

Trabajo de Fin de Master

RESTAURACIÓN DE  
IMÁGENES ASTRONOMICAS  
DE ABERRACIONES  
ADMOSFERICAS A TRAVÉS  
DE INTELIGENCIA ARTIFICIAL

Deep Learning en Física Solar

Descripción

*Prueba y expansión del método de la diversidad de fase (Phase Diversity) para detectar la PSF (Point Spread Function) atmosférica que afecta a imágenes astronómicas por medio de redes convolucionales y Lstm (Long short – term memory)*

Alejandro Parra García  
alex.clr17@gmail.com

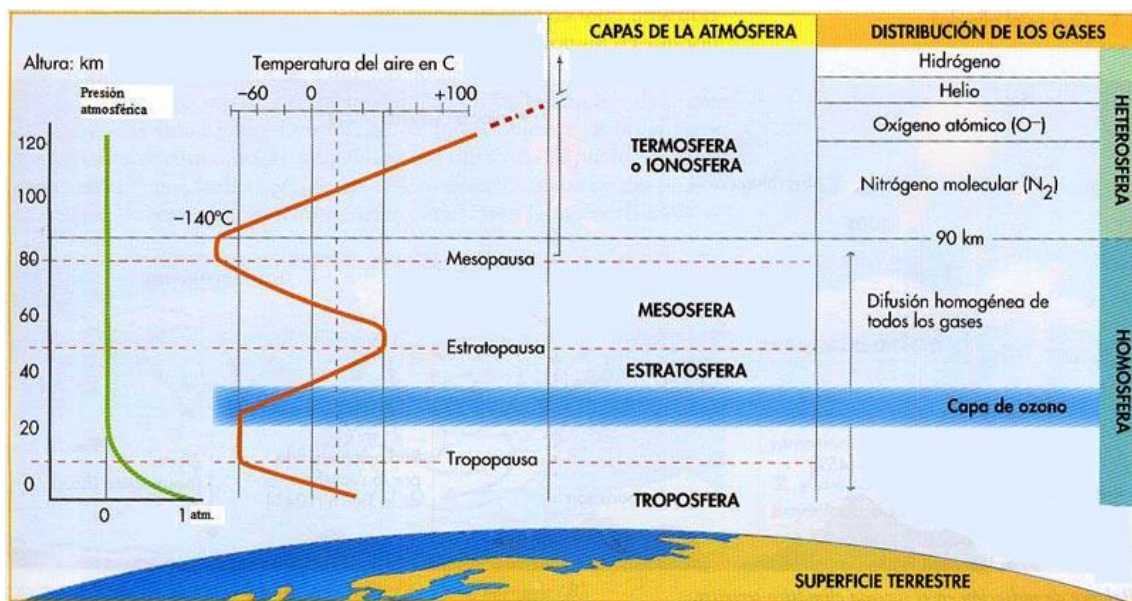
# Índice

<b>Introducción</b> .....	1
<b>DCGAN</b> .....	3
<b>Física del problema</b> .....	6
<i>PSF (Point Spread Function)</i> .....	6
<i>Aberraciones</i> .....	9
<i>Polinomios de Zernike</i> .....	10
<i>Descomposición de Karhunen–Loeve</i> .....	12
<b>Respuesta al problema</b> .....	14
<b>Creación del database</b> .....	16
<b>Modelo de Red</b> .....	19
<i>Modelo</i> .....	19
<i>Evaluación</i> .....	22
<b>Implementación</b> .....	26
<b>Resultados</b> .....	27
<b>Conclusiones</b> .....	31
<b>Referencias</b> .....	33

## Introducción

Uno de los problemas con los que ha tenido que lidiar la astrofísica desde que se comenzó a usar instrumentos (¡Y puede que incluso antes!) es el efecto de la propia atmósfera de nuestro planeta sobre la luz recibida de los objetos celestes. Y aún más importante si cabe, es el efecto del mismo instrumento sobre la luz que colecta y dirige hacia los detectores.

Esto se debe a la naturaleza ondulatoria de luz, por lo que es imposible evitar que difracte en los telescopios, o que se refracte al pasar de un medio a otro. Hasta cierto punto, la difracción puede ser controlada pues podemos estudiar su efecto según la geometría de los telescopios. Sin embargo, el fenómeno de la refracción no es tan amable, ya que podemos considerar como medios distintos volúmenes de atmósfera a distinta temperatura o presión. Y el problema gana más complejidad cuando caemos en que la atmósfera es un fluido que, debido a su extensión, la gravedad, y su contacto con el suelo, presenta una estructura compleja:



“Estructura simplificada de la atmósfera y representación de algunas de sus características  
Imagen obtenida de: <https://www.biologiasur.org/index.php/teoria/los-sistemas-fluidos-terrestres-externos/la-atmosfera>”

Donde, además, habría que introducir las corrientes de aire provocadas por el intercambio de calor con el suelo, tener en cuentas las características topográficas de la zona, nivel de humedad, etc.

Debido a esto, se han hecho grandes esfuerzos por parte de la comunidad científica para enfrentarse a este problema. Y las soluciones van desde métodos directos como poner los telescopios en órbita fuera de la atmósfera, a métodos más complejos que pasan por estudiar y comprender la naturaleza de la atmósfera, siendo un claro ejemplo de ello la

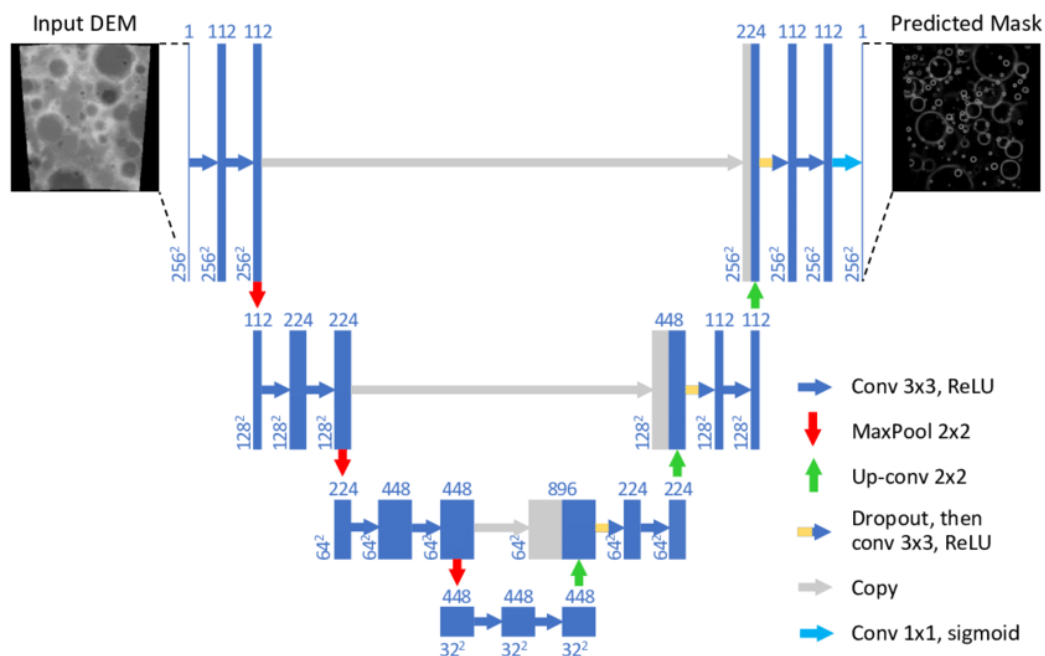
óptica adaptativa, que mide constantemente las perturbaciones del frente de onda para contrarrestarlos a altas frecuencias (kHz) a través de deformar los elementos ópticos que reciben el haz de luz. Por desgracia, la óptica adaptativa también sufre de limitaciones, que vienen dadas por las limitaciones en la señal y el ruido que el detector del frente de onda recibe; el número limitado de actuadores en los elementos ópticos; posibles retrasos entre el frente de onda medido y su corrección; y la limitación que supone tener que encontrar objetos de referencia cercanos al objeto de estudio. Esto último puede solucionarse con la creación de estrellas artificiales producidas por un haz laser, pero de nuevo, esto supone dirigir el laser a una altura determinada, con lo que se pierden contribuciones de otras capas atmosféricas.

Otra rama por la que se intenta atacar a este problema es de la corrección de las imágenes tomadas a-posteriori, a través de algoritmos dados por la física, intentando encontrar como deshacer las aberraciones de las imágenes a partir de la información de las propias imágenes o de imágenes extra.

El trabajo que aquí se presenta es un estudio sobre cómo los avances en hardware e inteligencia artificial podrían permitir la creación de métodos para eliminar o paliar este efecto en imágenes astronómicas. Se probará un tipo de red neuronal convolucional (*CNN*) utilizada mayormente en la modificación de imágenes, y se dará con ello paso a una discusión sobre porqué es necesario aplicar métodos más complejos que integren la física en su funcionamiento. Además, se intentará crear una arquitectura de red que sea insensible al ruido, lo cual probaremos su eficacia y aguante haciendo pruebas con señales de ruido cada vez más elevadas.

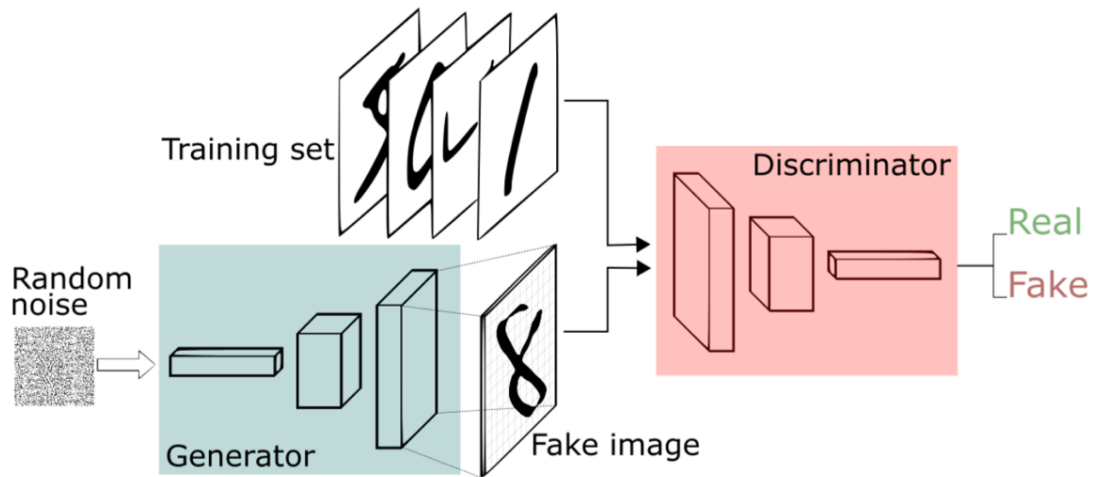
## DCGAN

Este trabajo se abordó, como toma de contacto, y como intento inocente por así decirlo, con la utilización de una red convolucional adversaria generativa DCGAN (*Deep Convolutional Generative Adversarial Network*) dada por el famoso artículo de “*Image-To-Image Translation with conditional adversarial networks*”, la cual pone a trabajar a dos redes convolucionales con trabajos principalmente opuestos. En este intento, una de estas redes era lo que se conoce como un autoencoder, una U-net (red convolucional llamada así por la forma de su grafo) que codifica la información de una imagen hasta un vector cuyos elementos deberían contener la información esencial de dicha imagen; y un decodificador que sea capaz de coger este vector y saber recrear una imagen a partir de dicha información.



