



Universidad de Oviedo
Universidá d'Uviéu
University of Oviedo

Redes neuronales aplicadas a la detección de
candidatos a exoplanetas en series fotométricas
mediante la técnica de tránsito

Lidia Sainz Ledo

Dirigido por
Joaquín González-Nuevo González y Enrique Díez Alonso

UNIVERSIDAD DE OVIEDO
Facultad de Ciencias
Grado en Física

Julio de 2022

Índice general

1. Introducción	5
1.1. Motivación	5
1.2. Objetivos	7
1.3. Estructura del trabajo	7
2. Fundamentos teóricos	8
2.1. Exoplanetas	8
2.1.1. Historia	8
2.1.2. Kepler y K2	12
2.1.3. TESS	15
2.1.4. Zonas habitables	16
2.1.5. Tipos de exoplanetas	19
2.1.6. Conceptos básicos	23
2.1.7. Método de detección por tránsitos	26
2.1.8. Otros métodos de detección	31
2.2. Redes Neuronales	44
2.2.1. Entrenamiento de la red	45
2.2.2. <i>Overfitting</i> y <i>Underfitting</i>	53
2.2.3. Parámetros variables en una red neuronal	56
2.2.4. Normalización de los datos	58
2.2.5. <i>Fine-tuning</i>	59
2.2.6. Redes Neuronales Convolucionales	60
2.2.7. Redes Convolucionales 1D	65
3. Metodología	67
3.1. Generación de curvas de luz artificiales	67
3.1.1. Curvas de luz con tránsito	67
3.1.2. Curvas de luz sin tránsito	72
3.2. Definición Red Neuronal para la detección de exoplanetas	73
4. Resultados y discusión	75
5. Conclusión	82
Bibliografía	84

Índice de figuras

2.1.	[16]. Número de detecciones acumulativas cada año según el método de detección.	11
2.2.	[17]. Futuras misiones en la búsqueda de exoplanetas. Imagen original de: NASA/JPL/Caltech.	12
2.3.	[19] Esquema del telescopio Kepler.	13
2.4.	[23] Esquema de las 19 campañas de la misión extendida K2.	14
2.5.	[27] Esquema del estudio de la bóveda celeste en la misión TESS. . . .	16
2.6.	[29, Fig. 3] Diagrama esquemático de los límites de las distancias orbitales de la zona habitable para estrellas con diferente masa. Las líneas roja y naranja muestran los límites optimistas interior y exterior de la zona habitable, mientras que las azules muestran los conservativos. Se representan varios planetas como referencia. Créditos: Chester Harman.	18
2.7.	[30] Representación de la zona habitable galáctica.	18
2.8.	Representación gráfica de el hueco de Fulton. Fuente: [36]. Créditos: NASA/Ames Research Center/CalTech/University of Hawaii/B.J. Fulton.	22
2.9.	[40] Esquema de las estrellas de la secuencia principal.	23
2.10.	Esquema del plano orbital de un planeta que orbita una estrella con respecto al plano del cielo.	26
2.11.	[43] Esquema del plano orbital con respecto al plano de referencia: <i>face-on</i> vs <i>edge-on</i>	26
2.12.	[46, p. 56] Detección de exoplanetas con Tránsito.	27
2.13.	[46, p. 56] Esquema de un tránsito.	28
2.14.	Esquema de espectroscopía de transmisión para un planeta como la Tierra. Fuente: [49]. Créditos: NASA, ESA, CSA, STScI, Joseph Olmsted (STScI).	30
2.15.	[52] Esquema de la variación en el espectro de una estrella debido al efecto Doppler, causado por la presencia de un planeta orbitando la estrella.	32
2.16.	[54] Variación de la señal dependiendo de ϵ y ω	33
2.17.	[57] Detección de exoplanetas con Astrometría.	35
2.18.	[60, p. 5] Esquema de la óptica básica en microlentes gravitacionales. .	36
2.19.	[61] Imágenes observadas debido a la flexión en la trayectoria del haz causada por una estrella.	37
2.20.	Variación del brillo de una estrella al pasar entre la estrella y el observador otra estrella (la lente) y un exoplaneta que la orbita. Fuente: [62]. Adaptada de la web de David Bennett.	38

2.21. Sistema planetario entorno a la estrella HR 8799, detectados con el método de imagen directa. Fuente: [65]. Créditos: Jason Wang y Christian Marois.	40
2.22. [67] Curva de luz de una binaria eclipsante observada por el satélite CoRoT. Crédito: CNES.	41
2.23. [75] Imagen tomada por ALMA del disco entorno a la estrella HL Tau.	43
2.24. Esquema básico de una red neuronal. Imagen realizada usando <i>Paint</i> y la aplicación <i>FlowdiaDiagrams</i>	45
2.25. Esquema básico de una red neuronal sobre el que se indican los pesos de ciertas conexiones, para ver la dependencia del término 1 de la ecuación 2.31. Imagen realizada usando <i>Paint</i> y la aplicación <i>FlowdiaDiagrams</i>	49
2.26. Explicación esquemática funcionamiento del sesgo con una función de activación ReLU. Imagen realizada con la ayuda de <i>Python</i> , <i>Paint</i> , la aplicación <i>FlowdiaDiagrams</i> y con la figura de la fuente [78].	53
2.27. Representación de tres tipos de funciones de activación: Sigmoid, Tanh y ReLU. Imagen realizada con <i>Python</i>	57
2.28. Esquema de rasgos reconocidos por los filtros al profundizar en más capas convolucionales. Figura realizada con la aplicación <i>Paint</i> , empleando imágenes de las siguientes fuentes: [81] [82] [83].	61
2.29. Aplicación de un filtro sobre una imagen de entrada. Imagen realizada con las aplicaciones <i>Write</i> y <i>Paint</i>	62
2.30. Esquema del funcionamiento del <i>zero padding</i> . Imagen realizada con las aplicaciones <i>Write</i> y <i>Paint</i>	63
2.31. Esquema del funcionamiento de <i>Max Pooling</i> . Imagen realizada con las aplicaciones <i>Write</i> y <i>Paint</i>	64
2.32. Esquema representativo de una red convolucional 1D construida por un filtro 3x1, siendo aplicada sobre una serie temporal con un stride igual a 1. Imagen realizada con las aplicaciones <i>Write</i> y <i>Paint</i>	65
3.1. Relación entre el ruido de los datos en la curva de luz y la magnitud aparente de la estrella. Regresión lineal realizada a partir de la Figura 5 de [87].	70
3.2. Curva de luz de una estrella que presenta tránsitos correspondientes a un exoplaneta. Las características del sistema estrella-exoplaneta son las siguientes: $M = 1.135M_{\odot}$, $R = 1.100R_{\odot}$, $rp = 0.064$, semi-eje = $2.380R_{\odot}$, inclinación orbital = 89.526° , excentricidad = 0.093, $T = 12.682$ días. La estrella tiene una magnitud aparente = 12.300 y un ruido gaussiano = 63.954ppm.	71
3.3. Curva de luz de una estrella que presenta tránsitos correspondientes a un exoplaneta. Las características del sistema estrella-exoplaneta son las siguientes: $M = 2.317M_{\odot}$, $R = 1.878R_{\odot}$, $rp = 0.041$, semi-eje = $18.313R_{\odot}$, inclinación orbital = 90.544° , excentricidad = 0.029, $T = 5.990$ días. La estrella tiene una magnitud aparente = 16.000 y un ruido gaussiano = 372.648ppm.	71

3.4.	Curva de luz de una estrella que presenta tránsitos correspondientes a un exoplaneta. Las características del sistema estrella-exoplaneta son las siguientes: $M = 0.586M_{\odot}$, $R = 0.670R_{\odot}$, $rp = 0.054$, semi-eje = $29.530R_{\odot}$, inclinación orbital = 91.810° , excentricidad = 0.095, $T = 24.390$ días. La estrella tiene una magnitud aparente = 15.800 y un ruido gaussiano = 338.785ppm.	72
3.5.	Curva de luz correspondiente a una estrella sin tránsito, con una magnitud aparente = 16.800 y un ruido gaussiano = 545.504ppm.	73
3.6.	Esquema de los datos que toman los parámetros en el entrenamiento de la red. Imagen realizada con la aplicación <i>FlowdiaDiagrams</i>	74
4.1.	Representación de la evolución de <i>train accuracy</i> y <i>val accuracy</i> con los <i>epochs</i> . Modelo con parámetros (3,64,11,1).	78
4.2.	Representación de la evolución de <i>train accuracy</i> y <i>val accuracy</i> con los <i>epochs</i> . Modelo con parámetros (5,64,7,1).	78
4.3.	Representación de la evolución de <i>train accuracy</i> y <i>val accuracy</i> con los <i>epochs</i> . Modelo con parámetros (3,16,7,3). Se representa en rojo la precisión en la que se estabiliza el modelo.	79
4.4.	Representación de la evolución de <i>train loss</i> y <i>val loss</i> con los <i>epochs</i> . Modelo con parámetros (3,16,7,3). Se representa en rojo el valor de la función coste en el que se estabiliza el modelo.	79
4.5.	Representación de los aciertos y fallos en la detección de exoplanetas con el modelo (3,16,7,3), en función del periodo orbital.	79
4.6.	Representación de los aciertos y fallos en la detección de exoplanetas con el modelo (3,16,7,3), en función del cociente <i>rp</i>	80

Capítulo 1

Introducción

1.1. Motivación

Son muchas las preguntas sobre la existencia las que gobiernan al ser humano, preguntas tales como de dónde venimos y a dónde vamos, o si estamos sólo en el Universo. A pesar de que ninguna de ellas ha sido resuelta científicamente con la precisión deseada, cada vez es más el conocimiento que se tiene al respecto, debido a los avances en la ciencia y en la tecnología.

