RECONSTRUCCION TOMOGRAFICA DE LA TURBULENCIA ATMOSFERICA POR MEDIO DE TECNICAS DE MINERIA DE DATOS

Francisco Javier De Cos Juez¹

Fernando Rodríguez Perez¹

Jesús Laine Cuervo¹

Valeriano Alvarez Cabal¹

¹Área de Proyectos de Ingeniería. Universidad de Oviedo

Abstract

Adaptive optics for Astronomy has as its main aim the correction of the optical aberrations produced by atmospheric turbulences, thus increasing significantly the angular resolution that can be reached by a telescope.

The new optic technologies perform the correction of atmospheric perturbations by means of deformable mirrors placed on conjugated planes to the correspondent turbulent atmospheric layers. By means of computer tomography techniques, it is possible to measure the turbulence in different directions, using for this purpose several references inside the vision field. This is the way of reaching a 3D reconstruction

In order to simplify the process, the latest researches consider the tomographic reconstruction as a lineal process, assuming a turbulence distribution by layers. The atmosphere does not always behave in this way and the present paper proposes the study of the efficiency of rebuilding the turbulence using data mining and artificial intelligence techniques that allows the modeling of turbulence phenomena in a more realistic way without being constrained to the lineal cases under research in this field up to now. This goal will be achieved through simulations.

Keywords: *adaptive optics; atmospheric turbulence; data mining; artificial intelligence*

Resumen

Gran parte de los avances científicos se deben al estudio del espacio y los cuerpos celestes. La mayor parte de estos estudios se realizan en los grandes telescopios terrestres. Desafortunadamente la información recibida por estos está condicionada por la atmósfera que amortigua y deforma las observaciones. La óptica adaptativa astronómica tiene como objetivo corregir las aberraciones ópticas producidas por la turbulencia atmosférica, aumentando significativamente la resolución angular que un telescopio puede alcanzar.

Los métodos ópticos de nueva generación realizan la corrección de la perturbación atmosférica con espejos deformables situados en los planos conjugados a las capas atmosféricas turbulentas correspondientes. Mediante técnicas de tomografía se mide la turbulencia en distintas direcciones utilizando para ello varias referencias dentro del campo de visión. De esta manera se consigue una reconstrucción 3D y utilizando

varios espejos deformables se realiza una compensación tridimensional (por capas) de la perturbación

Con objeto de simplificar el proceso, las últimas investigaciones consideran que la reconstrucción tomográfica es un proceso lineal, asumiendo una distribución de turbulencia en capas. Sin embargo, la atmósfera no siempre se comporta de esta manera, por lo que el presente trabajo muestra la efectividad de reconstruir la turbulencia utilizando técnicas de minería de datos e inteligencia artificial.

Palabras clave: *óptica adaptativa; turbulencia atmosférica; minería de datos; inteligencia artificial*

1 Introducción

Un telescopio es básicamente un sistema óptico que amplifica el tamaño aparente de un objeto distante controlando el frente de onda de la luz incidente mediante una serie de lentes y espejos.

Antes de alcanzar la lente del telescopio, la luz ha de atravesar el espacio vacío y penetrar en la atmósfera. Es durante este paso por la atmósfera, donde por efecto de los iones, turbulencias y desplazamientos de masa a diferentes alturas, la onda sufre una serie de distorsiones de diferente magnitud, que modifican la morfología del frente de onda y dificultan su correcta observación en el telescopio a nivel terrestre. Para alcanzar una calidad de imagen aceptable, es necesario corregir dichas desviaciones.

Para ello, en primer lugar es necesario caracterizar las mencionadas desviaciones. En Óptica Adaptativa (OA) normalmente se utiliza a tal efecto un diseño especial de sensor de frente de onda denominado Shack Hartmann, que determina la desviación del frente respecto de un frente plano (Figura 1). Con los datos obtenidos por estos sensores, es posible conocer la distorsión aplicando alguna técnica de reconstrucción existente en OA, para más adelante modificar la superficie de un espejo deformable de tal manera que se compense la aberración y el frente de onda resultante sea lo más plano posible. (Wilson y Jenkins, 1996).