“Ejemplo de estructura de una U-net  
Imagen tomada del artículo: *Lunar Crater Identification via Deep Learning*”

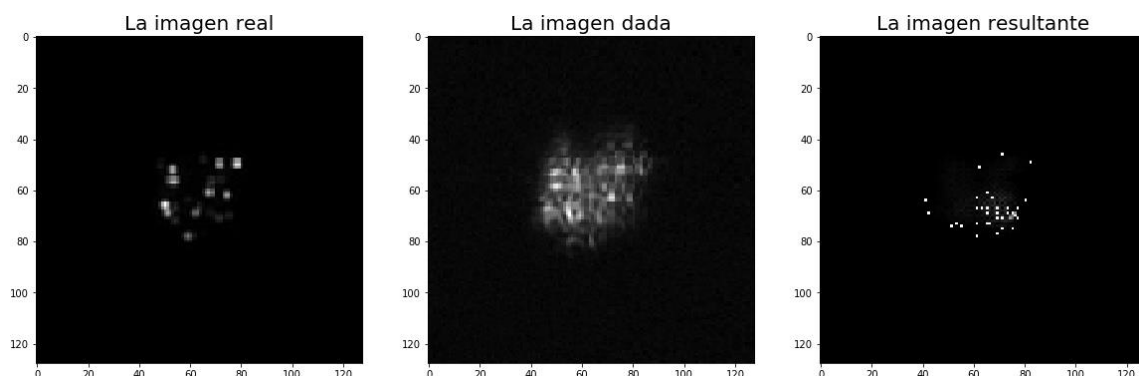
Mientras que la red adversaria a este última se la conoce como “discriminador”, y es una red encargada de aprender a distinguir si la imagen obtenida en la U-net es una imagen que podría ser parte del conjunto de entrenamiento o no. Es decir, el discriminador se encarga de juzgar si la imagen creada en la U-net es falsa o podría hacerse pasar por una verdadera, dando una señal de error en caso de juzgarla falsa.



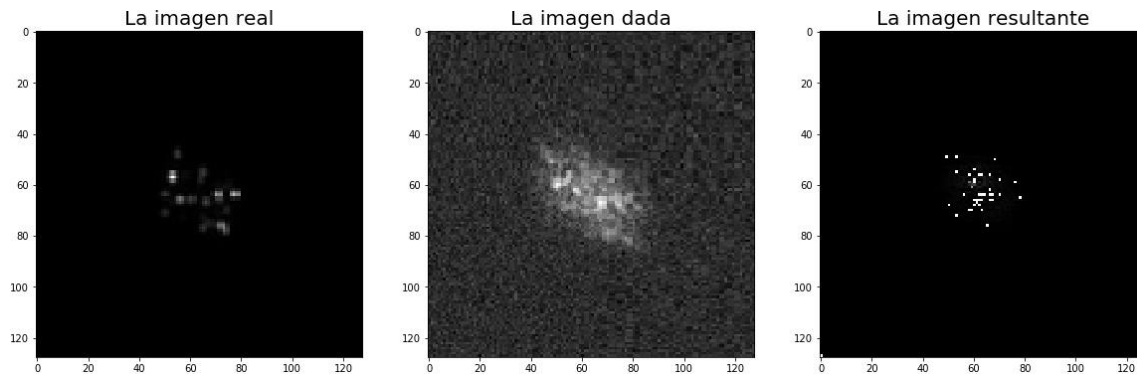
“Ejemplo de estructura de red convolucional adversaria generativa  
 Imagen obtenida de: <https://castelle.org/pages/social-theory-for-generative-networks-and-vice-versa.html>”

Dicho esto, queda claro el porqué del nombre de “redes adversarias”, pues la U-net aprenderá a crear imágenes cada vez más parecidas al set de entrenamiento para que el discriminador no dispare el error; y el discriminador aprende a distinguir mejor las imágenes falsas. La confrontación lleva a ambas redes a una precisión mucho mayor.

Sin embargo, y como ya se adelantó, este fue solo un intento, pues los resultados (para entrenamientos de 50 épocas con 50 imágenes) fueron:



“Caso para imagen con señal de ruido de 0,1 su media. A la izquierda: Imagen ideal que queríamos recuperar, solo la difracción de la luz producida en el telescopio afecta a la imagen. En el centro: Imagen observada en el foco del telescopio. A la derecha: Imagen recuperada por la red neuronal”



*“Caso para imagen con señal de ruido de 0,7 su media. A la izquierda: Imagen ideal que queríamos recuperar, solo la difracción de la luz producida en el telescopio afecta a la imagen. En el centro: Imagen observada en el foco del telescopio. A la derecha: Imagen recuperada por la red neuronal”*

La red parece aprender a eliminar la luz dispersada por la imagen y el ruido, incluso para intensidades altas, pero falla en recrear exactamente no solo el número y posición de las estrellas si no también su distribución de intensidades. Pensándolo, esto era de esperar. Pues dado que la única información con la que trabaja esta red es la imagen dada y su comparación con la imagen ideal a recuperar, solo trabaja con las intensidades en la imagen, con lo que en la mayoría de los casos difícilmente puede saberse cuantas estrellas hay realmente en la imagen y donde están.

Por lo que con este intento queda claro que, de querer reconstruir una imagen astronómica de efectos atmosféricos, debemos no solo quedarnos en la forma visual de las imágenes y atender algo más profundo. Algo que tenga que ver con las distorsiones que queremos corregir. Y es entonces cuando es inevitable pensar en la *PSF*.

## Física del problema

### PSF (Point Spread Function)

Se define la PSF como la función que representa la distribución de intensidades en el plano imagen que produce el sistema observacional de un objeto puntual situado en el infinito y observado en el eje óptico del sistema. Físicamente, representa la modificación del frente de onda (el cual suponemos que llega al telescopio como onda plana, dado que el objeto observado se sitúa en el infinito). Matemáticamente esta idea queda clara:

Si la imagen que se forma en la pupila de salida del telescopio se representa por la distribución de amplitudes

$$P_a(W) = p e^{iW}$$

Donde "W" representa la aberración el frente de onda; y "p" es la función de la abertura del telescopio, que vale 1 si el punto está dentro de la pupila, y 0 si está fuera.

Lo que vemos en el plano imagen del telescopio es:

$$P_b(W) = \mathcal{F}(P_a(W))$$

Donde  $\mathcal{F}()$   $\equiv$  Transformada de Fourier.

Pues bien, la PSF se define como:

$$S(x, y) = |\mathcal{F}(P_b(W))|^2$$

Fíjense que la PSF es simplemente el espectro de potencia de la imagen formada en el plano imagen del telescopio.

Y dado que la PSF da la distribución de intensidades para un objeto puntual, en caso de tener un objeto extenso, habrá que aplicar la PSF a cada uno de los puntos de la imagen. Matemáticamente consiste en efectuar el producto de convolución:

$$I(x, y) = \iint_{\text{Objeto}} O(x', y') S(x - x', y - y') dx' dy'$$

Por lo que se pueden aprovechar las propiedades matemáticas de las transformadas de Fourier para escribirlo de forma más manipulable y sencilla como:

$$i(\vec{f}) = o(\vec{f}) \cdot s(\vec{f})$$

Con:

$O(x', y')$  la distribución de intensidades real del objeto observado.

$I(x, y)$  la distribución de intensidades del objeto observado dada por el telescopio.



$i(\vec{f}), o(\vec{f}), s(\vec{f})$  son las transformadas de Fourier de  $I(x, y), O(x, y)$ , y  $S(x, y)$ .

$(x, y)$ , son las coordenadas espaciales, y  $\vec{f}$  son las frecuencias espaciales en el plano imagen.

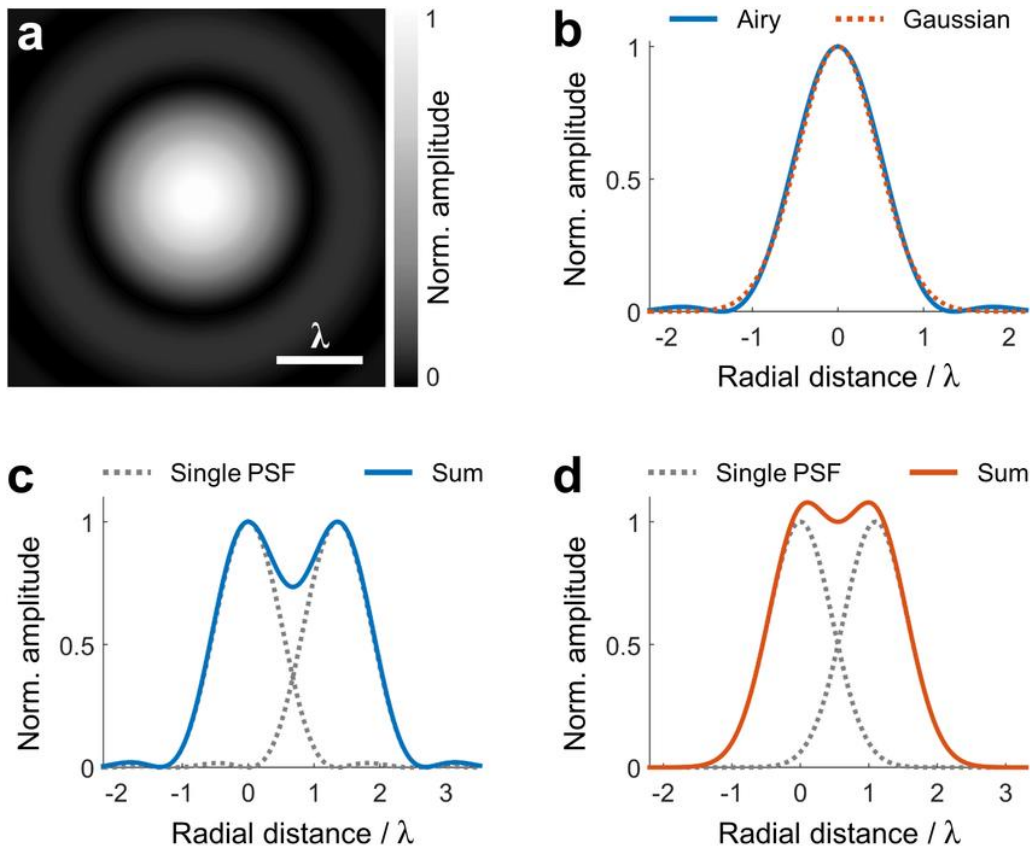
Hay que resaltar que  $s$  recibe el nombre de *Función de Transferencia Óptica (OTF)* y, en resumidas cuentas, da la magnitud en la que las frecuencias espaciales del objeto se transmiten a la imagen. Es sumamente importante, no solo porque será una de las principales herramientas que usemos, sino porque disminuye para altas frecuencias. Esto último quiere decir que, dado cualquier sistema óptico, existe una *frecuencia de corte* para la cual las frecuencias superiores no se transmiten.

$$f_c = D/\lambda \quad (D, \text{tamaño máximo de abertura}; \lambda, \text{longitud de onda observado})$$

Como queda patente, el cambio de representación de intensidades al dominio de frecuencias (es decir, suma de infinitas frecuencias sinusoidales dadas por las transformadas de Fourier), no solo es útil para la representación matemática, si no que alivia considerablemente el coste computacional, al tener que realizar ahora multiplicaciones donde antes había integrales de convolución.

Típicamente, para un telescopio de abertura circular, la PSF de un objeto se describe por la función de Airy, concentrando la luz en un punto central rodeado de anillos concéntricos. Este hecho hace que los patrones de objetos próximos se superpongan, por lo que solo se podrán resolver e identificar como objetos individuales si guardan una cierta distancia angular entre ellos. Esta distancia mínima define la máxima resolución espacial que es capaz de alcanzar el telescopio, y por tanto viene dada por el tamaño de la PSF, menor cuanto mayor sea el diámetro del objetivo, y menor cuanto menor sea la longitud de onda a la que se observa.

Por desgracia, esta distancia mínima para separar dos objetos podía tomarse de manera subjetiva, por lo que surgieron varios criterios que tratan de dar una justificación a cuando dos objetos pueden considerarse resueltos: *Criterio de Rayleigh*, dos objetos estarán resueltos cuando el máximo en el patrón de intensidad de un objeto está, por lo menos, en el mínimo del patrón de intensidad del otro objeto; *Criterio de Sparrow*, dos objetos estarán resueltos cuando los máximos de cada uno estén separados entre sí por una distancia superior a su FWHM (anchura a media altura del máximo).