Todas estas incógnitas son las que han movido siempre a los investigadores, llevándoles a desear conocer otros cuerpos diferentes a la Tierra. La primera sonda en explorar otro cuerpo fue Luna 1, la sonda soviética que sobrevoló nuestro satélite en Enero de 1959. Unos años más tarde, en 1961, la URSS logra sobrevolar un planeta distinto de la Tierra, Venus, con la sonda Venera 1. Estas metas se consiguen alcanzar con otros planetas posteriormente, durante la década de los años 70 [1, pp. 20-21].

Dentro del Sistema Solar, destaca Marte por tener probabilidad, por pequeña que sea, de albergar o haber albergado vida en algún momento, debido a su distancia con respecto al Sol, a la presencia de agua y a las características de la superficie. Tras conocer los planetas que nos rodean, llega un momento en el que se quieren superar las fronteras del Sistema Solar, y la tecnología lo permite. Además de nuestro Sol, existen 100000 millones de estrellas en nuestra Galaxia, por lo que la probabilidad de encontrar más sistemas planetarios y alguno en la zona habitable, parecía ser muy alta [2].

Así, en el año 1995, se detecta el primer exoplaneta alrededor de otra estrella de Secuencia Principal diferente al Sol, 51 Pegasi. A partir de entonces, la búsqueda de exoplanetas no ha hecho más que avanzar, conociendo a día de hoy 5054 planetas confirmados fuera de nuestro Sistema Solar [3] [4].

Gracias al avance tecnológico, hoy en día se dispone de una gran cantidad de datos que han logrado obtenerse a partir de diferentes métodos según el ámbito en el que se esté trabajando. Sin embargo, la posesión de dichos datos no aporta ninguna información si estos no se tratan. De tal forma que, para obtener beneficio de este logro, se han de desarrollar métodos que nos permitan elaborar conclusiones a partir de ellos [5, pp. 1-2].

Gran parte de los datos que resultan de un estudio, siguen un patrón que es precisamente lo que se quiere encontrar. El desarrollo del aprendizaje automático o *machine learning* ha sido clave en este campo, destacando la parte de minería de datos, es decir, “la aplicación de *machine learning* a grandes bases de datos” [5, pp. 1-2]. Una cuestión fundamental en esta línea de trabajo es la toma de decisiones, y es aquí donde interviene el *procedimiento de clasificación* [6, p. 1].

Son numerosas las ocasiones en las que los seres humanos clasifican inconscientemente en su día a día [7]. Simplemente, al ver a una persona, por sus rasgos se intuyen características como el país de procedencia o la edad. Lo mismo pasa al ver cualquier otra entidad que se pueda dividir en diferentes clases.

En la ciencia, estos procedimientos son fundamentales. Por ejemplo, en el campo de la medicina, el diagnóstico de un enfermo puede realizarse a partir de la experiencia pasada y las características de su enfermedad [5, p. 2]. En astrofísica, a partir de las propiedades de un exoplaneta se puede intuir si es habitable o no, comparándolas con las del Planeta Tierra. En este último campo, también puede detectarse la presencia de un exoplaneta orbitando una estrella dependiendo de las características de la curva de luz registrada.

Se han desarrollado varios métodos para llevar a cabo el procedimiento de clasificación, entre los que se encuentran las redes neuronales, un método supervisado que consiste en entrenar a la red a partir de datos de entrenamiento para obtener un modelo que sea capaz de predecir la clase de un objeto que no haya visto antes.

1.2. Objetivos

Entre los objetivos de este trabajo destacan los siguientes:

- Estudiar la historia de la investigación de los exoplanetas, así como las misiones más importantes en este campo de la ciencia.
- Conocer los tipos principales de exoplanetas.
- Estudiar los diferentes métodos de detección de los exoplanetas, destacando el método de detección por tránsitos.
- Desarrollar una red neuronal capaz de identificar la presencia de exoplanetas en curvas de luz con y sin tránsitos, creando de forma sintética los datos necesarios para su entrenamiento.

1.3. Estructura del trabajo

En base a los objetivos del trabajo, la estructura de este será la siguiente: el capítulo 2 comienza desarrollando la historia de detección de exoplanetas, destacando las misiones Kepler y TESS, por marcar un cambio importante en este campo de la ciencia. Se presentan también los diferentes tipos de exoplanetas y se definen las zonas habitables estelares y galácticas.

A continuación, se estudian los diferentes métodos de detección, comenzando con el que es de principal interés en el trabajo, el método de detección por tránsitos.

En la segunda parte de este capítulo se introducen las redes neuronales. Se desarrolla el proceso de entrenamiento de una red y se tratan diferentes aspectos de este modelo computacional, desde los diferentes problemas que se pueden encontrar en la programación de la red, hasta los parámetros variables de un modelo y diferentes tipos de red. Se termina el capítulo con las redes neuronales convolucionales y, en particular, las de una dimensión, debido a que será la que se construya en este trabajo.

En el tercer capítulo se expone la metodología seguida para la generación de curvas de luz artificiales y la definición de la red neuronal. Por último, en el capítulo 4 se analizan los resultados y el rendimiento del modelo obtenido.

Capítulo 2

Fundamentos teóricos

2.1. Exoplanetas

2.1.1. Historia

En un principio, tal y como es característico del egocentrismo del ser humano, se pensaba que la Tierra era el centro del Universo. Esta idea fue formalizada por Aristóteles en el año 340 a.C., que defendió la Tierra como centro de un Sistema, alrededor de la cual giraban el resto de planetas, el Sol y la Luna, así como todas las estrellas fijas [8].

Más tarde, en el año 1543, Nicolás Copérnico, propuso un modelo más cercano a la realidad basado en la idea de que la Tierra no era el centro del Universo, como ya había razonado Aristarco de Samos en el año 200 a.C. Sentenciaba que el cuerpo estático y central era el Sol, mientras que los planetas orbitaban a su alrededor [8].

Así, la idea de ser el centro del Universo fue quedando atrás, y se empezó a reflexionar sobre la posibilidad de existencia de otros sistemas estelares y de más planetas. En esta línea de pensamientos destaca Giordano Bruno, quien no solo defendía a Copérnico, sino que añadía a su sistema nuevas ideas, como el movimiento del Sol respecto a otros astros. Giordano predicaba también que existían muchos más sistemas estelares como el nuestro, con posible vida y formando un universo infinito y sin centro. Estas ideas en su época le llevaron a ser condenado a morir en la hoguera, el 17 de Febrero de 1600 [9].

En 1609, Galileo Galilei, fabricó el telescopio con el que pudo llevar a cabo las primeras observaciones del cielo, descubriendo las lunas de Júpiter, que giraban a su alrededor en una órbita circular. Así, el descubrimiento de un nuevo sistema con su propio centro, cambió definitivamente la forma de entender la Astronomía [8].

Uno de los objetivos principales del ser humano es saber si existe vida extraterrestre, por lo que no solo intrigaba la idea de presencia de exoplanetas (planetas fuera del Sistema Solar), sino que se pensaba en una posible vida en ellos. Esta intriga sigue hoy en día moviendo a los investigadores, que intentan entender la formación del Universo, su disposición y llegar a la evidencia de vida en aquellos planetas que parecen habitables.

Fue en 1992 cuando se descubrió el **primer exoplaneta**, aunque este no orbitaba una estrella de la secuencia principal. La estrella materna, PSR B1257+12, es una **estrella de neutrones** y, en particular, un púlsar. Este sistema tiene dos planetas, los cuales fueron detectados por los astrónomos Aleksander Wolszczan y Dale Frail, empleando el método de la velocidad radial [10].