En estas técnicas es necesario contar con una fuente de luz en las proximidades del objetivo a medir con la intensidad suficiente para permitir la determinación de la desviación del frente respecto un frente plano. Estas fuentes de luz se denominan estrellas guía y pueden ser naturales o artificialmente generadas con un laser de alta potencia. La proximidad entre el objetivo y la estrella guía determina el ratio de solapamiento entre sus conos de luz, lo que es un indicador de qué grado de distorsión comparten, es decir, cómo de válidas serán las correcciones efectuadas sobre el objetivo.

Existen en la actualidad múltiples configuraciones ópticas destinadas a corregir las aberraciones introducidas por la atmosfera, siendo una de las más prometedoras la denominada Multi Object Adaptive Optics (MOAO) (Assemat, Gendron & Hammer, 2007; Gendron et al., 2011). En MOAO, se miden las desviaciones de múltiples estrellas guía distribuidas por todo el campo de observación mediante un sistema de sensores Shack Hartmann carente de conexión directa con un espejo (ciclo abierto). Con dicha información, se reconstruye la aberración total en la dirección de cada objetivo a medir y la información se envía a un espejo deformable. (Morris et al. 2009; Gendron et al. 2011)

En este artículo se expondrá el desarrollo de un nuevo sistema de reconstrucción de la distorsión de la onda basado en redes neuronales de tipo perceptron multicapa, utilizando los datos obtenidos por una serie de sensores Shack Hartmann en MOAO. Demostraremos la competitividad de este nuevo sistema comparándolo con los

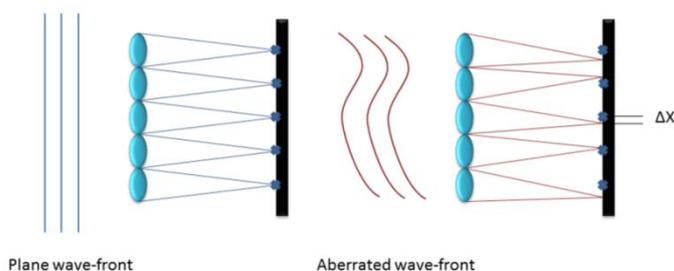
resultados obtenidos por dos técnicas de reconstrucción que actualmente se encuentran implantadas o en desarrollo.

2 Metodologías y materiales

2.1 Sensor de frente de onda Shack Hartmann

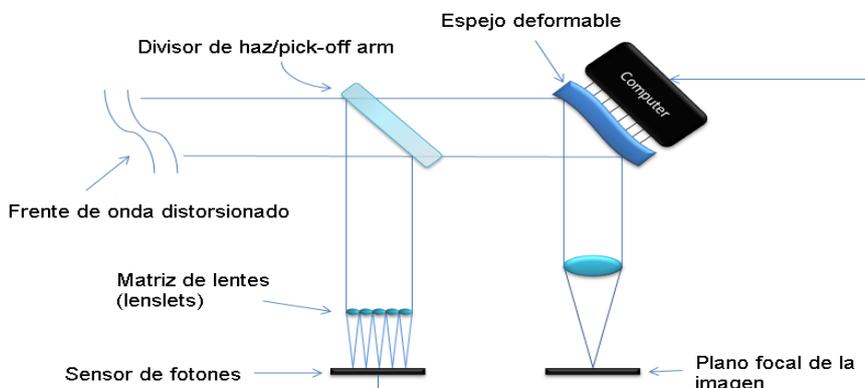
Comúnmente utilizados en astronomía para caracterizar frentes de onda, estos sensores consisten en una matriz de lentes con la misma distancia focal (lenslets), enfocadas a un sensor de fotones. Cuando el frente de onda incide en la matriz de lentes, se divide en pequeños rayos que se proyectan sobre el sensor de fotones. Si el frente de onda es plano, los rayos inciden en el centro geométrico del área designada para cada una de las lentes. Si por el contrario el frente de onda se encuentra deformado, el punto focal de las lentes varía respecto del centro del área y es posible medir dicha desviación tal y como muestra la Figura 1, conformando una matriz de desviaciones característica de la aberración del frente de onda (Platt y Shack, 2001). Con estos datos, la distorsión inducida por la turbulencia atmosférica puede ser reconstruida en forma de ecuación matemática (generalmente como polinomios de Zernike).

Figura 1: Medida de los desplazamientos del frente de onda.