**a).** Disco de Airy, PSF dada por un sistema óptico al observar un objeto puntual en el infinito. **b).** Sección transversal del disco de Airy (línea azul) comparada con el perfil de una distribución de Gauss (línea discontinua naranja). **c).** Sección cruzada del perfil de dos objetos cercanos separados a la distancia mínima dada por el criterio de Rayleigh (perfiles individuales en líneas grises discontinuas, y la suma en línea continua azul). **d)** Misma situación que en el caso c), pero ahora la separación es la dada por el criterio de Sparrow.

Imagen sacada de: [https://www.researchgate.net/figure/Point-spread-function-and-resolution-a-Airy-pattern-representing-the-2D-transverse\\_fig9\\_323835351](https://www.researchgate.net/figure/Point-spread-function-and-resolution-a-Airy-pattern-representing-the-2D-transverse_fig9_323835351)

No obstante, la mejor resolución que se puede alcanzar en una imagen tomada suele venir limitada no por el telescopio, si no por la atmósfera terrestre (obviamente no se aplica a telescopios espaciales). Esto significa que, en su paso por la atmósfera, un haz de luz se verá irremediabilmente modificado, alejándose más del frente de onda plano ideal esperado cuanto mayor sea la atmósfera tenga que atravesar el haz. Estas alteraciones, que se traducen en distorsiones que degradan aún más la resolución de las imágenes observadas, es lo que llamamos *aberraciones*.

## Aberraciones

Como ya se dijo, las aberraciones son defectos producidos en las imágenes resultantes debido a modificaciones del frente de onda del haz de luz por el telescopio y/o atmósfera.

Gracias a los estudios de laboratorio las aberraciones producidas por los instrumentos ópticos se conocen muy bien, pudiendo clasificarlas según sus efectos, y pudiéndose corregir en muchos casos. Hablo de la aberración *Esférica*, *Coma*, *Astigmatismo*, *Curvatura del campo*, *Distorsión* y aberración *Cromática*.

Debido al alto conocimiento sobre estas (y al objetivo de este trabajo), nos centraremos más en las aberraciones producidas por la atmósfera.

Las aberraciones atmosféricas cambian constantemente con el tiempo, lo que da el efecto de que a cada fotograma se le han añadido aberraciones de forma aleatoria. Podemos representar su efecto como:

$$\mathbf{s}_T(\vec{f}) = \mathbf{s}_{ins}(\vec{f}) \mathbf{s}_{atm}(\vec{f})$$

Es decir, la *OTF* resultante  $\mathbf{s}_T(\vec{f})$ , será el resultado de tener en cuenta la *OTF* del sistema instrumental  $\mathbf{s}_{ins}(\vec{f})$ , y la  $\mathbf{s}_{atm}(\vec{f})$  que llamaremos *Función de Transferencia Atmosférica (ATF)*.

Debido a la naturaleza estocástica de la atmósfera, la *ATF* tendrá que venir dada por una función estadística denominada *Función de estructura de Fase*,  $D_\varphi(\vec{r})$ , de forma que:

$$\mathbf{s}_{atm}(\vec{f}) = \exp [-0.5D_\varphi(\lambda\vec{f})]$$

Donde, para este trabajo, tomaremos la *Función de estructura de Fase* del modelo de Kolmogorov:

$$D_\varphi(\vec{r}) = 6.88 \left( \frac{|\vec{r}|}{r_0} \right)^{5/3}$$

Con  $r_0$  el *Radio de coherencia atmosférico*, o también llamado *Parámetro de Fried*, que da una medida de a partir de que tamaño de abertura de telescopio la resolución de las imágenes obtenidas quedará limitada por los efectos atmosféricos sobre los instrumentales. Esto sucederá para telescopios cuyo diámetro de abertura sea superior a  $r_0$ . Además, hace las veces de longitud de escala a la cual los fenómenos turbulentos en la atmósfera son significativos. Su valor típico depende de la longitud de onda observada,  $r_0 \propto \lambda^{6/5}$  (Para el rango visible toma valores 20 cm a 5 cm típicos a nivel del mar).

Por este motivo, suponiendo el modelo de Kolmogorov, existe una relación entre la varianza encontrada en el frente de onda observado y la abertura del telescopio:

$$\sigma^2 = 1.2099 \left( \frac{D}{r_0} \right)^{5/3}$$

Y entre el *seeing* y el *parámetro de Fried*:

$$r_0 = 0.98 \frac{\lambda}{\text{seeing}}$$

El modelo de Kolmogorov, fue desarrollado por Valerian Tatarskii (*Wave Propagation in a Turbulent Medium*) a partir los estudios sobre turbulencias de Andreï Kolmogorov, y se basa en que las perturbaciones del frente de onda son producto de continuos cambios del índice de refracción debido a turbulencias debidas a gradientes de temperatura y/o presión. El cambio es continuo en cada punto y para cada tiempo, y se traducen directamente en las fluctuaciones de fase que se observan.

### Polinomios de Zernike

Una manera muy útil y práctica de describir las aberraciones de un frente de onda consiste en describir estas a través de los polinomios de Zernike. Los polinomios de Zernike, nombrados en honor al físico ganador del premio nobel en 1953 Frits Zernike, son una familia de polinomios ortonormales, que, escritos en coordenadas cartesianas su forma general es:

$$W(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_{nm} x_n y_n$$

Por lo que geoméricamente el polinomio representa una superficie dada por la suma de superficies más simples que representan los monomios. Dado entonces que la superficie final queda entonces definida por los coeficientes  $x_n y_n$ , por lo que una aproximación por mínimos cuadrados nos lleva a la ordenación de Noll, con la cual podemos ordenar los términos de forma que pasamos de tener los dos índices: “ $n$ ” (índice radial) y “ $m$ ” (índice azimutal); a necesitar un solo índice “ $j$ ”. Cualquier función en la pupila se puede escribir entonces de la forma (escritos en coordenadas polares):

$$W(\rho, \varphi) = \sum_{j=1}^{\infty} a_j Z_j(\rho, \varphi)$$

Donde:

$W(\vec{r})$  sería la representación del frente de onda aberrado.

$a_j$  son los coeficientes de Zernike o también llamados *momentos de Zernike*.

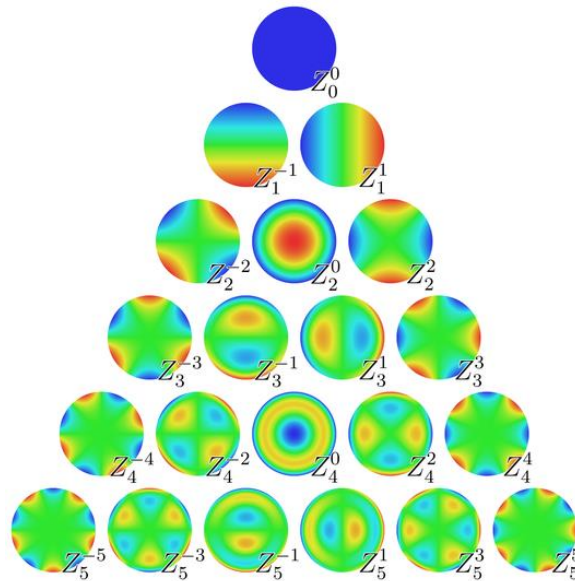
$Z_j(\rho, \varphi) = R_n^m(\rho) \cos(m\varphi)$  Para los términos  $j$  pares

$Z_j(\rho, \varphi) = R_n^m(\rho) \sin(m\varphi)$  Para los términos  $j$  impares

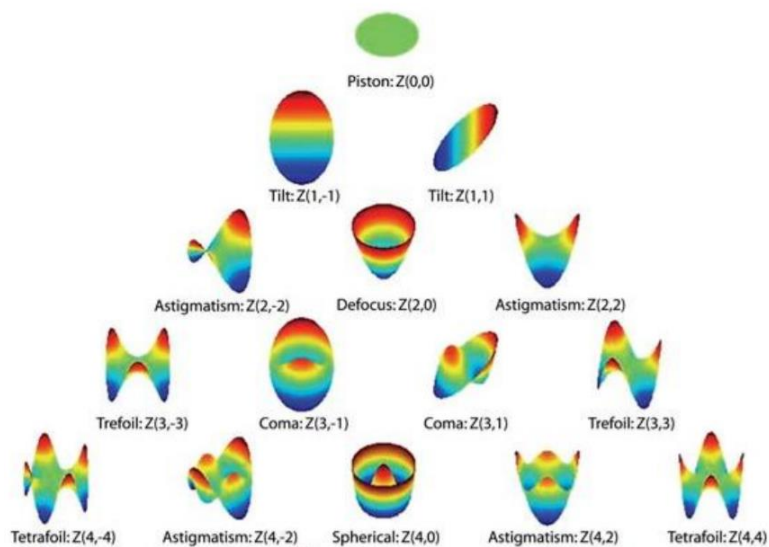
Con:

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} \rho^{n-2k}$$

De forma que el resultado final es:



“Imagen de los efectos de los primeros 21 términos de Zernike en dos dimensiones coloreados según la intensidad. Imagen obtenida de Wikipedia: [https://en.wikipedia.org/wiki/Zernike\\_polynomials#Radial\\_polynomials](https://en.wikipedia.org/wiki/Zernike_polynomials#Radial_polynomials)”



“Representación tridimensional de los efectos de los primeros 15 términos de Zernike sobre el frente de onda. Imagen obtenida de: Trabajo de fin de grado de Ester Soldevilla Foj, “Estudio de la relación entre Aberraciones Oculares y Acomodación””

En estas imágenes puede verse además la simetría que encontramos a ambos lados de la pirámide, hecho que puede explotarse para reducir los tiempos de cálculo y carga de procesamiento del ordenador.

Nos interesa, además, conocer las propiedades estadísticas de los coeficientes (*los momentos de Zernike*)  $a_j$ . Por lo que, si suponemos que estos coeficientes pueden considerarse como variables gaussianas aleatorias de media cero, la matriz de covarianza de estos queda:

$$\langle a_j^* a_i \rangle = \int d\vec{r}_j \int d\vec{r}_i A(\vec{r}_i) A(\vec{r}_j) Z_j(r_j) C(\vec{r}_j, \vec{r}_i) Z_i(r_i)$$

Donde se integra al área de la imagen dentro de la abertura del telescopio.

$C(\vec{r}_j, \vec{r}_i)$  es la función de covarianza de fase:

$$C(\vec{r}_j, \vec{r}_i) = \langle W(\vec{r}_j) W(\vec{r}_i) \rangle$$

$$Y \quad A(r_j) = \begin{cases} \frac{1}{\pi} & \text{Dentro de la imagen de la abertura del telescopio} \\ 0 & \text{Fuera} \end{cases}$$

La matriz de covarianza formada es simétrica.

### Descomposición de Karhunen–Loève

Los polinomios de Zernike son perfectos para representar aberraciones de orden bajo del frente de onda, pero tienen el inconveniente de no preservar las componentes de frecuencias espaciales altas. Esto, en la práctica, no nos afectará dado que nos limitaremos los 40 primeros términos de Zernike.

Sin embargo, en teoría, la forma óptima de representar procesos estocásticos son un conjunto de funciones conocidas como los modos de Karhunen–Loève, que consisten en representar el frente de onda como una combinación lineal de funciones ortonormales:

$$W = \sum_{\text{modos}} b_m = \sum_{\text{modos}} \sum v_i Z_i$$

Donde  $W$  sería la representación del frente de onda; los  $b_m$  son los *modos de Karhunen–Loève*; y la descomposición de estos últimos se compone de funciones ortonormales  $Z_i$ , y los términos  $v_i$  que serían las componentes, para cada modo, de los autovectores resultantes de la diagonalización de la matriz de covarianza de las funciones elegidas.