El **primer exoplaneta alrededor de una estrella de la secuencia principal (51 Pegasi)** fue descubierto en 1995 por Didier Queloz y Michel Mayor, empleando también la velocidad radial. El planeta está a 50 años luz de la Tierra, y se denomina 51 Pegasi b, siguiendo la norma de nomenclatura, que consiste en llamarlos con el nombre de la estrella seguido de una numeración alfabética (que comienza con la letra *b*, reservando la *a* para la estrella). Una propiedad muy sorprendente fue la masa del planeta, pues es aproximadamente la mitad del de Júpiter y no se entendía cómo, con esa masa, podía producir esos efectos sobre la estrella. La respuesta a esta cuestión es la distancia de su órbita al planeta, que es de 0.05 AU, lo que produce unas variaciones en la velocidad radial de una amplitud media de 56m/s [11, p. 33-34] [3]. Las características del planeta llevaban a clasificarlo como gigante gaseoso, sin embargo, el periodo de 4,23 días (en vez de años, como era habitual en estos casos), y la pequeña distancia a la estrella madre, daban lugar a un tipo diferente de exoplaneta nunca antes visto: un Júpiter caliente (ver sección 2.1.5 *Tipos de exoplanetas*) [12].

Tres años más tarde se descubrió el **primer planeta alrededor de una enana roja** (ver sección 2.1.6 *Conceptos básicos*), Gliese 876 b, que formaba parte de un

sistema que después revelaría que posee al menos cuatro planetas en total [13, p. 18]. Pasado un año, dos equipos de investigación (liderados por David Charbonneau y Greg Henry) detectaron de forma independiente el **primer exoplaneta por el método de tránsito**, HD 209458 b, el cual ya había sido hallado a través de la velocidad radial [3] [13, p. 18].

En 1999 otros dos equipos de investigación (San Francisco State University y Harvard-Smithsonian Center for Astrophysics), de forma independiente también, observaron el primer sistema con varios planetas (tres), los cuales orbitan la estrella Upsilon Andromedae [3].

En Abril de 2001 se detecta el primer planeta que orbita a la misma distancia que la Tierra lo hace del Sol. Es el caso de HD 28185 b, que es aproximadamente seis veces más masivo que Júpiter. Dos años más tarde se producen dos lanzamientos de telescopios. En Junio, se lanza **MOST** (*Microvariability and Oscillations of Stars*), un telescopio pequeño creado en Canadá, cuya función principal era detectar variaciones en el brillo de exoplanetas. En Agosto, se lanzó **Spitzer Space Telescope**, que se emplearía para detectar radiación infrarroja; sus aplicaciones en el campo de los exoplanetas van dirigidas a conocer datos de sus atmósferas y tamaños [3] [14]. Por otro lado, se descubre ese mismo año el primer exoplaneta a través del método de tránsito, OGLE-TR-56 b [13, p. 19].

Posteriormente, en Diciembre de 2006, se lanza otro satélite, el **CoRoT**, cuyo nombre proviene de “convección”, “rotación” y “tránsitos planetarios”, haciendo así referencia a las principales ideas de investigación que se perseguía con su lanzamiento [3] [15]. Poco después, en Febrero de 2007, se estudió por primera vez el espectro de dos planetas, HD 209458 b y HD 189733 b, observados ambos espectros con el telescopio Spitzer Space Telescope. En Mayo de este mismo año, a través del mismo telescopio, David Charbonneau y Heather Knutson consiguen crear el primer mapa térmico de un exoplaneta, de nuevo el HD 189733 b [3].

Al año siguiente, y siguiendo con descubrimientos en el planeta HD 189733 b, se consigue detectar materia orgánica (metano) por primera vez en un planeta fuera del Sistema Solar [13, p. 18].

En el año 2009 comienza una de las misiones más importantes en la historia de la

Astrofísica: la misión **Kepler**, la cual marca un antes y después en la investigación de exoplanetas. Una misión posterior, fundamental también en este campo, es la misión **TESS**. Ambos proyectos, Kepler y TESS, se desarrollan en dos secciones posteriores, profundizando más en ellas.

Hasta el día de hoy, se han detectado 5054 exoplanetas confirmados. En la tabla 2.1 se puede observar el número de planetas descubierto con cada uno de los métodos, destacando el método de tránsito, que será en el que se centre este trabajo.

Método	Número de exoplanetas
Astrometría	1
Imágenes	59
Velocidad Radial	928
Tránsito	3858
Variaciones en el tiempo de tránsito	22
Variaciones en el tiempo de eclipse	18
Microentes Gravitacionales	130
Variaciones en el tiempo pulsar	7
Variaciones en el tiempo pulsación	2
Modulaciones brillo orbital	9
Cinemática del disco	1
Polarización luz estelar	0

Tabla 2.1: Número de exoplanetas descubiertos hasta ahora con cada método. Datos obtenidos de [4] y [11].

La gran diferencia entre el éxito del método de tránsito y el resto se debe a la misión Kepler, como puede confirmarse a partir de la Figura 2.1, donde se observa que, a partir del año en el que comienzan a registrarse datos gracias a esta misión, aumenta bruscamente el número de planetas.

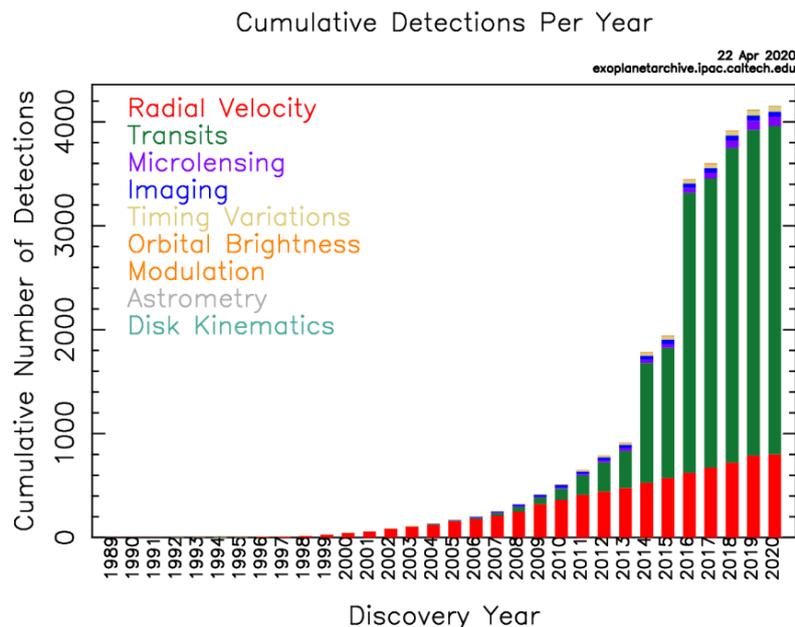


Figura 2.1: [16]. Número de detecciones acumulativas cada año según el método de detección.

Por último, se presenta a continuación un esquema de las misiones futuras que se

harán en este campo:

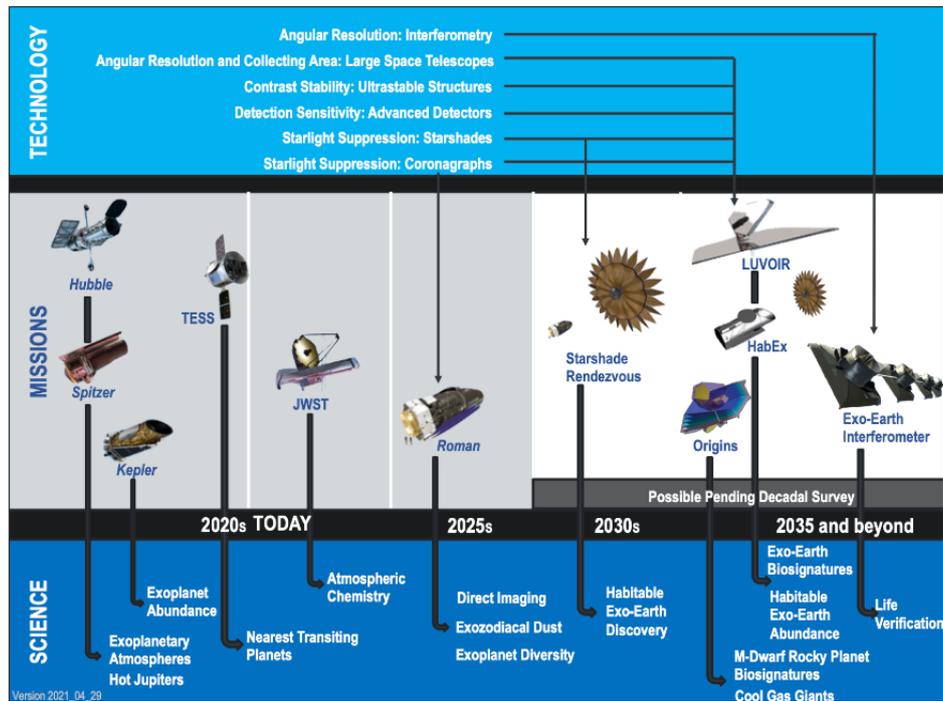


Figura 2.2: [17]. Futuras misiones en la búsqueda de exoplanetas. Imagen original de: NASA/JPL/Caltech.