En la Figura 2 puede verse un esquema simplificado de un sistema MOAO. El frente de onda distorsionado penetra en el telescopio y se separa en dos flujos. Uno de ellos es utilizado para medir la desviación del frente mediante la acción de los sensores Shack Hartman y deformar el espejo para corregir el segundo flujo y formar la imagen final para observación (Platt y Shack 2001).

Figura 2: Sistema de optica adaptativa en ciclo abierto



Tal y como se ha mencionado antes, los sistemas MOAO utilizan una serie de sensores de frente de onda cuyos campos de observación están solapados con el del

objetivo a medir, de forma que es posible obtener información de la distorsión a lo largo de todo el cono de luz del objetivo. Esa información es procesada para reconstruir la aberración de la onda mediante alguna de las técnicas de reconstrucción y se utiliza para deformar un espejo que compense la onda incidente para obtener un frente de onda más plano. (Assemat et al., 2007; Gendron et al., 2011).

2.2 Técnicas de reconstrucción existentes basadas en Sensores de frente de onda Shack Hartmann

Existen varias formas de reconstruir el frente de onda a partir de la información de los sensores Shack Hartmann. Compararemos nuestro sistema de reconstrucción con 2 tipos de técnicas diferentes: Standard least squares type matrix vector multiplication (LS) (Ellerbroek, 1994) y Learn and apply (L+A) (Vidal, Gendron & Rousset, 2010).

Learn and apply es una técnica de reconstrucción en 2 pasos que consiste en calcular la matriz de covarianza que combina las desviaciones medidas en los sensores de frente de onda fuera del eje entre sí y con otro sensor montado sobre el eje del telescopio. De esta forma tanto el perfil de turbulencias (“fuerza en función de la altitud”) como la posición de las estrellas guía son tenidas en cuenta a la hora de la reconstrucción. El mayor inconveniente que presenta esta técnica es la necesidad de recalcularse las matrices de covarianza cuando el perfil atmosférico sufre cambios significativos, ya que durante ese periodo es imposible realizar la medición. Es posible lidiar con este problema estimando la matriz covarianzas del sensor en el eje en base a información geométrica, permitiendo que el sistema funcione durante transitorios turbulentos, pero con un rendimiento menor.

Standar Least Squares type Matrix vector multiplication (Ellerbroek, 1994; Fusco et al., 2001) es el método estándar para la reconstrucción tomográfica. Consiste en una matriz de control que cuando se multiplica por los valores obtenidos en los sensores Shack Hartman, convierte las desviaciones medidas en voltaje para los actuadores del espejo. Puede calcularse offline pero requiere mucha capacidad computacional, lo que representa un problema para las futuras generaciones de telescopios (extra grandes) debido al gran número de sensores y de espejos deformables que presentan.

2.3 Métricas de rendimiento óptico

Se han utilizado dos métricas diferentes para cuantificar el rendimiento óptico de cada sistema evaluado

- Error cuadrático medio del frente de onda (EFO, [nm])
- Strehl ratio de la función de dispersión puntual (Point Spread function o PSF)

El error cuadrático medio se define como la diferencia entre la media de los cuadrados de las desviaciones respecto al valor real y el cuadrado de la media de dichas desviaciones. Mide la variación media del error a lo largo de toda la esfera de referencia y como tal es un indicador de la calidad de la imagen. (Wyant, 1992),

$$RMS = \sqrt{\overline{W^2} - \bar{W}^2}, \quad (1)$$

Donde W es la desviación del frente de onda.

El Strehl ratio de la función de dispersión es una medida la calidad óptica de los telescopios y otros sistemas ópticos. Se define como el ratio entre la intensidad pico medida en el plano de detección y la intensidad máxima teórica del punto fuente de un instrumento trabajando en el límite de difracción (carente de aberraciones) (Wyant, 1992).

Para pequeñas desviaciones se puede expresar como

$$S \equiv e^{-(2\pi\sigma/\lambda)}, \quad (2)$$

Donde λ es la longitud de onda y σ el error cuadrático medio del frente de onda. Varía entre 0 y 1 y es directamente proporcional a la calidad de la imagen (Wyant, 1992).