Dicho esto, es fácil ver que lo que se usará en este trabajo para representar los frentes de onda será los modos de Karhunen–Loève usando los polinomios de Zernike como las bases de la combinación lineal, obteniendo los coeficientes  $v_i$  de los autovectores de la matriz de covarianza de los Zernike, que teniendo además en cuenta que se aplicará el modelo de atmósfera de Kolmogorov, los elementos de esta matriz de covarianza se calcularían como:

$$C(\vec{r}_j, \vec{r}_i) = \langle W(\vec{r}_j) W(\vec{r}_i) \rangle = C G$$

$$C = 0.0072 \left(\frac{D}{r_0}\right)^{\frac{5}{3}} (-1)^{\frac{(n_i+n_j-2m_i)}{2}} \sqrt{(n_i+1)(n_j+1)} \pi^{\frac{8}{3}} \delta_{m_i} \delta_{m_j}$$

$$G = \frac{\Gamma\left(\frac{14}{3}\right) \Gamma\left(\frac{n_i+n_j-\frac{5}{3}}{2}\right)}{\Gamma\left(\frac{n_i-n_j+\frac{17}{3}}{2}\right) \Gamma\left(\frac{n_j-n_i+\frac{17}{3}}{2}\right) \Gamma\left(\frac{n_i+n_j+\frac{23}{3}}{2}\right)}$$

Para  $i - j$  par (el caso impar es cero), los índices  $n$  y  $m$  son los correspondientes índices de Zernike para el índice de Noll para el modo  $j$ .  $\Gamma$  es la función gamma de Euler, y  $\delta$  es la delta de Kronecker.

## Respuesta al problema

Reduciendo el problema que se plantea a su forma más básica, queremos obtener la información de un objeto observado sin la pérdida o modificación de esa información por parte de la atmósfera. Como ya sabemos:

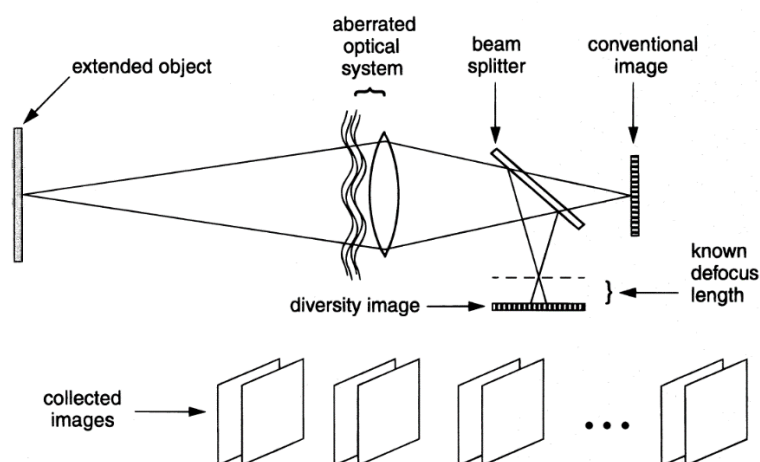
$$i(\vec{f}) = o(\vec{f}) \cdot s(\vec{f})$$

Por lo que el objetivo consistirá en hallar  $o(\vec{f})$ .

Sin embargo, las matemáticas son claras: un sistema con más incógnitas que ecuaciones no puede ser determinado. No obstante, y como ya se comentó, se han desarrollado métodos para tratar con este problema. Entre ellos destacan la técnica de moteado (*Speckle technique*) (Labeyrie 1970; von der Lühe 1993); el método de la diversidad de fase (*Phase diversity*) (Gonsalves 1979, 1982); y la técnica del moteado de fase diversa (*Phase-diverse Speckle*) (Paxman, Schulz, & Fienup 1992; Paxman & Seldin 1994).

Teniendo esto mente, nos inspiramos en el artículo “*Object-independent image-based wavefront*” [QI XIN, GUOHAO JU, CHUNYUE ZHANG, y SHUYAN XU], donde proponen una forma de desarrollar el método de la diversidad de fase aplicando inteligencia artificial.

El método de diversidad de fase (*Phase diversity*) conceptualmente es muy sencillo: añadimos al sistema una ecuación más, que es la misma pero aplicada a otro plano focal. Es decir, que contaríamos con la imagen convencional del plano focal, degradada por aberraciones desconocidas; y una segunda imagen del mismo objeto, obtenida mediante la división del haz de luz recogido por el telescopio con un splitter y un segundo detector. De esta forma, esta segunda imagen contendría también la información de las aberraciones desconocidas más el efecto añadido de un desenfoque que puede ser controlado.



“Esquema de la instrumentación necesaria para la obtención de las imágenes necesarias para el método de la diversidad de fase.

Imagen obtenida del artículo *Evaluation of phase-diversity techniques for solar-images restoration*”



De esta forma, observando el mismo objeto, no solo nos quitamos el problema de la dependencia temporal de la PSF, sino que, además, tal y como muestran en el artículo mencionado, aplicando el método de reducción vemos que:

$$\begin{cases} i(\vec{f})_1 = o(\vec{f})_1 \cdot s(\vec{f})_1 \\ i(\vec{f})_2 = o(\vec{f})_2 \cdot s(\vec{f})_2 \end{cases} \quad \frac{i(\vec{f})_1}{i(\vec{f})_2} = \frac{s(\vec{f})_1}{s(\vec{f})_2}$$

Que aplicando la transformada de Fourier, y recordando lo que sabemos de la *PSF*:

$$f = \frac{\mathcal{F}(i(\vec{f})_1)}{\mathcal{F}(i(\vec{f})_2)} = \frac{P_1(W) \otimes P_1(W)}{P_2(W) \otimes P_2(W)}$$

Donde  $f$  resulta ser una imagen en intensidad que contendría la información característica de las imágenes dadas.  $P_i(W)$  es la función pupila, la imagen en la pupila del telescopio. El índice  $i$  sirve para representar los distintos planos focales a los que hagamos referencia, y  $\otimes$  representa la operación de auto-correlación. La idea fundamental es que las imágenes de la pupila la escribimos como:

$$P_1(W) = p e^{iW} \quad P_2(W) = p e^{i(W+\Delta W)}$$

En el artículo (*“Object-independent image-based wavefront”*), esto les lleva a la conclusión de que, conocido el desenfoque entre los dos planos focales  $\Delta W$ , podían recuperar el frente de fase aberrado  $W$  del sistema a partir de la información que recibida de  $i(\vec{f})_1$  e  $i(\vec{f})_2$ , y contenida en  $f$ .

Pues bien, nuestra idea es exactamente la misma, con el cambio de que nuestras imágenes ya cuentan con ruido intrínseco, e intentaremos corregir aberraciones mucho más complejas, usando contribuciones de hasta 40 polinomios de Zernike. Nuestra propuesta, además, es darle a la red neuronal el par de imágenes foco-desenfocada, para que sea la propia red la que a través de un codificador extraiga un mapa de características propio. El objetivo es dar lugar a una red que aprendería a lidiar con el ruido de la imagen.

Aplicaremos esta idea en imágenes astrofísicas de campos estelares, en los que incluiremos diferentes efectos de ruido de lectura, con el fin de, una vez obtenido el frente de onda aberrado, poder corregir ambos efectos. Es por ello que haremos uso de la tecnología de la inteligencia artificial, primero, para obtener las características del frente de onda guardada en las imágenes  $i(\vec{f})_1$  e  $i(\vec{f})_2$ , y segundo, para obtener los momentos de Zernike que representen este frente.

## Creación del database

Una de las partes fundamentales del desarrollo de redes neuronales consiste en la elección o creación del dataset, es decir, los datos o imágenes con las que la red trabajará para ir ajustando sus parámetros internos en pos de resolver el problema dado. O en otras palabras: aprender.

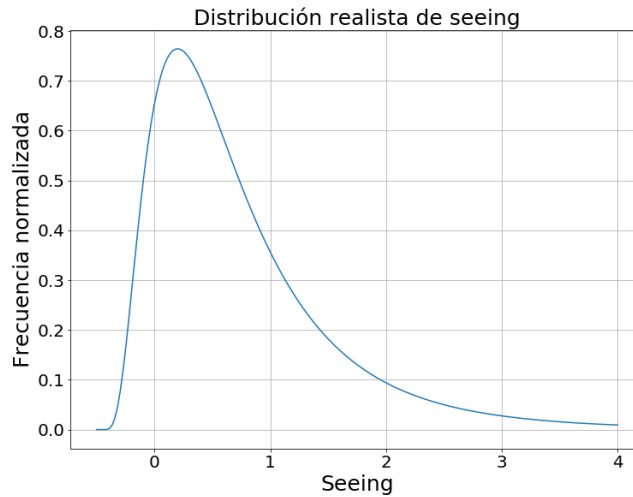
Lo fundamental de un conjunto de entrenamiento es que contenga un espacio rico de ejemplos de lo que queremos que sea aprendido, pero que a la vez no se extienda mucho a otros ámbitos.

Y debido al planteamiento de este trabajo, para poder tener las imágenes necesarias, hemos tenido que crear imágenes artificialmente. Esto se debe, en parte al alto volumen que las inteligencias artificiales necesitan en su entrenamiento, y cuya obtención de otro modo llevaría mucho tiempo de observación; y en parte a poder tener control total del contenido de estas imágenes. Además, de esta manera quedamos libres de problemas asociados a las medidas como contribuciones de ruido, o problemas al medir en un rango de longitudes de onda por muy estrechos que sean los filtros.

Crear el campo de estrellas puntuales en si es sencillo, basta tener una matriz nula del tamaño de las imágenes de modo que cada elemento representa un pixel, y colocar elementos aleatorios (siguiendo una distribución uniforme) con valores comprendidos entre  $10^{-3}$  y 1. Dado que estos valores representarían el valor del flujo medido por los pixeles, se obtiene un campo de objetos puntuales de magnitudes visuales que varían en un rango de  $\Delta m = 7.5$ .

Sin embargo, para que pueda servir de remplazo a una imagen real necesita ser convolucionada con una *PSF*, y es aquí donde hacemos uso de lo explicado anteriormente para construir un frente de onda aberrado, descrito por los polinomios de Zernike conociendo perfectamente los momentos de Zernike que lo conforman.

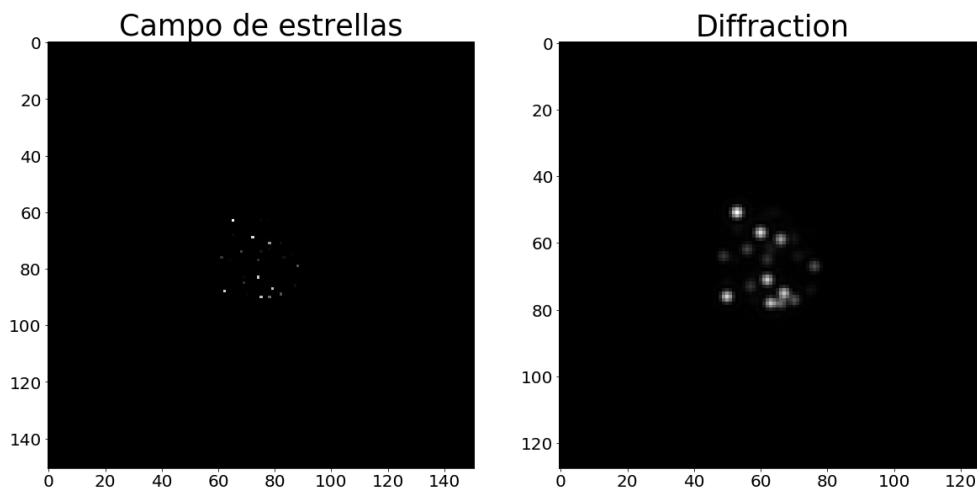
Para ello, se escoge un valor de *seeing* de forma aleatoria de una distribución lognormal con media  $-0.4638$ , y sigma  $0.4105$  (Datos obtenidos de estudios de calidad del cielo en dicho observatorio para el observatorio del Teide, Santa Cruz de Tenerife, España):



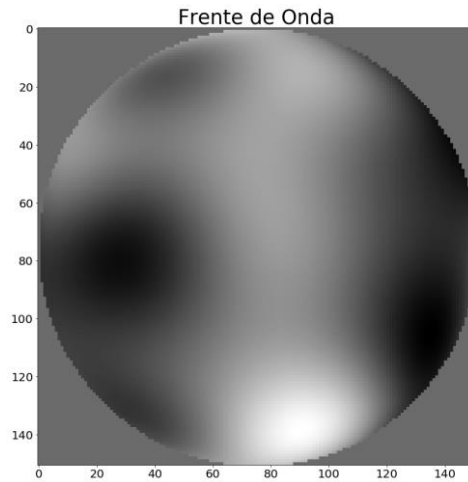
“Distribución realista de posibles valores de seeing atmosférico en una observación para el observatorio del Teide (Santa Cruz de Tenerife, España)”

Y si tenemos determinado el *seeing* tenemos el *parámetro de Fried*, el cual como vimos antes, nos servirá para determinar la una matriz de covarianza realista para tal *seeing*. Con esto, podemos hacer uso de los *modos de Karhunen–Loève* para representar un frente de onda aberrado natural para este *seeing* a partir de la matriz de covarianza de los polinomios de Zernike.