Los objetivos de las misiones futuras se basan en avanzar en el estudio de las características de los exoplanetas, especialmente de sus atmósferas, centrándose siempre en la búsqueda de planetas habitables. Para ello, las principales mejoras tecnológicas se quieren realizar en la precisión angular de detección y en el contraste de las imágenes, así como en el aumento de la estabilidad de los telescopios espaciales [17].

2.1.2. Kepler y K2

El telescopio Kepler fue lanzado en 2009 desde Cabo Cañaveral, en Florida. La misión con el nombre de Johannes Kepler (1571-1630), gran astrónomo y matemático, fue llevada a cabo principalmente por tres entidades: *Jet Propulsion Laboratory* estudió su puesta en órbita, *Ball Aerospace* se dedicó al sistema de vuelo, y el Centro de Investigación de *Ames* desarrolló su sistema a tierra. [18]

Kepler tenía una masa de 1039 kg y orbitaba alrededor del Sol con un periodo de **372.3 días**, definiendo una órbita elíptica a una distancia aproximada de **0.5 AU de la Tierra**. Se diferencian dos partes principalmente: la **nave espacial** y el **fotómetro**.

La nave, con forma hexagonal, estaba formada por todos los dispositivos que permiten la navegación, la transmisión de datos o la propulsión (con hidrazina como propelente) [13, pp. 86-87] [18] [19].

El fotómetro es la otra parte que constituye el telescopio, y estaba formado por elementos que le capacitaban para detectar variaciones en el brillo de las estrellas. Destaca la pantalla solar, que ayudaba a la precisión de las mediciones impidiendo la posible la entrada de luz solar; así como el Corrector Schmidt, con 0.95m de apertura y un espejo primario de 1.4m. La Cámara CCD constaba de 95 millones de píxeles y barría un área de 100 grados [13, p. 86] [18].

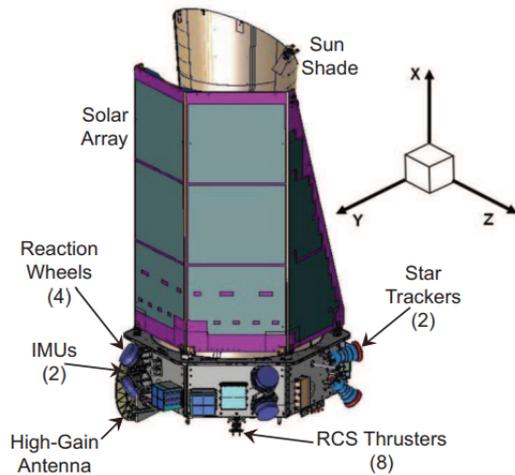


Figura 2.3: [19] Esquema del telescopio Kepler.

El principal objetivo de esta misión era encontrar exoplanetas de tamaño similar a la Tierra y presentes en la zona habitable por medio del **método de tránsito**. Además, se deseaba **monitorear la mayor cantidad de estrellas posibles al mismo tiempo**, de manera que se eligió la región de las constelaciones **Cygnus y Lyra**, cubriendo una cantidad de **150000 estrellas**. Por otra parte, el campo de visión del telescopio no podía ser bloqueado durante el estudio, para poder captar el brillo de la forma más continuada posible, por lo que debía observar fuera del plano definido por la órbita eclíptica [20].

La misión Kepler se llevó a cabo sin ningún problema hasta el año 2013. En Julio de 2012 dejó de funcionar una de las cuatro ruedas giroscópicas que poseía, y en Mayo de 2013 dejó de funcionar otra. Estos volantes de inercia o giroscopios permitían la orientación y la estabilidad del telescopio. Sin embargo, la misión pudo continuar con una solución que llevó a lo que se conoce como **misión extendida K2, o la segunda parte de la misión Kepler con dos ruedas** [19] [21].

El principal objetivo de K2 era poner **solución a la pérdida de estabilidad del telescopio** y poder seguir llevando a cabo el estudio. Para ello, se empleó la presión de la radiación solar.

Con esta nueva misión, lejos de apuntar siempre a la misma región, el telescopio recorre **18 campos diferentes** (18 campañas completas) situadas entorno a la órbita eclíptica, recogiendo datos en cada una de ellas durante aproximadamente **80 días**. Cabe destacar que, a diferencia de la misión Kepler original, la cual tomaba datos con una cadencia de 2 minutos, en la versión extendida K2, la gran mayoría de estrellas son estudiadas con una cadencia de 30 minutos¹. Las observaciones finalizaron en Octubre de 2018, debido a la falta de hidrazina, mientras llevaba a cabo la décimo-novena campaña [22].

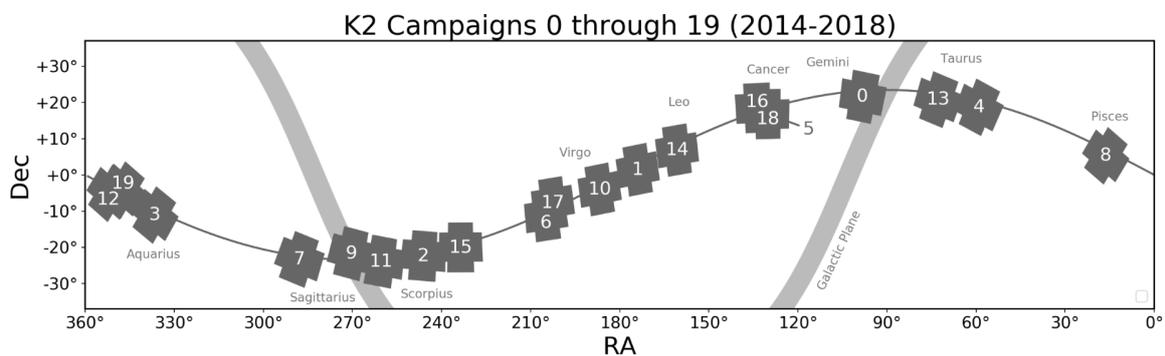


Figura 2.4: [23] Esquema de las 19 campañas de la misión extendida K2.

La información obtenida con K2 cubre también otros objetivos diferentes de los alcanzados con la primera parte de la misión Kepler. A partir de este proyecto, se profundiza en la investigación de cúmulos jóvenes, estrellas brillantes, y planetas habitables alrededor de enanas M. Se avanzó también en el estudio de sistemas binarios y en la variabilidad de AGN, así como en la luminosidad de supernovas de tipo Ia. Por otra parte, ayudó también a avanzar en el entendimiento de la formación planetaria [20] [22].

¹Aunque en la versión extendida K2 la gran mayoría de estrellas se estudien con una cadencia de 30 minutos, para cada campaña se tenían ciertos *targets* cuya fotometría se daba con una cadencia de 2 minutos.

2.1.3. TESS

Tras la misión Kepler, se aprobó otra misión fundamental en la búsqueda de exoplanetas, el satélite *Transiting Exoplanet Survey Satellite* (TESS), que fue lanzado el 18 de Abril de 2018 y la misión fue dirigida por el MIT (*Massachusetts Institute of Technology*).

El estudio con esta misión se centra en **planetas de tamaño menor que Neptuno**, basándose en el **método de detección de tránsito**, al igual que la misión Kepler. Se busca en esta ocasión poder llevar a cabo **investigaciones detalladas**, así como la medición de las masas de los exoplanetas y la composición atmosférica. Para ello, se necesita que las **estrellas sean brillantes**, entre 30 y 100 veces más que las estudiadas en la misión anterior [24].