2.4 Red neuronal de perceptron multicapa

Las Redes Neuronales Artificiales (RNA) son modelos computacionales inspirados en las redes neuronales biológicas, y que consisten en una serie de elementos de procesamiento simple denominados neuronas o nodos conectados entre sí. (Sanz Molina y Martín del Brío, 2006)

El Perceptrón Multicapa es un tipo específico de red neuronal sin realimentación. Los nodos se encuentran organizados en capas (capa de entrada, capa oculta y capa de salida) y cada neurona se conecta con una o varias neuronas de cualquiera de las capas siguiente. Existe un tipo especial de neurona llamada "bias", que no es alimentada por ninguna neurona de las capas anteriores. Se utiliza para modificar el término independiente de la función de activación en cada neurona y mejorar la adaptabilidad del conjunto de la red.

Cada neurona recibe una serie de datos de neuronas de capas anteriores o de una fuente externa, los transforma localmente utilizando una función, denominada función de activación, y envía el resultado a una o más neuronas de las capas siguientes. Este ciclo se repite hasta que se alcanzan las neuronas de salida. Cada conexión entre nodos tiene un valor numérico que representa la importancia de la neurona predecesora en el resultado de la actual, llamado peso sináptico o simplemente "peso". Es en estos valores donde se almacena la fracción más importante de conocimiento de la red. (Haykin, 1999)

Matemáticamente una neurona de una red puede modelarse como

$$Z_j(t) = g(S_j) = g\left(\sum_{i=1}^n W_{ji} \cdot Z_i(t-1) + W_j\right) \quad (3)$$

donde Z_j es la salida de la neurona j , W_{ji} el peso sináptico entre la neurona j y la neurona i , t la capa en la que se encuentra la neurona Z_j , Z_i la salida de la neurona i de la capa $(t-1)$ y W_j el bias. La función $g(\cdot)$ es la llamada función de activación.

En el aprendizaje supervisado, el entrenamiento se efectúa introduciendo a la red un conjunto de datos de entrada y salidas esperadas y modificando los pesos para adaptarlos a la estructura de la función inherente a dicha serie de datos. La forma en la que los pesos varían para amoldarse a los resultados se denomina "algoritmo de aprendizaje" y es una de las características más importantes para el rendimiento de una red. (Ghahramani, 2004)

2.5 Métricas de rendimiento de la red neuronal

Para medir el rendimiento de la red se utilizan 3 métricas diferentes: Raíz del error cuadrático medio, error normalizado y precisión (accuracy)

La raíz del error cuadrático medio se utiliza en regresiones y operaciones similares, mide cuánta desviación presentan los valores de salida respecto a los valores objetivo

$$RECM = \sqrt{\frac{1}{N} \sum_{j=1}^N (y_j - d_j)^2} \quad (4)$$

El problema con el error cuadrático es que se ve influenciado por la varianza de los valores objetivo y no permite la comparación con otras redes que trabajen en situaciones diferentes.

El error normalizado es una métrica que elimina el efecto de las varianzas del objetivo y resulta independiente de la configuración de la red, devolviendo valores entre 0 y 1 (Pineda, 1988). Puede verse como una medida de varianza del resultado que es provocada por el error en sí mismo y no por la varianza de los valores objetivo o la arquitectura de la red (Eberhart y Dobbins, 1990). Para calcular el error normalizado es necesario calcular la suma de las desviaciones cuadráticas de la media.

$$E_{medio} = \sum_{j=1}^N \sum_{i=1}^M (d_{ji} - \mu_j)^2 \quad (5)$$

Donde N es el número de salidas y M el número de casos. El error total cuadrático de la red es

$$E_t = \sum_{i=1}^M \sum_{j=1}^N (y_j - d_j)_i^2 \quad (6)$$

Por lo que el error normalizado puede calcularse como

$$E_n = E_t / E_{medio} \quad (7)$$

Valores cercanos a 0 indican que el patrón se está aprendiendo correctamente. Por el contrario, un valor cercano a 1 indica que la red está entregando la media como valor de salida para todos los conjuntos de entradas. Las redes con el algoritmo de entrenamiento de retropropagación del error tienden a aprender este patrón de medias con relativa facilidad, por lo que esta métrica es particularmente útil para este tipo de redes (Eberhart y Dobbins, 1990). Puede aplicarse a la red de forma global o individualmente a cada salida.