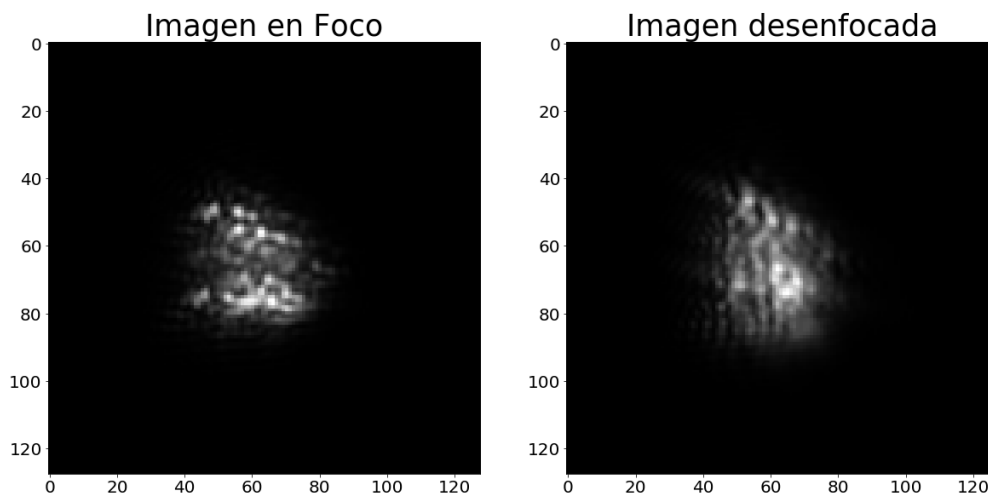
Ya con el frente de onda, calculamos la *PSF*, la *OTF*, y convolucionamos con ella el campo de estrellas puntuales:



“Ejemplo de imágenes artificiales creadas. De izquierda a derecha: Campo de estrellas puntuales; el mismo campo convolucionado con la PSF del telescopio, sin efecto atmosférico”



*“Frente de onda aberrado en el área de la pupila del telescopio”*



*“Ejemplo de imágenes artificiales creadas. De izquierda a derecha: Campo de estrellas puntual convolucionado con la PSF resultante del frente de onda aberrado; el mismo campo convolucionado con un frente aberrado con los momentos de Zernike aumentados”*

La forma en la que se ha creado el dataset por tanto es la siguiente: Se obtiene un valor de *seeing*; para ese *seeing* se crean  $n$  campos de estrellas puntuales; y aplica a cada campo de estrellas  $m$  PSF distintas. Los valores de  $n$  y  $m$  dependen del volumen de imágenes artificiales para el entrenamiento que se quieran crear, en nuestro caso, para 100 valores de *seeing*,  $n = 50$  y  $m = 20$ . Además, se debe tener en cuenta que serán el doble de lo sugerido aquí, pues como se comentó en la sección anterior, por cada imagen creada en el foco del telescopio, necesitamos crear otra desenfocada para poder resolver el problema. Esta imagen desenfocada se crea fácilmente aumentando el valor del tercer momento de Zernike del conjunto (el término de desenfoco) que conforman el frente de onda aberrado para la imagen en el foco. En la creación de nuestro dataset, este aumento fue de  $\pi/\sqrt{3}$ , valor para el cual, la diferencia entre las *rms* (*root-mean-square*) para los frentes de onda en la pupila de salida toma el valor de 1 lambda de error.

La inclusión de la simulación de ruido intrínseco en las imágenes se elaboró mediante la adición de una imagen que solo contuviera valores aleatorios en sus píxeles según una distribución gaussiana de media 0 y varianza igual a la media de intensidades en la imagen del foco multiplicado por un parámetro  $S_R$  cuyo valor podemos modificar para crear distintas intensidades de ruido. Al contar solo con ruido gaussiano estamos incluyendo tan solo ruido que podría provenir de la lectura del detector, y es cierto que podríamos simular ruido mucho más realista incluyendo ruido de señal, dado por una distribución Poisson, o incluso añadir ruido de patrón fijo (*Fixed Pattern Noise*), pero recordamos que este trabajo se centra en probar la eficacia de una inteligencia artificial detectando la *PSF*. En caso de verse potencial, podrán integrarse modelos de ruido más complejos y realistas.

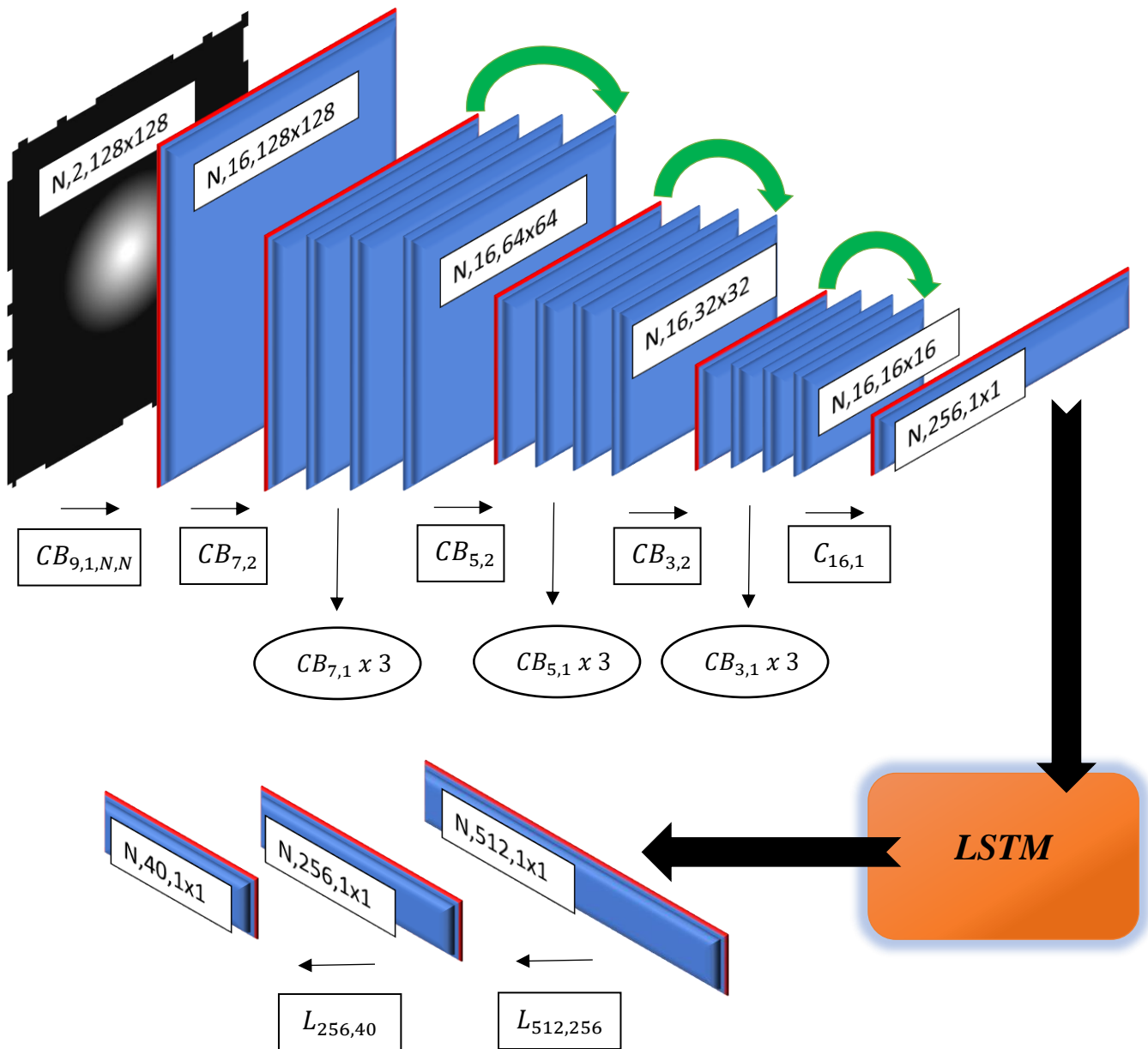
Como nota final a este apartado, queda decir que para poder tener ejemplos en los que conozcamos la contribución de cada polinomio en una imagen, se ha recreado la pupila de salida de un telescopio de 1.5 m de diámetro, y se ha tenido en cuenta como detector el instrumento *Fastcam*, con escala del detector  $\approx 0.042$  arcsec/pixel) observando a una longitud de onda de 800 nanómetros, con la idea de que posteriormente se pudiera probar la red creada con imágenes reales para pruebas de validación.

## Modelo de Red

### Modelo

Se ha utilizado una CNN (*Convolutional Neural Network*), dada la ya demostrada eficacia de estas redes a la hora de conseguir extraer los mapas característicos de imágenes. En nuestro caso, la entrada que le proporcionaremos a la red será una secuencia de  $N$  tensores, en el que cada cual contiene las dos imágenes que necesitamos (enfocada y desenfocada). Falta decir, que cada imagen contendrá un único canal, el que da la intensidad de la señal recibida.

Para una mayor comprensión, visualmente nuestro modelo tendría la siguiente forma:



De esta manera, la primera parte de la CNN será una parte *encoder* (codificación), que irá contrayendo la información hasta obtener un vector. -  $CB_{k,s,bn,act}$  - indica que hay que efectuar un bloque de convolución con:  $k$ , el tamaño del kernel;  $s$ , tamaño del paso del kernel;  $bn$ , si se realiza una normalización a las imágenes;  $act$ , si se pasa el resultado del bloque por la función unidad lineal rectificadora (*ReLU*).  $bn$  y  $act$  solo se indicará con  $N$  cuando no se realicen. Más explícitamente, el bloque de convolución  $CB$  consiste en:

If  $bn$ :  
 $bn \rightarrow ELU \rightarrow C_{3,1}$   
 Else:  
 $C_{3,1} \rightarrow (If\ act: ReLU)$

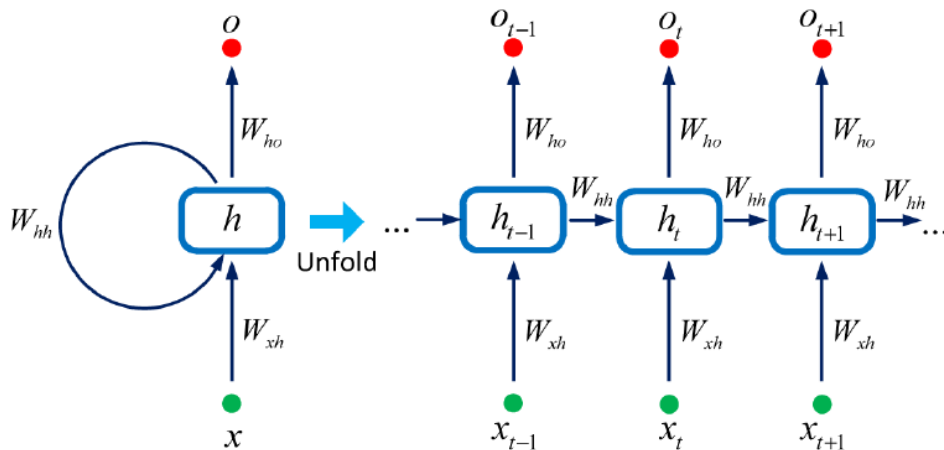
El caso  $C_{k,s}$  representa una convolución en dos dimensiones de tamaño de kernel  $k$ , y de paso  $s$ .