Para la definición de la órbita que sigue el satélite se tuvieron en cuenta muchos aspectos. Se deseaba evitar zonas de alta radiación, además de prevenir eclipses de larga duración. Por otra parte, con el fin de simplificar la máximo posible la transmisión de los datos, se necesitaba que en alguna región de la órbita el telescopio se acercara lo máximo posible a la Tierra. Por todo ello, la órbita es elíptica con un **perigeo y apogeo de $17R_{\oplus}^2$ y $59R_{\oplus}$** respectivamente. Presenta una inclinación de 37° con respecto al plano de la eclíptica y describe un periodo de **13.7 días**, presentando una resonancia 2:1 con la Luna³. Otro factor que hace posible la estabilidad del telescopio es el hecho de que la Luna está en fase de 90° con respecto al apogeo de la órbita, lo que promedia las perturbaciones [27, pp. 5, 8] [28].

El hecho de que las estrellas más brillantes estén distribuidas uniformemente en el cielo, lleva al deseo de que TESS estudie **todo el cielo**. Para ello, se dividirá la bóveda celeste en **26 segmentos**, con unas dimensiones de $24^{\circ} \times 96^{\circ}$, estudiando cada uno de

²El símbolo R_{\oplus} se emplea para denotar *radios terrestres*.

³Cuando dos cuerpos poseen una relación entre sus periodos orbitales de números enteros (entorno a otro cuerpo central en común), se da el fenómeno de resonancia [25]. Esto ocurre también en nuestro Sistema Solar. Es el caso de Plutón y Neptuno, cuya resonancia es 2:3, lo que significa que 2 vueltas de Plutón entorno al Sol se corresponden con 3 de Neptuno [26].

ellos durante 27 días. Dependiendo de la situación del sector el tiempo de su estudio puede aumentar, debido a superposiciones en los polos de la eclíptica, alcanzando los 351 días en los puntos más extremos (ver Figura 2.5).

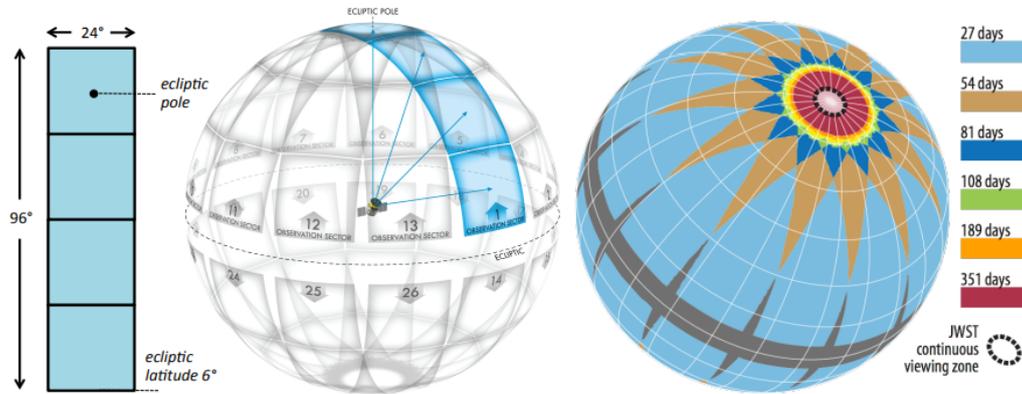


Figura 2.5: [27] Esquema del estudio de la bóveda celeste en la misión TESS.

El rango de estrellas que se estudian es, dentro de la **secuencia principal, F5-M5**. Aproximadamente **200000** estrellas componen la *Candidate Target List (CTL)*, en una región del cielo 400 veces mayor que la estudiada por Kepler, cuya fotometría se obtiene con una cadencia de 2 minutos. Por otra parte, se obtienen las *Full Frame Images (FFI)* de millones de estrellas con una cadencia de 30 minutos [24].

La inspección de TESS estaba pensada para durar dos años, sin embargo, en 2020 comenzó la misión extendida de TESS, que durará 27 meses más. Actualmente hay 227 planetas confirmados descubiertos en esta misión, además de 3854 planetas pendientes de confirmación [4] [27].

2.1.4. Zonas habitables

En esta sección se describe el tipo de zonas donde se puede llegar a encontrar vida. Para ello, se emplea la definición clásica de *zona habitable*, basada en las características de la Tierra, puesto que es el único planeta con vida que se conoce.

Un factor primordial en el estudio de una posible vida en un planeta es la presencia de agua en estado líquido, debido a sus características favorables para que se den reacciones entre moléculas [13, pp. 64-65]. Por ello, la distancia a la que se encuentre el planeta de la estrella es muy importante. Si la distancia en cuestión es demasiado

pequeña, entonces el planeta puede perder todo el agua debido a su gran exposición a los efectos de la estrella; esta zona se denomina el borde interior conservador de la zona habitable (IHZ). Por otra parte, en el borde exterior conservador de la zona habitable (OHZ), se alcanza el límite máximo de CO₂ en el efecto invernadero.

Sin embargo, no solo se debe tener en cuenta la distancia entre el planeta y la estrella, sino que muchos otros factores influyen en la existencia de agua líquida en la superficie, como la temperatura y la presión. Los cambios en la temperatura se producen con cada desequilibrio entre la energía estelar que llega al planeta y la radiación térmica que este expulsa. Suele expresarse el balance energético en términos de *la temperatura de equilibrio*, que es “la temperatura del planeta asumiendo que está irradiando al espacio toda la radiación que recibe, como un cuerpo negro” [29, pp. 13-14].

La presencia de la atmósfera en un planeta influye directamente en la posibilidad de que sea o no habitable, debido a que lleva consigo un efecto invernadero mayor o menor, que permite elevar la temperatura por encima a la de equilibrio, de forma que puede alcanzarse la necesaria para la presencia de agua en estado líquido [29, p. 14].

Como se mencionó al comienzo de este apartado, las características desarrolladas hasta el momento se consideran favorables debido a su comparación con la situación de la Tierra; sin embargo, otras composiciones atmosféricas y presiones en la superficie podrían llevar también a planetas habitables. Por ejemplo, en planetas que giran sincrónicamente o que están a una distancia grande de su estrella, las presiones altas podrían llevar a superficies habitables debido a una disminución de la parte del planeta que de otro modo estaría congelado [29, p. 14].

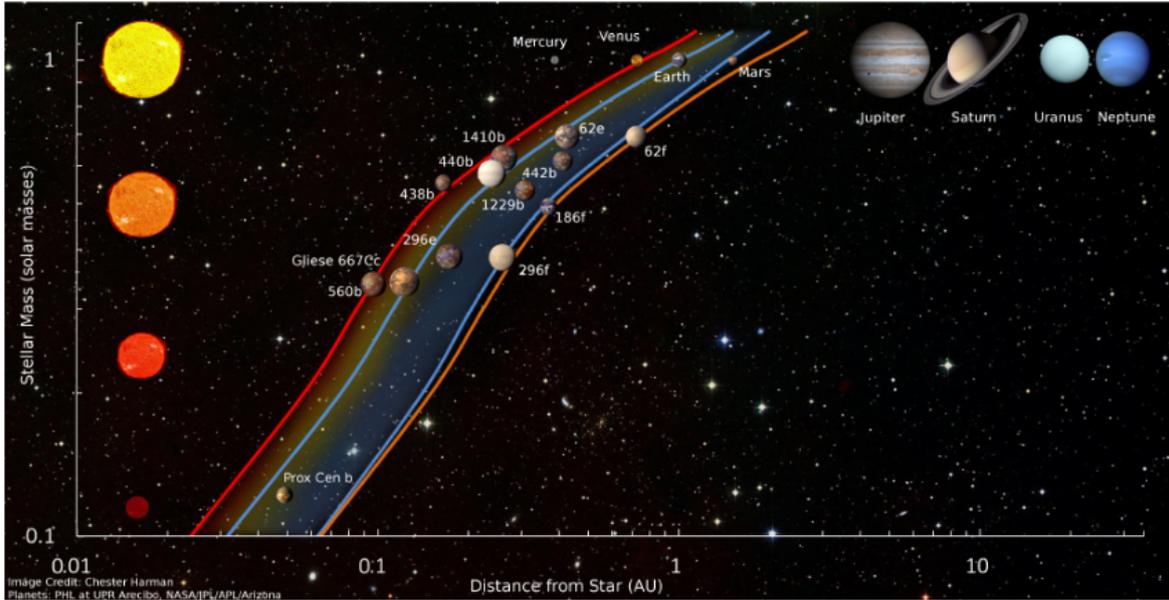


Figura 2.6: [29, Fig. 3] Diagrama esquemático de los límites de las distancias orbitales de la zona habitable para estrellas con diferente masa. Las líneas roja y naranja muestran los límites optimistas interior y exterior de la zona habitable, mientras que las azules muestran los conservativos. Se representan varios planetas como referencia. Créditos: Chester Harman.