Se define la precisión (accuracy) como la proporción de valores predichos correctamente con respecto al tamaño del set de datos. Se utiliza normalmente en redes neuronales de salida discontinua, ya que es fácil determinar cuándo un resultado es cierto o falso. En el caso de salidas continuas, es necesario definir un valor umbral del error, a partir del cual se considere un valor como correcto. Para este trabajo se ha utilizado un valor umbral del 5% del intervalo completo (diferencia entre el valor máximo y el mínimo) para cada salida. (Caruana & Niculescu-Mizil 2004)

3 Parámetros de la simulación

El reconstructor ha sido diseñado para devolver los coeficientes del polinomio Zernike para los 6 primeros órdenes (si incluir el piston) utilizando las desviaciones medidas en las 7x7 subaperturas de tres sensores Shack Hartmann.

La red neuronal seleccionada es un perceptron multicapa con retropropagación del error como algoritmo de entrenamiento. Se han hecho pruebas con varias morfologías de red: una o dos capas ocultas con diferente número de neuronas, learning rates y funciones de activación. Problemas más complejos requieren en general arquitecturas más complicadas. En nuestro caso la complejidad aumenta en función del número de capas turbulentas, su altura y la diferencia entre sus propiedades

Para el entrenamiento se han utilizado datos de simulaciones, de forma que es posible tener el control total de las parejas entrada-salida a alimentar y evitar así valores desviados, datos perdidos y otras problemáticas de la medición. El escenario más realista sería simular varias capas turbulentas a diferentes alturas y con diferentes fuerzas relativas. Otra posibilidad es simplificar este esquema simulando una capa

dominante a nivel de superficie con una segunda capa que va variando su altitud. (Le Louarn et al., 2000; Hubin et al., 2005; Osborn et al., 2011)

Se han realizado pruebas con muchos escenarios diferentes de entrenamiento incluidos aquellos que se citan anteriormente, con diferentes arquitecturas y parámetros de red. Finalmente, el mejor rendimiento se obtuvo mediante la combinación de simular una capa turbulencia que varía su cota, comenzando desde la superficie hasta los 15500 m con un paso de 150 m y utilizar los parámetros de red recogidos en la Tabla 1. Se generaron 1000 parejas aleatorias para cada altitud, de forma que finalmente se obtuvo un set de entrenamiento de 155000 parejas, que incluyen todas las posiciones posibles de la capa en la atmosfera.

La **Tabla 1** muestra un resumen de los parámetros utilizados para el entrenamiento de la red

Tabla 1: Resumen de los parámetros de la red neuronal

Parámetro	Valor
Red neuronal	Perceptrón multicapa
Número de capas ocultas	1
Neuronas	222(entrada)-222(oculta)-27 (salida)
Function de activación	Sigmoide continua
Algoritmo de entrenamiento	Retropropagación del error
Learning rate	0.01
Epochs	10000

Debido al hecho de que los pesos iniciales son aleatorios, todas las redes neuronales presentan diferentes valores de salida para la misma entrada incluso aunque fueran entrenadas utilizando los mismos parámetros y datos. Con el objetivo de disminuir la varianza y mejorar la precisión de la red, se entrenaron 7 redes diferentes con los mismos parámetros y datos, de forma que a la hora de utilización se promedian los resultados obtenidos. (Hashem y Schmeiser, 1995; Hashem, 1997; Naftaly, Intrator & Horn, 1997).

Todas las simulaciones se han realizado asumiendo tres estrellas guía naturales fuera del eje situadas a una distancia angular de 30 arcosegundos en un telescopio de 4.2 m. La posición del objetivo es en el centro del campo de visión y todos los sensores de frente de onda son Shack Hartmann con 7x7 subaberturas, 100 fotones y 20x20 pixels por subapertura. Estas condiciones son similares a las del experimento CANARY (Gendron et al, 2011) para posibilitar la comparación de resultados con datos de valores on-sky del reconstructor Learn and Apply.