Como se puede ver en el diagrama, en las series de convolución que mantienen el tamaño constante aparece una flecha verde que une el primero resultado del bloque con el último. Esto representa una *Skip connection*. Introducidas por el artículo “*Deep Residual Learning for Image Recognition*” se basan en agregar información de capas anteriores a capas más profundas. Esto puede realizarse concatenando ambas capas en distintos canales o sumándolas. Las diferencias entre ambos métodos son sutiles, pero en general ambas formas permiten que la capa más profunda “recuerde” información pasada o la “corrija”. En nuestro caso se usará la suma, ya que permite además evitar que los gradientes de la retropropagación en el cálculo de errores se vuelvan muy pequeños, lo que ayuda a la convergencia de los parámetros del modelo.

Otra cosa a comentar es que puede verse que los  $CB$  mantiene el número de canales de imagen constantes (16 en nuestro caso), hasta que en  $C_{16,1}$  se pasa de 16 a 256. Y las imágenes con las que trabajamos se recortan o rellenan (proceso de *padding*) de modo que su tamaño antes de entrar a la red es de  $128 \times 128$ , por lo que al final de la parte de los  $CB$  estas se han reducido a un tamaño de  $1 \times 1$ . Es entonces cuando, para cada tensor de la secuencia dada, tenemos un vector de 256 elementos que guarda la información esencial de la imagen.

Llegados a este punto, es cuando hacemos uso de la *Lstm*. *Lstm* (*Long short – term memory*) es una variación de red neuronal recurrente (*RNN*) cuyo fin, en comparación con la cognición humana, es la de aportar el contexto aprendido de una situación o problema para procesar correctamente nueva información según una coherencia temporal. Un claro ejemplo sería el visionado de una película. De un fotograma cualquiera poca información de la obra puedes obtener, pero el caso es muy diferente si has visto todos los fotogramas previos, pues ya tienes información sobre todo lo que pueda ofrecer ese fotograma concreto: personajes, lugares, objeto, etc. Así, al darle una secuencia de vectores característicos, esta red aprenderá que información pertenece al objeto, y que parte pertenece a las aberraciones, consiguiendo con ello obtener unos nuevos vectores característicos mucho más precisos en su información sobre las aberraciones que afectaban a cada imagen de la secuencia.

La *Lstm* se estructura como:



“RNN “desenrollando” el bucle de recurrencia.  
Obtenida del artículo “Object-independent image-based wavefront””

$$\begin{aligned}
 i_t &= \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \\
 f_t &= \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \\
 g_t &= \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \\
 o_t &= \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \\
 c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\
 h_t &= o_t \odot \tanh(c_t)
 \end{aligned}$$

“Secuencia de cálculo en la *Lstm* para cada elemento de la secuencia introducida  
<https://pytorch.org/docs/master/generated/torch.nn.LSTM.html>”

Dado un vector de entrada  $x = (x_1, \dots, x_T)$  la *RNN* calculará el vector oculto  $h = (h_1, \dots, h_T)$ , y devolverá el vector  $o = (o_1, \dots, o_T)$ .  $\sigma$  es la función sigmoide, y  $\odot$  es el producto de Hadamard. Los términos  $W$  son los parámetros de la red *Lstm*, y el resto son variables intermedias.

Cabe mencionar, que dado el conjunto de imágenes de entrenamiento que hemos creado, estas no guardan relación temporal alguna entre ellas, pero, aun así, la *Lstm* debería ser capaz de reconocer patrones haciendo que la red “recuerde” que la intensidad de los píxeles no es independiente. La inclusión de la *Lstm* esta pensada para pruebas futuras cuando se tomen imágenes reales.

Finalmente, los vectores salientes de la *Lstm* pasan por aplicaciones lineales  $L$ , que van reduciendo el número de elementos del vector a la mitad hasta obtener un vector con tantos elementos como número de momentos de Zernike queramos extraer de las imágenes.



## Evaluación

Con esto ya cubrimos la parte calcularía los momentos de Zernike que queremos, con lo que podremos recrear la PSF que afecta a las imágenes, pero falta una parte fundamental en el desarrollo de las inteligencias artificiales: evaluar de alguna forma el resultado obtenido, de modo que la red pueda reajustar sus parámetros y realizar el proceso de aprendizaje.

Entramos así en una de las partes con mayor discusión del trabajo. Y es que, aunque a primeras uno podría pensar que bastaría hacer uso del aprendizaje supervisado comparando los momentos de Zernike obtenidos por la red con los momentos que conforman la imagen dada (pues recordemos que son conocidos. Las imágenes de entrenamiento son artificiales creadas por nosotros mismos), esto queda descartado si pensamos en que se puede escribir el frente de onda  $W(\vec{r})$  como combinación lineal de los polinomios de Zernike:

$$W(\vec{r}) = \sum_{j=1}^{\infty} a_j Z_j(\rho, \varphi)$$

Y de este frente de onda se forma la *PSF*,  $S(x, y)$ , tal y como se enseñó en los otros apartados:  $S(x, y) = |\mathcal{F}(P_b(W))|^2$

Con:  $P_b(W) = \mathcal{F}(P_a(W)) \quad P_a(W) = p e^{iW}$

Por lo que un frente  $W(\vec{r})$  daría exactamente la misma *PSF* que otros frentes  $W(\vec{r}) + 2\pi n$ . Es decir, vemos entonces que la *PSF* esta degenerada en cuanto a frentes de onda que pueden resultar en la misma *PSF*. Y este hecho llevado a la práctica, donde tenemos que quedarnos con un conjunto de 40 polinomios de Zernike, y que estemos en un problema de reconocimiento de esta *PSF*, hace que no solo entren las combinaciones que dan la misma *PSF*, si no también aquellas que se le parezcan.

Es por ello que tomaremos un camino algo más complejo al informar el entrenamiento usando la física de la formación de imagen y evaluaremos el resultado de la red mediante una función de coste MOMFBD (*Multi-Object Multi-Frame Blind De-convolution*). Para ello, partimos de la diferencia cuadrática media entre la imagen dada y la estimación de esta que podemos hacer con lo obtenido de la red:

$$L = \sum_{m=1}^M \sum_{i=1}^2 |y_{input,i} - y_{red,i}|^2 =$$

$$= \sum_{m=1}^M |i_1(\vec{f}) - o(\vec{f}) s_1(\vec{f})|^2 + \gamma |i_2(\vec{f}) - o(\vec{f}) s_2(\vec{f})|^2$$

Donde  $i_i(\vec{f})$  es la imagen dada a la red ( $i = 1,2$  imagen enfocada y desenfocada);,  $s_i(\vec{f})$  es la *OTF* que formamos con los momentos de Zernike calculados por la red;  $o(\vec{f})$  es la imagen del objeto sin la distorsión atmosférica;  $\gamma$  es una medida de la diferencia señal-ruido de las dos imágenes, calculada como el cociente de sus varianzas. El sumatorio sobre  $M$ , indica sumar el resultado obtenido en todas los  $M$  tensores que contenga la secuencia dada a la red como entrada.

Esta expresión de por sí sola no es muy útil, dado que queremos realizar lo que se conoce como una deconvolución ciega. Es decir, la expresión depende de la *OTF* que queremos calcular y evaluar, y a la vez, de la imagen objeto que a priori no conocemos pues es la información que queremos recuperar.

Este problema puede solventarse si pasamos a pensar en términos de probabilidad. Para ello identificamos nuestra función de evaluación  $L$  como el  $-\log(\text{Likelihood})$  (*Negative log likelihood*), interpretando el *Likelihood* como la probabilidad de que el resultado de la red sea correcto. Teniendo esto claro, podemos aplicar el método de la máxima verosimilitud (*maximum Likelihood*) que nos permite estimar  $o(\vec{f})$  diferenciando  $L$  respecto a  $o(\vec{f})$ , de forma que obtengamos el valor de  $o(\vec{f})$  que maximiza el *Likelihood*. Obtenemos que  $o(\vec{f})$  debe ser:

$$o(\vec{f}) = \frac{\sum_m^M (i_1(\vec{f}) s_1(\vec{f})^* + \gamma i_2(\vec{f}) s_2(\vec{f})^*)}{\sum_{m=1}^M (|s_1(\vec{f})|^2 + \gamma |s_2(\vec{f})|^2 + var)}$$

Aunque aquí pudiera haber algo de confusión al ver que necesitamos tener una *PSF* para cada imagen mientras que nuestra red solo calcula la de la imagen en el foco, recordamos que la imagen desenfocada se obtenía simplemente sumando un término conocido de desenfoco a la *PSF* del foco. Es por ello que realmente tenemos las dos. Además, introducimos un termino *var*, que es, para cada imagen, la varianza en el foco. para evitar posibles divisiones por cero por parte de elementos de la *PSF* que puedan permanecer nulos, pero siendo un valor que aporta información de la imagen. Este es un mal necesario, pues la *PSF* no ocupa toda la imagen, pero modificamos algunas frecuencias, las cuales perdemos la información contienen. Solo queda ver si la inteligencia artificial es capaz de afrontar también este inconveniente.

Dicho esto, si sustituimos la expresión de  $o(\vec{f})$  en la diferencia cuadrática media, tenemos una expresión que no depende directamente de la imagen objeto  $o(\vec{f})$ :

$$L = \sum_{m=1}^M (|i_1(\vec{f})|^2 + \gamma |i_2(\vec{f})|^2) - \frac{|\sum_{m=1}^M (i_1(\vec{f}) s_1(\vec{f})^* + \gamma i_2(\vec{f}) s_2(\vec{f})^*)|^2}{\sum_{m=1}^M (|s_1(\vec{f})|^2 + \gamma |s_2(\vec{f})|^2 + var)}$$

Esta será la medida utilizada para evaluar el trabajo de la red neuronal. No obstante, este método también nos supone enfrentarnos a algunos problemas:

Primero, la alta sensibilidad al ruido, que viene descrito sobre todo por las frecuencias altas, las cuales perdemos al recrear la *PSF*. *PSF* que además viene dada por la minimización de nuestra función de evaluación y formada además por un número limitado de momentos de Zernike. Si a esto añadimos la degeneración antes vista de la *PSF* con el frente aberrado calculado, es fácil pensar que la solución encontrada puede ser múltiple o estar mal definida.

Segundo, estaríamos viendo si la imagen deconvolucionada estimada,  $o(\vec{f})$ , se parece a lo que buscamos, pero seguiríamos sin tener claro que la red haya calculado los momentos de Zernike correctos.

La respuesta al primer problema podríamos encontrarla en definir mejor la solución del *maximum likelihood* por medio de darle a la red como entrada muchos más pares de imágenes del mismo objeto. Otra solución a probar sería la aplicación de otros métodos, como el *Phase-diverse Speckle*. Estas ideas quedan como propuestas para futuras pruebas.

Respecto al segundo problema, en este caso la degeneración de la *PSF* en cuanto al frente de onda que lo produce juega a nuestro favor, de forma que, si estadísticamente la red logra reconstruir las imágenes dadas para el entrenamiento, querrá decir que la red ha aprendido a obtener la *PSF* más o menos correcta que sufre la imagen, que es el verdadero fin buscamos, y no obtener una combinación explícita de momentos de Zernike.

## Implementación

Todos los códigos usados para este trabajo se escribieron en Python, usando las librerías de Pytorch para aprendizaje automático de código abierto basada en la biblioteca Torch.

Para la optimización de los parámetros de la red se ha usado el método de gradiente estocástico de Adam, con un decaimiento de los parámetros (*weight decay*) de  $10^{-5}$  a modo de regularización, evitando que los valores de parámetros de la red se vuelvan muy pequeños al penalizar estos casos en la evaluación del trabajo de la red.