En cuanto a la zona habitable en la galaxia, destacan tres factores a tener en cuenta: la radiación, la presencia de supernovas y la existencia de metales. Por las dos primeras razones no se puede dar la vida en el centro de la galaxia, aunque esta región tenga una composición química muy propicia. Por otro lado, la escasez de elementos pesados en la parte más exterior de la galaxia impide la creación de planetas en esta región. Se concluye así que la zona habitable galáctica es el anillo presentado en la Figura 2.7 [13, pp. 66-67].

La radiación y la presencia de supernovas son algunos de los fenómenos altamente energéticos que llevan consigo la destrucción de moléculas orgánicas, dificultando así

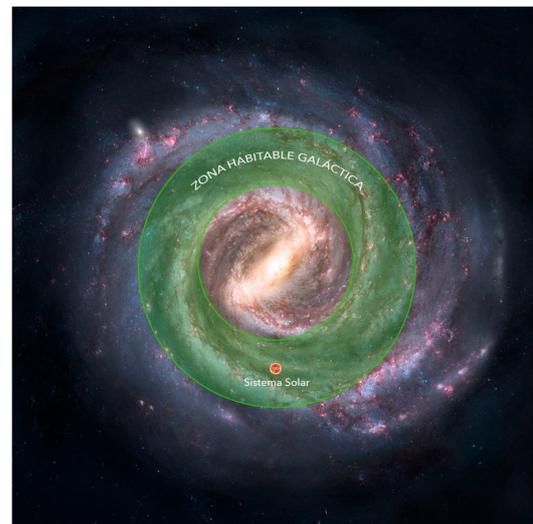


Figura 2.7: [30] Representación de la zona habitable galáctica.

la presencia de vida. Por otra parte, los elementos pesados⁴ son necesarios para la formación de química orgánica.

2.1.5. Tipos de exoplanetas

Actualmente, a través de varios métodos de detección, se pueden conocer propiedades de los exoplanetas como la masa, el tamaño y parámetros orbitales. Sin embargo, aún es difícil conocer otras características. Es por ello que la clasificación más usual de los exoplanetas depende casi en su totalidad de la masa y el tamaño, destacando cuatro tipos [13, pp. 32-33]. Aún así, esta clasificación puede variar dependiendo del autor o fuente que se consulte.

- **Gigantes gaseosos:** se caracterizan por tener una masa aproximadamente como la del **planeta Júpiter o mayor**. Su composición es fundamentalmente de hidrógeno y/o helio, formado por un núcleo sólido y elementos volátiles a su alrededor. Ejemplos de gigantes gaseosos en nuestro Sistema Solar son Saturno y el mismo Júpiter [31] [13, pp. 32-33].

Dentro de este tipo destacan los llamados **Júpiter calientes**, que, como su propio nombre indica, son planetas de estas características que orbitan muy cercanos a su estrella, de modo que pueden alcanzar temperaturas de miles de grados (Celsius o Fahrenheit). Debido a su tamaño y a su pequeña órbita, los Júpiter calientes son fáciles de detectar por los métodos de detección tradicionales (velocidad radial y tránsitos) [31].

Destaca el exoplaneta HIP 67522 b (descubierto en Junio 2020), el cual orbita una estrella aproximadamente igual de masiva que el Sol, formada hace 17 millones de años, lo que lleva a deducir que HIP 67522 b tiene unos pocos millones de años menos, convirtiéndose en el más joven detectado hasta la fecha. Tiene un

⁴En Astrofísica los elementos pesados son aquellos más pesados que el helio. La abundancia de estos elementos se describe con el concepto de *metalicidad*.

periodo de 7 días y un diámetro 10 veces mayor que el terrestre, de forma que se intuye que está compuesto principalmente por elementos volátiles [31].

- **Planetas neptunianos:** son planetas con una masa como mínimo **10 veces mayor que la masa de la Tierra, pero menor que la de Júpiter**; es decir, comparado con los planetas de nuestro Sistema Solar, se asemejan a Urano y Neptuno, como su propio nombre indica. A pesar de que su composición puede variar, todos ellos presentan núcleos rocosos y una atmósfera rica en hidrógeno y helio [32] [13, pp. 32-33].

Cabe destacar el exoplaneta HAT-P-11b, descubierto en 2017. En su atmósfera se pudieron detectar moléculas de vapor de agua mediante la técnica de espectroscopía de transmisión [32].

- **Super-Tierras:** el nombre de estos planetas está relacionado únicamente con su tamaño, comprendido entre **3 y 10 veces la Tierra**, es decir, son mayores que la Tierra y menores que Neptuno. Es bastante común encontrar exoplanetas de este tipo, sin embargo, no hay ningún ejemplo en nuestro Sistema Solar. Aunque la composición de las diferentes Super-Tierras puede variar mucho, muchos suelen ser planetas rocosos rodeados de gas denso [33] [13, pp. 32-33].

Un ejemplo de Super-Tierra es 55 Cancri e, descubierto en 2004, cuyo radio es el doble del de nuestro planeta. Este exoplaneta destaca por describir una órbita con un periodo de 17 horas y 40 minutos, a una distancia de su estrella 26 veces menor que la que separa a Mercurio del Sol. El planeta presenta temperaturas muy diferentes dependiendo de la parte de la superficie; la cara más caliente presenta una temperatura cercana a 4,400 grados Fahrenheit, mientras que la más fría está a 2,060 grados Fahrenheit. La baja densidad de este exoplaneta y las características en las que se encuentra, lleva a pensar que los materiales que lo constituyen están en un estado supercrítico, lo que se ha confirmado también a partir los datos de Spitzer [34].

Existe un grupo de planetas que se encuentran en la frontera entre Super-Tierras y Neptunianos, estos reciben el nombre de **mini-Neptunos**. Suelen ser planetas compuestos principalmente por gas, pero mucho más pequeños que Neptuno.

Se tienen como ejemplos TOI 270 b, TOI 270 c y TOI 270 d, que orbitan una estrella enana tipo M (ver sección 2.1.6 *Conceptos básicos*) [33].

- **Planetas terrestres:** los planetas terrestres o exotierras presentan un **tamaño similar al de la Tierra** (entre la mitad y dos veces su radio), de ahí su nombre. Además, están compuestos principalmente de roca, silicato, agua o carbono. Su superficie suele ser sólida o líquida, y pueden tener atmósfera gaseosa [35] [13, pp. 32-33].

Como se vio en el apartado 2.1.4., la búsqueda de vida en exoplanetas se basa en su gran mayoría en los conocimientos que se tienen sobre la Tierra y las condiciones a las que está sometida. Por esto mismo, resulta muy interesante estudiar planetas terrestres en la zona habitable estelar, es decir, a la distancia correcta de su estrella para que pueda presentar agua en estado líquido [35]. En 2017 se descubrieron siete exotierras orbitando la estrella TRAPPIST-1, una enana roja. Este es un objeto de estudio realmente interesante, debido a que cuatro de ellos (los planetas d, e, f y g) presentan características que hace probable que alberguen vida [35].

Por otro lado, cabe destacar que existe un hueco notable, ***Fulton gap*** (Hueco Fulton), en los descubrimientos de exoplanetas con un radio entre 1.5 y 2 veces el de la Tierra. Como se puede ver en la figura 2.8, la mayoría de los planetas pequeños están comprendidos en el rango delimitado por 2 y 4 veces el radio terrestre; es decir, entre el planeta Wolf 503b y Neptuno [35].

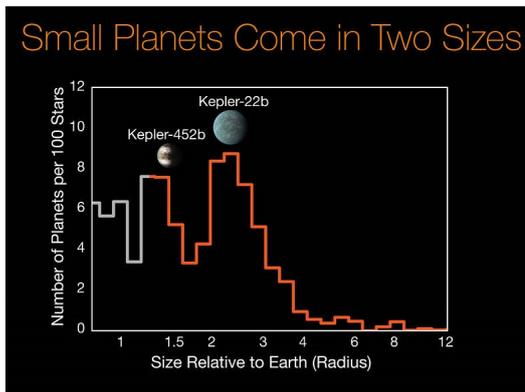


Figura 2.8: Representación gráfica de el hueco de Fulton. Fuente: [36]. Créditos: NASA/Ames Research Center/CalTech/University of Hawaii/B.J. Fulton.

La presencia del *Fulton gap* puede llevar a discernir entre dos planetas dependiendo de su formación. Los planetas de mayor radio pueden atraer hidrógeno y helio formando su atmósfera gaseosa. Por otro lado, los planetas más pequeños, similares a la tierra, no tienen esta capacidad y por ello son principalmente rocosos [37][38].