A la hora de realizar las pruebas, se utilizó una simulación de Montecarlo para generar tres perfiles diferentes: buenas condiciones, condiciones normales y malas condiciones atmosféricas, extraídos del anteriormente mencionado experimento CANARY en La Palma. Todas tienen 4 capas turbulentas pero las alturas y su intensidad varían en cada uno. Las otras técnicas mencionadas fueron también sometidas a estos perfiles para comprobar su rendimiento. Tal y como se mencionada anteriormente, estas técnicas requieren una recalibración cuando varía el perfil atmosférico si se desea un buen rendimiento. Ambos sistemas fueron reconfigurados después de cada prueba para poder afrontar la siguiente con su mejor rendimiento, mientras que la red neuronal permaneció inalterada durante todas las pruebas. La Tabla 2 muestra los parámetros de las capas turbulentas en cada caso.

Tabla 2: Parámetros de las capas turbulentas para los tres casos de estudio

Capa	Parámetro	Valores			Unidades
Común	Test	test1	test2	test3	
	r_0	0,16	0,12	0,085	m
Capa 1	Altitud	0	0	0	m
	Intensidad relativa	0,65	0,45	0,8	
	Velocidad del viento	7,5	7,5	10	m/s
	Dirección del viento	0	0	0	grados
Capa 2	Altitud	4000	2500	6500	m
	Intensidad relativa	0,15	0,15	0,05	
	Velocidad del viento	12,5	12,5	15	m/s
	Dirección del viento	330	330	330	grados
Capa 3	Altitud	10000	4000	10000	m
	Intensidad relativa	0,1	0,3	0,1	
	Velocidad del viento	15	15	17,5	m/s
	Dirección del viento	135	135	135	grados
Capa 4	Altitud	15500	13500	15500	m
	Intensidad relativa	0,1	0,1	0,05	
	Velocidad del viento	20	20	25	m/s
	Dirección del viento	240	240	240	grados

4 Resultados

4.1 Métricas de rendimiento de la red neuronal

La Tabla 3 muestra las medidas de rendimiento realizadas a la red neuronal con los diferentes casos de test, respectivamente desde la neurona 1 hasta la 5 (coeficientes de Zernike)

Tabla 3: Métricas del rendimiento de la red para los tests 1, 2 y 3

Test	Métrica	Coef. 1	Coef. 2	Coef. 3	Coef. 4	Coef. 5
Test 1	RECM	0,8976	0,8464	0,6917	0,6159	0,6303
	Error normalizado	0,0374	0,0345	0,1007	0,0844	0,0765
	Precisión	94,8	97,13	77,22	80,1	83,63
Test 2	RECM	1,0445	1,0387	0,7746	0,6891	0,7121
	Error normalizado	0,0314	0,0327	0,0773	0,0661	0,0614
	Precisión	96,49	95,44	84,52	84,9	86,99
Test 3	RECM	1,0941	1,0902	1,0082	0,8701	0,9312
	Error normalizado	0,0195	0,0200	0,0743	0,0589	0,0586
	Precisión	99,29	99,46	85,94	85,91	89,7

Tal y como se esperaba, la raíz del error cuadrático medio aumenta desde el test 1 hasta el 3 debido a la mayor complejidad de estos últimos. Dado que el test 3 es el de

peores condiciones atmosféricas, su varianza y su intervalo es mayor que en los otros dos casos. Esta mayor variabilidad provoca que el error normalizado es menor que en los tests 1 y 2 dada la mayor diferencia cuadrática que existe con la media. Esta mejora de rendimiento del test 3 sugiere que la red rinde más cuanto mayor sea la complejidad de la situación atmosférica pese a que debido a dicha complejidad el error total continúa siendo mayor.

Existen también diferencias según el orden del zernike que se evalúe. Se puede observar que la raíz del error cuadrático disminuye al aumentar el orden, mientras que el error normalizado y la precisión aumentan debido a una mayor varianza e intervalo.

4.2 Métricas tomográficas

Learn and Apply y LS también fueron sometidos a las mismas pruebas para poder evaluar las diferencias de rendimiento entre las técnicas. Los resultados se muestran en la Tabla 4

Se puede observar claramente que la red neuronal es capaz de corregir los 3 perfiles con un rendimiento ligeramente superior a las otras técnicas, incluso sin proporcionarle ninguna información adicional entre las diferentes pruebas. En una situación real, donde el perfil cambia con el tiempo de una forma desconocida y a una velocidad sin determinar, las otras técnicas de reconstrucción pueden no ser capaces de manejar estos cambios de la misma forma que lo hace la red neuronal, ya que precisan de una recalibración.