Por otra parte, el parámetro de aprendizaje (*learning rate*) se dispone inicialmente de  $10^{-4}$ .

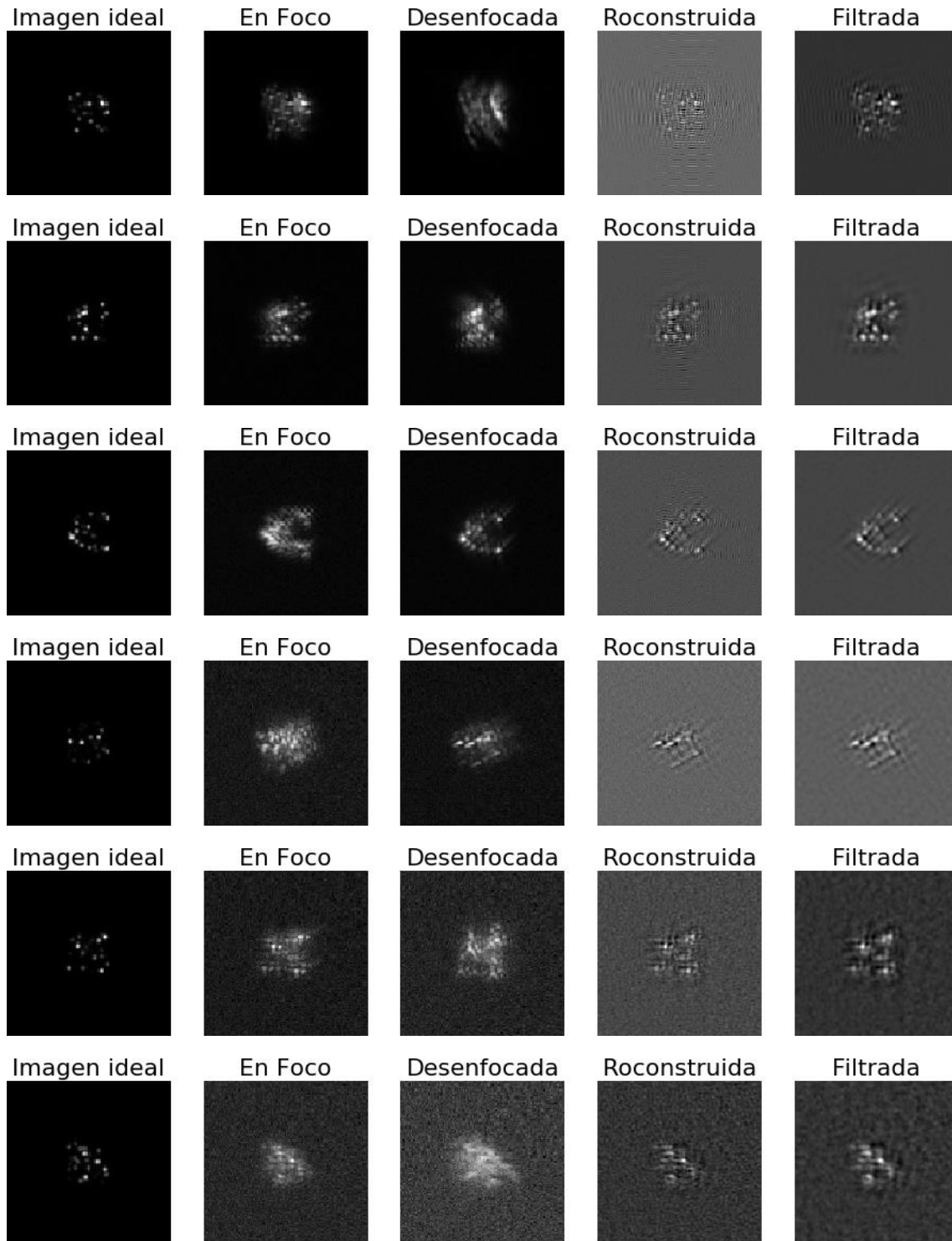
Los códigos fueron ejecutados en un procesador Intel core i7, y una tarjeta gráfica NVIDIA GEFORCE GTX 960M.

Para obtener los parámetros finales del modelo se realizaron varios entrenamientos distintos:

- Entrenamiento con 5000 secuencias de imágenes de misma señal de ruido (baja), durante 40 épocas a un único par de imágenes foco-desenfocado por secuencia. El entrenamiento duró 6 horas.
- Entrenamiento con 1000 secuencias de imágenes, cada una con 20 pares de imágenes foco-desenfocado del mismo objeto, con ruido aleatorio en cada secuencia, durante 40 épocas, actualizando el *learning rate* a la mitad de su valor cada 30 épocas. El entrenamiento duró 5 horas.

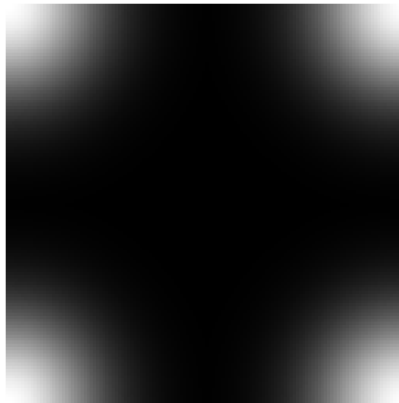
## Resultados

Los resultados obtenidos en el primer entrenamiento muestran:



*“Resultados para distintas intensidades de ruido caracterizado por la varianza de la distribución gaussiana igual a  $S_R 10^{-4}$ . Los valores de  $S_R$  en orden descendente: 0.01, 0.05, 0.1, 0.3, 0.5, 0.7”*

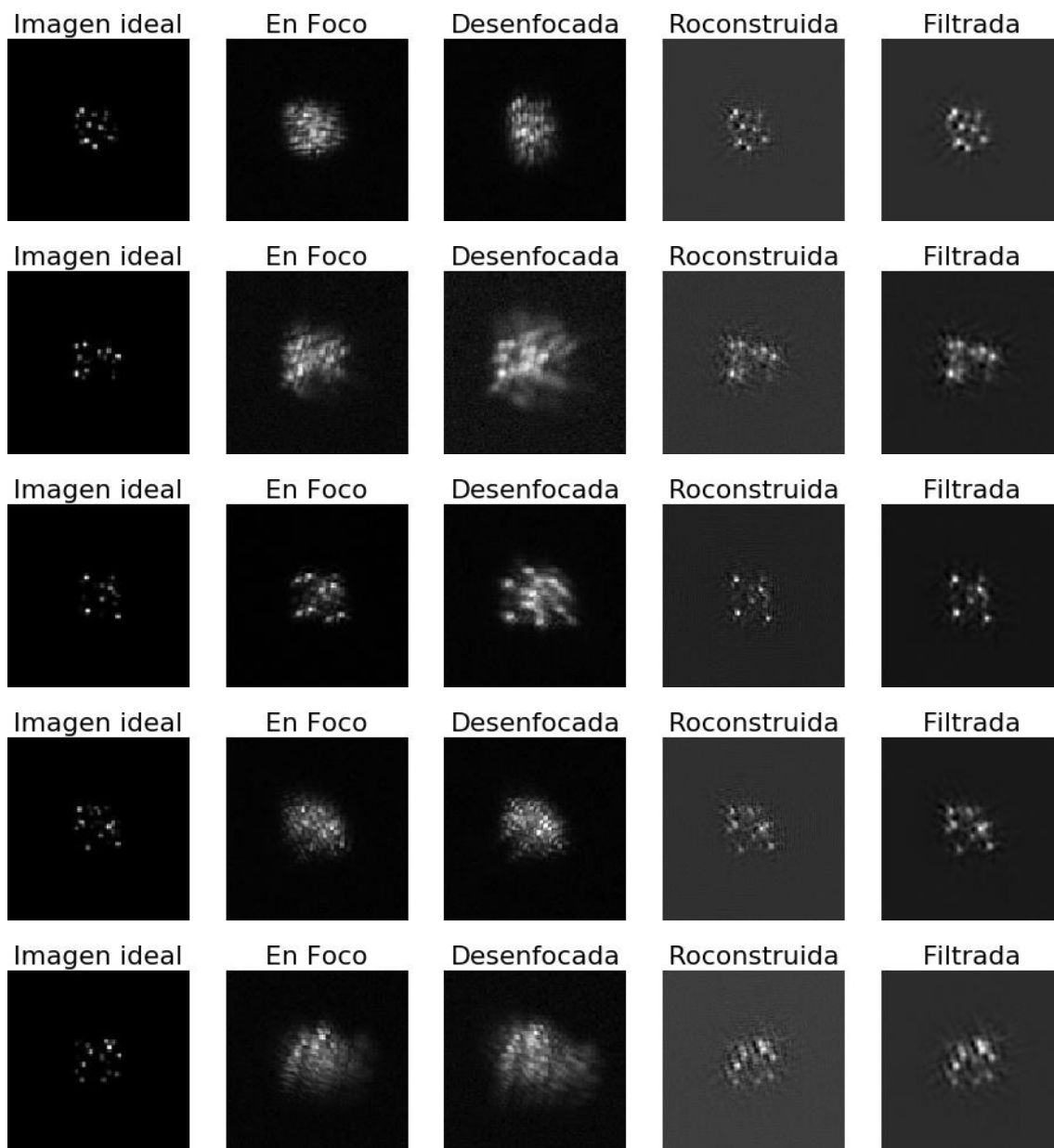
En los resultados se muestran: la imagen ideal a recuperar (recordemos que en nuestro caso disponemos de tal imagen por haber creado las imágenes de forma artificial, pero a priori debería ser desconocida); las imágenes dadas en el tensor de entrada a la red, la imagen en el plano focal del telescopio, y la imagen desenfocada; la imagen que reconstruye la inteligencia artificial a partir del par de imágenes mencionadas; y por último, la imagen reconstruida tras la aplicación de un filtro que intente eliminar las altas frecuencias resultantes en las operaciones. Este filtro aplicado consiste en una gaussiana de media 0 y varianza 0.05.



*“Imagen del filtro aplicado”*

En el caso ideal deberíamos recuperar una imagen en la que solo quedará el efecto de difracción provocado por el telescopio afectando a las estrellas, lo cual sería el límite de resolución. Bien, es obvio que dicho objetivo no se ha cumplido, y que la inclusión de niveles altos de ruido haga que el efecto del filtrado parezca inútil.

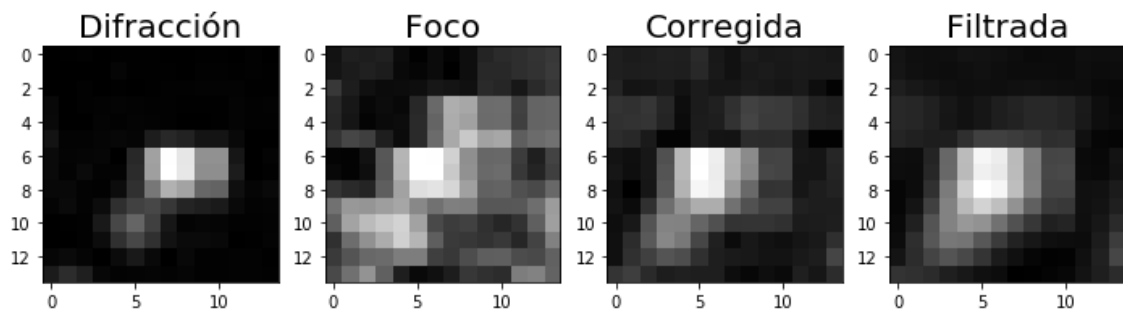
No obstante, el verdadero potencial de este modelo debería aflorar al hacer uso de la *Lstm*, y así lo vemos en los resultados del segundo entrenamiento, al darle como entrada a la red 20 pares de imágenes foco-desenfocada del mismo objeto, e introduciendo además aleatoriedad entre el ruido que presenta cada secuencia de imágenes dentro de un rango de 0.1 hasta 0.4 veces la media de la primera imagen del foco del telescopio de cada secuencia.