Por último, además de los cuatro tipos de exoplanetas mencionados, también existen las **Sub-Tierras**. Como su propio nombre indica, se caracteriza por ser **menos masivos que la Tierra**. Por ejemplo, en nuestro Sistema Solar se puede enmarcar dentro de esta clase a Mercurio y Marte [13, p. 33].

2.1.6. Conceptos básicos

Para la comprensión de la física estelar es imprescindible conocer la clasificación de las estrellas según la temperatura de la superficie y la luminosidad absoluta. Esta clasificación se representa en el **diagrama de Hertzsprung-Russell** (ver Figura 2.9), el cual lleva el nombre de los dos astrónomos que lo definieron en 1911 [39, p. 28].

En general, aunque depende de la región estelar que se estudie, entre el 80-90 % de las estrellas se encuentra en la **secuencia principal**⁵, conjunto de estrellas caracterizadas por su masa. Las estrellas en esta etapa cumplen la relación $L \sim T_E^8$, siendo L la luminosidad y T_E la temperatura efectiva de la superficie de la estrella. Al principio, se pensaba que la evolución de las estrellas era tal que comenzaban siendo calientes y se convertían en frías. Es por esto que aún hoy en día se suele denominar a las estrellas de tipo O y B *estrellas de tipo temprano*, y a las de tipo F,G,K y M *estrellas de tipo tardío* [39, p. 29].

Las estrellas que pertenecen a la secuencia principal pueden tener una masa desde

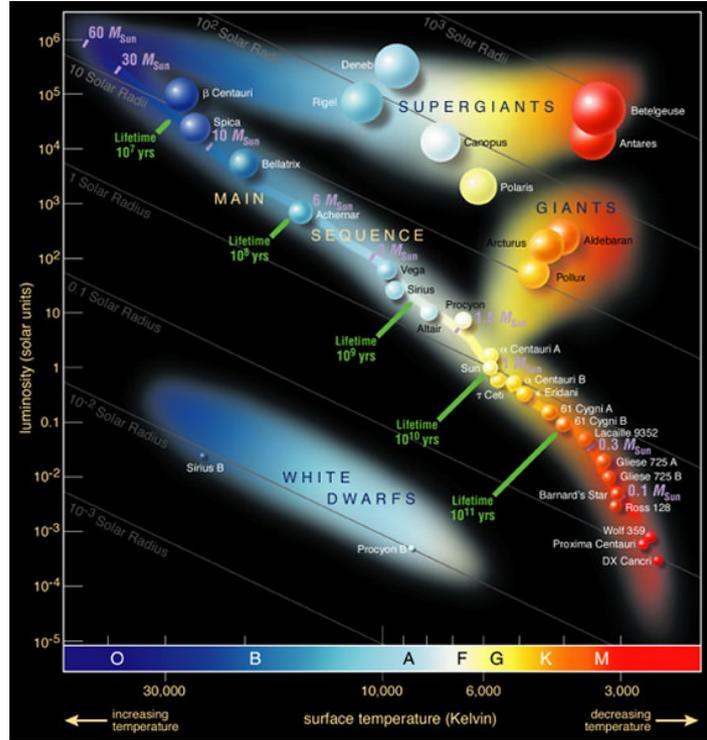


Figura 2.9: [40] Esquema de las estrellas de la secuencia principal.

⁵Cabe destacar que, antes de que las estrellas evolucionen hasta la secuencia principal, las estrellas son lo que se conoce como enanas marrones. Si la masa de la estrella es mayor a $0.1M_{\odot}$, entonces llegará a la secuencia principal, y esta fase comenzará cuando empiece la fusión de hidrógeno en helio. Por otro lado, si la masa es menor que la indicada, entonces las estrellas no llegarán a la secuencia principal, y no tendrá lugar la fusión [41].

$0.1-100M_{\odot}$. La relación de la luminosidad con la masa es la siguiente [39, p. 29]:

$$L \sim M^{\alpha} \text{ con } \left\{ \begin{array}{l} \alpha = 3 \text{ para } M > M_{\odot} \\ \alpha = 5 \text{ en otro caso} \end{array} \right\}. \quad (2.1)$$

Las estrellas pasan la mayor parte de su existencia en la secuencia principal. En este periodo el hidrógeno se convierte en helio por fusión termonuclear en el núcleo de la estrella. Una vez ha finalizado la sintetización de hidrógeno en helio, durante un periodo corto de tiempo aumenta en gran cantidad la luminosidad de la estrella, además, las capas externas se expanden, llevando a sí a la fase de estrellas gigantes [39, p. 29] [41].

Las **gigantes rojas** son estrellas frías, cuyo radio es aproximadamente 100 veces el radio solar, y la masa suele estar comprendida entre 1-2 M_{\odot} . Estas estrellas, para una misma temperatura, presentan una luminosidad mayor que las de la secuencia principal [39, p. 29].

Dependiendo de la masa que presenten las estrellas en la fase ZAMS (*Zero Age Main Sequence*)⁶, su evolución podrá tomar caminos diferentes. Si la masa es menor o igual que $8M_{\odot}$, en la fase de gigante roja tendrá lugar la contracción y el aumento de temperatura en el corazón de la estrella; y, por otro lado, la expansión y el enfriamiento de las capas externas. Más tarde, si la masa estelar es suficiente ($M > M_{\odot}$), entonces tendrá lugar la fusión de helio, expulsando finalmente la envoltura y convirtiéndose en una enana blanca de CNO con masa $0.6M_{\odot} < M < 1.1M_{\odot}$. En otro caso, el corazón se contrae y pasa a ser una enana blanca de helio con una masa $0.2M_{\odot} < M < 0.4M_{\odot}$. Las **enanas blancas** son estrellas más calientes, con radios del orden del radio terrestre y una luminosidad más baja que las de la secuencia principal (para la misma temperatura) [41].

Por otro lado, si la masa de la estrella es mayor que $8M_{\odot}$, entonces presenta una envoltura mucho mayor y pérdidas de energía por generación de neutrinos. Evolucionan

⁶La fase ZAMS (*Zero Age Main Sequence*) es el momento en el que la estrella tiene la temperatura suficiente para iniciar la fusión de hidrógeno, encontrándose así en la secuencia principal.

hasta el grupo de estrellas más luminosas, las **supergigantes** [39, p. 29]. Estas estrellas, al agotar las reacciones de fusión nuclear, se convertirán en estrellas de neutrones (si $10M_{\odot} < M < 30 - 40M_{\odot}$) o agujeros negros (si $30-40M_{\odot} < M$) [41].

La nomenclatura de las estrellas depende de su espectro y de su luminosidad. Por una parte, como se puede ver en la Figura 2.9, según su temperatura se les asignan las letras O, B, A, F, G, K, y M; siendo O las más calientes y M las más frías. Los números romanos se emplean para clasificarlas según la luminosidad, empleando el intervalo I-V. Las estrellas de la secuencia principal se corresponden con el número romano V [39, p. 29].

Una vez vista la clasificación de las estrellas y su evolución, antes de comenzar a explicar los métodos de detección, es conveniente describir los elementos y parámetros básicos que describen un sistema planetario.

Se diferencian dos planos: el **plano orbital del planeta** (el que contiene la órbita del planeta) y el **plano de referencia** (plano del cielo). En estos dos planos se pueden identificar los siguientes elementos (ver Figura 2.10):

- **Nodo ascendente:** punto en el que el objeto corta el plano del cielo, moviéndose desde el hemisferio sur al norte [42].
- **Nodo descendente:** punto en el que el objeto corta el plano del cielo, moviéndose desde el hemisferio norte al sur [42].
- **Línea de nodos:** recta que contiene al nodo ascendente y al descendente [42].
- **Periastro:** punto de la órbita del planeta en el que la distancia con la estrella es mínima.
- **Apoastro:** punto de la órbita del planeta en el que la distancia con la estrella es máxima.

Además, se tienen los siguientes parámetros, muy importantes en términos astrofísicos:

- i : inclinación del plano orbital con respecto al de referencia.
- ω : ángulo formado entre la línea de nodos y el periastro.
- a : semieje mayor entre la órbita del planeta y la de la estrella.
- ν : posición del planeta respecto al periastro.