Tabla 4: Resultados de los tests con las 3 técnicas de reconstrucción

Test	Técnica	EFO	Strehl ratio
Test 1	Sin corregir	644	0,048
	LS	293	0,296
	L+A	251	0,402
	Red neuronal	231	0,462
Test 2	Sin corregir	817	0,025
	LS	322	0,23
	L+A	289	0,3
	Red neuronal	262	0,37
Test 3	Sin corregir	1088	0,012
	LS	454	0,068
	L+A	409	0,1
	Red neuronal	387	0,125

Todas las métricas utilizadas indican que el mayor rendimiento se alcanza con la red neuronal, seguido por el Learn and Apply y finalmente el LS.

Sin embargo, estos tres casos de prueba son todos similares entre sí y no existe una gran diferencia en los rendimientos de las 3 técnicas. Para una comparación más exhaustiva se han diseñado tres perfiles de condiciones extremas no realistas, consistentes en una capa a nivel de superficie y otra que varía su altitud, ambas con la misma intensidad. De la misma forma que con los tres casos anteriores, LS y L+A fueron reconfiguradas entre las pruebas

Tabla 5: Error del frente de onda y Strehl ratio para las 3 técnicas de reconstrucción en los 3 casos de perfiles extremos

Reconstructor	Altitud de la capa superior (m)	EFO (nm)	Strehl ratio
Sin corregir		767	0,064
LS	5000	293	0,289
L+A		269	0,353
Red neuronal		211	0,52
Sin corregir		818	0,025
LS	10000	465	0,066
L+A		372	0,147
Red neuronal		297	0,287
Sin corregir		815	0,026
LS	15000	574	0,043
L+A		466	0,069
Red neuronal		390	0,127

La **Tabla 5** muestra que la red neuronal presenta un mejor comportamiento que las otras dos técnicas de reconstrucción en los tres perfiles. También puede verse que el rendimiento de todas las técnicas disminuye al aumentar la altitud de la capa móvil debido a la disminución de la superposición de los conos de luz del objetivo y la estrella guía (Smith, 2000).

5 Conclusiones

En el presente trabajo se demostró el potencial de las redes neuronales artificiales en la reconstrucción de frentes de onda mediante una serie de sensores tipo Shack Hartmann colocados fuera del eje del telescopio. Para ello se ha entrenado una red neuronal con una serie de datos de entrada diseñados para cubrir el mayor rango de vectores de entrada posible. Dicho set de datos se ha obtenido mediante simulación de una capa de turbulencia simple cambiando su altitud, generando de esta manera un conjunto de escenarios en los que el efecto de solapado de los conos de luz cambia.

Tras determinar el rendimiento de diferentes morfologías de red y algoritmos de entrenamiento se ha comprobado que la mejor combinación para los datos presentados es la más sencilla: Perceptrón multicapa con el mismo número de neuronas en la capa oculta que en la de entrada y con una sigmoide continua como función de activación. Otras morfologías arrojaban resultados favorables para algunos casos específicos de test, pero la antes citada presentaba un buen rendimiento para todos los casos. De todas formas aun existen multitud de combinaciones posibles que no se han comprobado en este trabajo, por lo que es posible continuar investigando en este aspecto.

La red neuronal entrenada en este trabajo parece mostrar mejor rendimiento con perfiles muy complejos que con aquellos más sencillos. A pesar de que la raíz del error cuadrático medio es inferior con los perfiles más simples, la fracción de varianza que es debida al error en si mismo es menor en el escenarios más complejos (test 3) que en los demás. De todas formas, el rendimiento de la red en los otros perfiles también es competitivo.

Comparando la red neuronal con otras técnicas de reconstrucción como Learn and Apply y LS, se puede observar que esta nueva técnica presenta unos valores de error de frente de onda menores y una mejor calidad de imagen.

Una de las mayores ventajas que presenta la red neuronal frente a los otros sistemas es que no requiere un retrenamiento cuando el perfil de atmosférico varía. La red fue capaz de predecir valores con un buen rendimiento para todos los perfiles puestos a prueba.

En el futuro el rendimiento de esta nueva técnica ha de ser contrastado con situaciones más realistas, en un banco de laboratorio y on-sky.