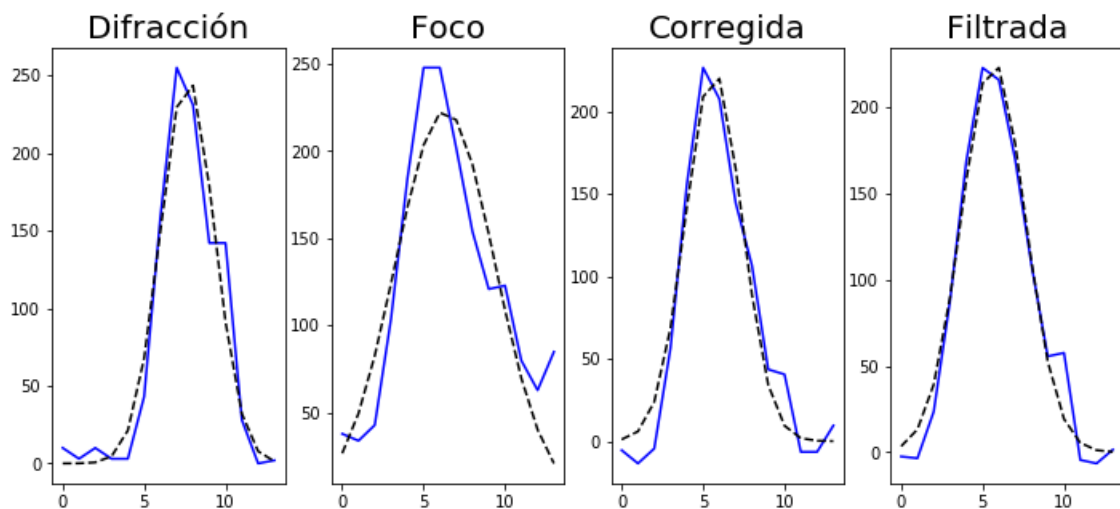
*“Resultados entrenamiento con secuencias de 20 imágenes y ruido aleatorio”*

Puede apreciarse una clara mejora notable respecto al ruido residual de las imágenes resultantes, y como el filtro parece aumentar la relación Señal-Ruido.

Dado que este es el tipo de resultado que esperábamos, conviene realizar un análisis algo más detallado para comprobar que las imágenes resultantes suponen alguna mejora respecto las imágenes de entrada. Para ello se realizará una estadística simple, en la que, para una imagen, seleccionaremos una estrella del campo, y efectuaremos un ajuste gaussiano al perfil de intensidad de esa estrella. Dicho ajuste se realizará para las imágenes de difracción, la imagen del foco, la imagen corregida obtenida por la red, y el filtrado de esta última; y compararemos los parámetros obtenidos del ajuste:



“Recorte de las imágenes entorno a la estrella más brillante”



“Ajuste gaussiano al perfil de intensidades de la estrella más brillante”

Los parámetros del ajuste:

Imágen	Media (S/R)	Varianza
Difracción	$248.63 \pm 16.82$	$2.72 \pm 0.42$
Foco	$223.25 \pm 18.17$	$9.40 \pm 1.82$
Corregida	$224.38 \pm 13.42$	$2.97 \pm 0.41$
Filtrada	$225.99 \pm 9.99$	$3.82 \pm 0.39$

Podemos ver varias cosas a tener en cuenta. Por una parte, queda claro, como se logra estrechar la distribución de intensidad tendiendo a la distribución que encontramos en la imagen de difracción. Lo cual significa que, en efecto, se ha conseguido un efecto de concentrar la luz entorno a los pixeles correctos donde están las estrellas, manteniendo además el valor medio de cuentas (que podemos tomar como indicador del cociente Señal-Ruido, pues se restó el ruido promedio de fondo a las imágenes para hacer este análisis).



En cuanto al filtro utilizado, si parece aumentar el número de cuentas, pero a costa de ensanchar la distribución. Debido a la simpleza del filtro no había gran cosa que esperar, sobre todo cuando (como ha sido el caso de los resultados de este entrenamiento) las imágenes corregidas que devuelve la red presentan esta calidad. Un filtrado más complejo de altas frecuencias pudiera mejorar aún más la imagen, pero de mejorar aún más el rendimiento del modelo tal vez sea innecesario.

Por último, aunque no menos reseñable, parece que sí se ha conseguido que la red sea insensible al ruido. Pues, aunque puede apreciarse una disminución en la señal según aumentamos el ruido, la capacidad de predicción de la *PSF* que afecta a la imagen queda intacta.

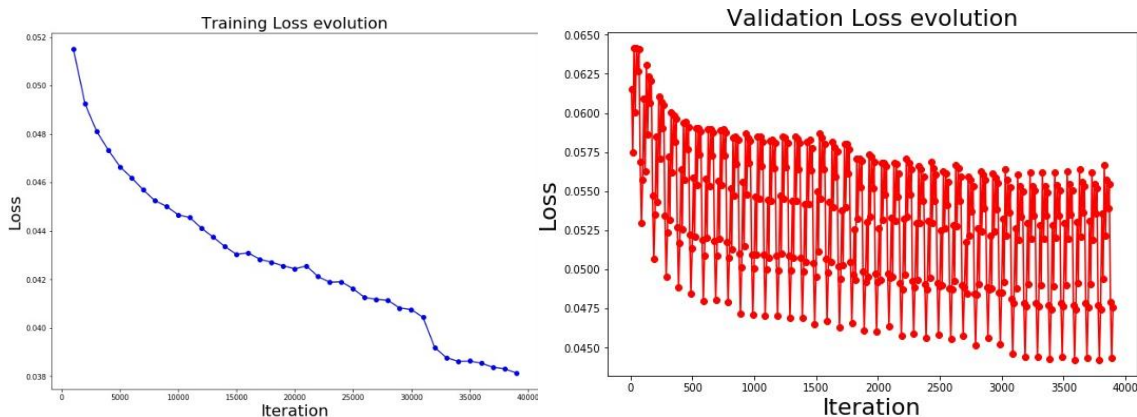
## Conclusiones

No podemos ignorar el dilema de preguntarnos si realmente la red ha aprendido a hacer el trabajo que le pedíamos. Como ya se mostró en la introducción, una red que simplemente tomara como entrada la imagen del foco del telescopio y que mediante un autoencoder generara otra imagen que podríamos comparar con la imagen del objeto solo afectada por la difracción causada por el telescopio (o incluso con la imagen del objeto sin aberración alguna), también disipa el ruido y concentra la luz. Por lo que queda preguntarse ¿Tienen la misma validez ambos métodos? ¿Solucionamos el problema al recurrir a métodos más complejos? La respuesta queda totalmente clara en este trabajo, radica en la esencia de lo queremos conseguir, y marca la gran diferencia entre ambos métodos: Uno trata de recuperar la información de la imagen, el otro solo cambia la apariencia de la imagen.

El resultado obtenido permite afirmar claramente que se pueden corregir imágenes a posteriori añadiendo información extra que nos permita aplicar la física a tecnologías como la inteligencia artificial, capaz de reconocer patrones con precisiones cada vez mayores.

Es por ello, que, aunque las imágenes-resultado obtenidas por la red disten del caso ideal, sí que se consigue volver a concentrar la luz tras paliar el efecto de una distorsión identificada con sentido físico a través de los momentos de Zernike.

Como ya se ha dicho entrenamientos con volúmenes de imágenes mucho mayores que permitan dar secuencias de imágenes mucho más largas, podrían mejorar los resultados. No obstante, si el entrenamiento realizado en este trabajo parece haberse quedado pequeño, se debe, no solo a limitaciones técnicas, sino por la observación de que en entrenamientos mayores el modelo parecía empezar a entrar en régimen de *overfitting*. Esto se observa claramente en la evolución de nuestra función de evaluación tanto en el entrenamiento como en la validación simultánea:



Donde, aunque el error de entrenamiento continúa bajando según continuamos haciendo iteraciones, puede observarse que para algunas imágenes de validación la tendencia empieza a ser ascendente. Punto por el cual se interrumpió el entrenamiento.

El *overfitting*, como suele ser general en el uso inteligencia artificial representa un serio problema al que enfrentarse. Dicho de forma simple, representa un estado del modelo en el que los parámetros de este se ha ajustado hasta tal punto que comienza a perder espacio de predicción para devolver siempre el mismo resultado. Las formas de impedirlo se basan en medidas que aumente la regularización del modelo. Se podría aumentar del decaimiento de los parámetros (*weight decay*), pero se corre el riesgo de limitar la capacidad de predicción; o se puede aumentar el número de parámetros del modelo, pero esto aumenta considerablemente el tiempo necesario para que el modelo converja y comience a “aprender”.

Sin embargo, y como hemos visto eso no ha sido inconveniente para obtener unos resultados claramente satisfactorios y alentadores que dan motivos para continuar con esta línea de trabajo.

## Referencias

### Artículos:

- SOLAR IMAGE RESTORATION BY USE OF MULTI-FRAME BLIND DE-CONVOLUTION WITH MULTIPLE OBJECTS AND PHASE DIVERSITY.  
[Michiel Van Noort, Luc Rouppe Van Der Voort, Mats G. Löfdahl]  
Solar Physics (2005) 228: 191–215
- OBJECT-INDEPENDENT IMAGE-BASED WAVEFRONT SENSING APPROACH USING PHASE DIVERSITY IMAGES AND DEEP LEARNING  
[Qi Xin, Guohao Ju, Chunyue Zhang, Shuyan Xu]  
Vol.27, No.18 | 2 Sep 2019 | Optics Express 26102
- EVALUATION OF PHASE-DIVERSITY TECHNIQUES FOR SOLAR RESTORATION  
[Richard G. Paxman, John H. Seldin, Mats G. Löfdahl, Göran B. Scharmer, Christoph U. Keller]  
The Astrophysical Journal, 466:1087-1099, 1996 Agosto 1
- EXTRAPOLATING ZERNIKE MOMENTS TO PREDICT FUTURE OPTICAL WAVE-FRONT IN ADAPTATIVE OPTICS USING REAL TIME DATA MINING  
[Akondi Vyas, M B Roopashree, B Raghavendra Prasad]  
Journal of Communication and Computer, 8(3):173-179, Enero 2011
- ZERNIKE POLYNOMIALS AND ATMOSPHERIC TURBULANCE\*  
[Rober J. Noll]  
Journal of the Optical Society of America, Vol. 66, Issue 3, pp. 207-211 (1976)
- IMAGE-TO-IMAGE TRANSLATION WITH CONDITIONAL ADVERSARIAL NETWORKS  
[Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros]  
University of California, Berkeley, In CVPR 2017
- ADAM: A METHOD FOR STOCHASTICS OPTIMIZATION  
[Diederik P. Kingma, Jimmy Lei Ba]  
Published as a conference paper at ICLR 2015
- POLINOMIOS DE ZERNIKE Y SU APLICACIÓN EN OFTALMOLOGÍA  
[Marlo Recarte]  
Revista de la escuela de Física, UNAH • Vol. 5, No. 1 • 21-25

Libros:

- Integrated Modeling of Telescopes, 3.6.2 Karhunen-Loève Expansion  
[Torben Andersen, Anita Enmark]

Páginas Web:

- Zernike Polynomials  
[https://en.wikipedia.org/wiki/Zernike\\_polynomials](https://en.wikipedia.org/wiki/Zernike_polynomials)
- Wiener Filtering  
<http://www.owlnet.rice.edu/~elec539/Projects99/BACH/proj2/wiener.html>
- The Wiener Filter  
[https://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/VELDHUIZEN/node15.html](https://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/VELDHUIZEN/node15.html)
- Imaging through turbulence  
<http://www.ctio.noao.edu/~atokovin/tutorial/part1/turb.html#SEC1.1>
- Astronomical seeing  
[https://en.wikipedia.org/wiki/Astronomical\\_seeing](https://en.wikipedia.org/wiki/Astronomical_seeing)
- Pytorch  
<https://pytorch.org/>
- Dot CSV  
<https://www.youtube.com/channel/UCy5znSnfMsDwaLIROnZ7Qbg>

Códigos base:

- Andrés Asensio Ramos