References

- Assemat, F., Gendron, E., & Hammer, F. (2007). The FALCON concept: multi-object adaptive optics and atmospheric tomography for integral field spectroscopy - principles and performance on an 8-m telescope. *Mon. Not. R. Astron. Soc.*, 376(1): ,287-312.
- Caruana, R., & A. Niculescu-Mizil (2004). Data mining in metric space: an empirical analysis of supervised learning performance criteria. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. Seattle, WA, USA, ACM: 69-78.
- Eberhart, R. C., & R. W. Dobbins (1990). Neural network performance metrics for biomedical applications. *Proceedings of third Annual IEEE Symposium on Computer-Based Medical Systems*, (282-289)
- Ellerbroek, B. L. (1994). First-order performance evaluation of adaptive-optics systems for atmospheric-turbulence compensation in extended-field-of-view astronomical telescopes. *J. Opt. Soc. Am. A* 11(2): 783-805.
- Fusco, T., Conan, J. M, Rousset, G. Mugnier, L.m., & Michov, V. (2001). Optimal wave-front reconstruction strategies for multiconjugate adaptive optics. *J. Opt. Soc. Am. A-Opt. Image Sci. Vis.* 18(10): 2527-2538.
- Gendron, E., Vidal, F. et al. (2011). MOAO first on-sky demonstration with CANARY. *Astron. Astrophys.* Vol. 529.
- Ghahramani, Z. (2004). Unsupervised Learning. En O. Bousquet, U. v. Luxburg y G. Rätsch (Eds.), *Advanced Lectures on Machine Learning*,(63) Springer-Verlag.
- Hashem, S. (1997). Optimal linear combinations of neural networks. *Neural Networks*, 10(4), 599-614.
- Hashem, S., & B. Schmeiser (1995). Improving model accuracy using optimal linear combinations of trained neural networks. *Ieee Transactions on Neural Networks*, 6(3), 792-794.
- Haykin, S. S. (1999). *Neural Networks: A Comprehensive Foundation*, Prentice Hall.
- Hubin, N., Arsenault, R., Conzelmann, R., Delabre, B., Le Louarn, M., Stroebele, S., & Stuik, R. (2005). Ground Layer Adaptive Optics. *Comptes Rendus Physique* 6(10), 1099-1109.
- Le Louarn, M., Hubin, N., Sarazin, M., & Tokovinin, A. (2000). New challenges for adaptive optics: Extremely large telescopes. *Monthly Notices of the Royal Astronomical Society* 317(3), 535-544.

- Morris, T., Hubert, Z., Myers, R. et al. (2009). CANARY: The NGS/LGS MOAO demonstrator for EAGLE. *1st AO4ELT Conference - Adaptive Optics for Extremely Large Telescopes*.
- Naftaly, U., Intrator, N., & Horn, D.. (1997). Optimal ensemble averaging of neural networks. *Network: Computation in Neural Systems* 8(3), 283-296.
- Osborn, J., Wilson, R. W., Dhillon, V.S., Avilo, R., & Love, G.D. (2011). Conjugate-plane photometry: reducing scintillation in ground-based photometry. *Mon. Not. R. Astron. Soc.* 411(2),: 1223-1230.
- Pineda, F. J. (1988). Dynamics and architecture for neural computation. *J. Complexity* 4(3): 216-245.
- Platt, B. C. & R. Shack (2001). History and principles of Shack-Hartmann wavefront sensing. *J. Refractive Surg.* 17(5): S573-S577.
- Sanz Molina, A., & Martín del Brío, B. (2006) *Redes Neuronales y Sistemas Borrosos*, Ra-Ma, Librería y Editorial Microinformática:España
- Smith, W. J. (2000). *Modern Optical Engineering, 3rd Ed.*, McGraw-Hill, Inc.: Nueva York.
- Vidal, F., Gendron, E., & Rousset, G. (2010). Tomography approach for multi-object adaptive optics. *J. Opt. Soc. Am. A-Opt. Image Sci. Vis.* 27(11): A253-A264.
- Wilson, R. W., & C. R. Jenkins (1996). Adaptive optics for astronomy: Theoretical performance and limitations. *Mon. Not. R. Astron. Soc.* 278(1): 39-61.
- Wyant, J. C. (1992). Basic Wavefront Aberration Theory for Optical Metrology En Academic Press, Inc., *Applied optics and Optical Engineering*. Volume XI, chapter 1