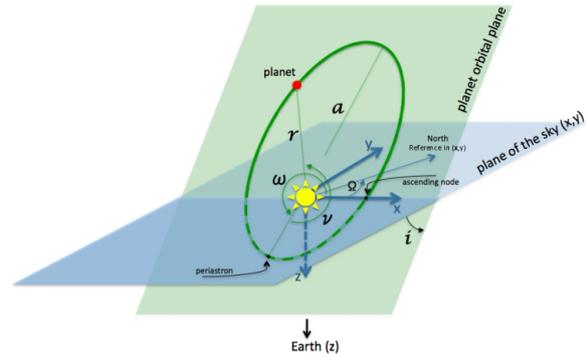
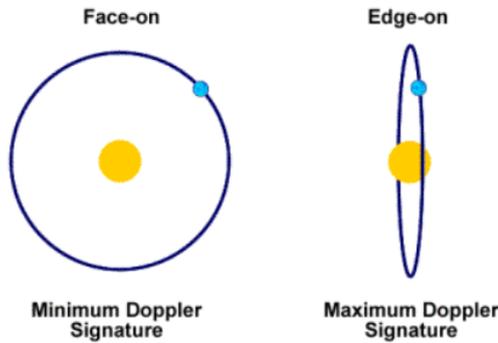


Figura 2.10: Esquema del plano orbital de un planeta que orbita una estrella con respecto al plano del cielo.

- Ω : ángulo entre la dirección Norte y el nodo ascendente.
- ϵ : excentricidad de la órbita. Dependiendo de su valor se tiene una órbita con forma más circular ($\epsilon = 0$) o más elíptica ($\epsilon > 0$).



Se diferencian dos configuraciones principales de los sistemas planetarios. Una de ellas se denomina **edge-on**, y hace referencia a la inclinación entre el plano de la órbita con respecto al del cielo de un ángulo de 90° . La otra forma se llama **face-on**, y ambos planos

Figura 2.11: [43] Esquema del plano orbital con respecto al plano de referencia: *face-on* vs *edge-on*.

forman un ángulo de 0° .

2.1.7. Método de detección por tránsitos

Los tránsitos se producen cuando un planeta pasa entre la estrella y el observador. En nuestro Sistema Solar, este fenómeno se puede observar desde nuestro planeta al transitar tanto Venus como Mercurio [44].

El paso de un exoplaneta entre la estrella y la Tierra puede observarse a partir de los cambios en la curva de luz, que es una gráfica que representa la intensidad de luz estelar en función del tiempo. Durante un tránsito se produce una disminución de

esta luz, como se puede observar en la Figura 2.12. Este primer descenso del brillo se conoce como **tránsito principal** o **tránsito** [45].

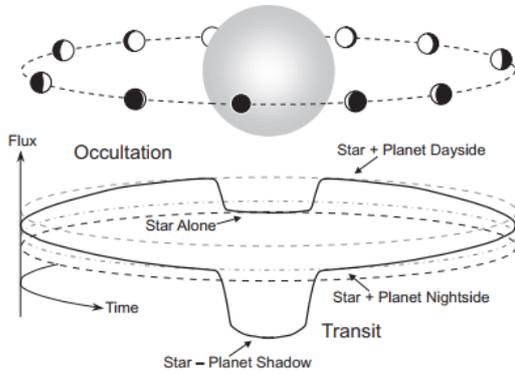


Figura 2.12: [46, p. 56] Detección de exoplanetas con Tránsito.

Estos efectos sobre la curva de luz pueden ser causados por más fenómenos, como estrellas binarias o por irregularidades en la emisión de radiación de una estrella. Sin embargo, cuando el tránsito se produce de forma regular y periódica temporalmente, entonces es muy probable que lo esté causando un planeta [48, p. 5][11, p. 56]. Para que este suceso pueda ser visto por el observador, tiene que darse la configuración *edge-on* del sistema, siendo más evidente cuanto más cercano sea el ángulo de inclinación i a 90° [47].

Para presentar algunas ecuaciones que definen los tránsitos de un planeta, se denotará el radio del planeta como R_p y el de la estrella, R . De la misma forma, la masa del planeta será M_p y la masa estelar, M .

El sistema de referencia empleado es centrado en la estrella, con el plano del cielo correspondiente al X-Y, y el sentido positivo del eje Z apuntando al observador. Además, se alinea el eje X con la línea de nodos, siendo $\Omega=180^\circ$. Por otra parte, se supondrá que todos los tránsitos ocurren centrados en $X=0$ [46, pp. 55-56].

Un dato que va a ser necesario en muchas de las ecuaciones es la distancia del planeta respecto a la estrella [46, p. 56]:

$$r = \frac{a(1 - \epsilon^2)}{1 + \epsilon \cos \nu}, \quad (2.2)$$

donde a es el semieje mayor entre la órbita del planeta y la de la estrella, como se definió en el apartado 2.1.6.

Por otra parte, cuando el planeta pasa por detrás de la estrella, se da la **ocultación** o **tránsito secundario**, en el que vuelve a observarse un decaimiento en la curva de luz, pero menor que la anterior, ya que en este caso la luz estelar llega en su totalidad al observador, y solo se omite el reflejo causado por el planeta[45] [47].

La distancia entre el tránsito y el origen del sistema, proyectada sobre el plano del cielo, se define como b (parámetro de impacto), tanto para el tránsito principal como secundario [46, p. 56]:

$$b_{tra} = \frac{a \cos(i)}{R} \left(\frac{1 - \epsilon^2}{1 + \epsilon \sin(\omega)} \right), \quad (2.3)$$

$$b_{occ} = \frac{a \cos(i)}{R} \left(\frac{1 - \epsilon^2}{1 - \epsilon \sin(\omega)} \right). \quad (2.4)$$

A partir del parámetro de impacto, b , se puede definir la situación del tránsito en el plano X-Y. En los casos en los que $R \ll a$, el tránsito seguirá una trayectoria recta de $X = -R\sqrt{(1 - b^2)}$ a $X = R\sqrt{(1 - b^2)}$, en $Y = bR$ [46, p. 56].

Otros datos importantes en este estudio son los tiempos de duración del tránsito. Destacan cuatro tiempos: t_I , t_{II} , t_{III} y t_{IV} . Todos ellos corresponden a un punto de contacto entre la estrella y el planeta, como puede verse en la Figura 2.13 [46, pp. 57-58].

Se puede calcular el tiempo correspondiente a la parte del tránsito más profunda a partir de la siguiente ecuación: [46, p. 57]

$$t_{III} - t_{II} = \frac{T}{2\pi\sqrt{1 - \epsilon^2}} \int_{\nu_{II}}^{\nu_{III}} \left[\frac{r(\nu)}{a} \right]^2 d\nu, \quad (2.5)$$

donde T , es el periodo orbital. Para obtener el valor de ν en los tiempos requeridos, se iguala

$$r = \frac{a(1 - \epsilon^2)}{1 + \epsilon \cos \nu} \sqrt{(1 - \sin^2(\omega + \nu) \sin^2(i))} \quad (2.6)$$

a $R \pm R_p$ [46, p. 57].

En el caso de órbitas circulares, las expresiones de los tiempos se simplifican como sigue:

$$t_{III} - t_{II} = \frac{T}{\pi} \text{sen}^{-1} \left[\frac{R \sqrt{(1 - k)^2 - b^2}}{a \text{sen}(i)} \right], \quad (2.7)$$

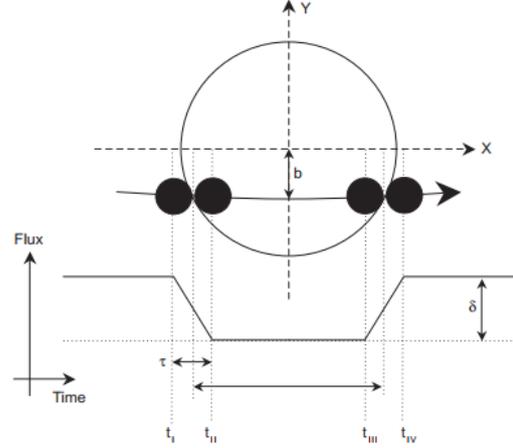


Figura 2.13: [46, p. 56] Esquema de un tránsito.

$$t_{IV} - t_I = \frac{T}{\pi} \text{sen}^{-1} \left[\frac{R}{a} \frac{\sqrt{(1+k)^2 - b^2}}{\text{sen}(i)} \right], \quad (2.8)$$

denotando $k \equiv \frac{R_p}{R}$ [46, p. 58].

Por último, se describe la pérdida de luz durante un tránsito, a partir de la intensidad media del planeta (I_p) y de la estrella (I), suponiendo que ambas son constantes en el tiempo de duración de un tránsito:

$$\delta_{tra} \approx k^2 \left(1 - \frac{I_p(t_{tra})}{I} \right), \quad (2.9)$$

que puede ser aproximado por $\delta_{tra} \approx k^2$ en la mayoría de los casos, debido a que la intensidad del planeta cuando no refleja la radiación de la estrella es despreciable [46, p. 58].

Por otra parte, en el caso de la ocultación: [46, p. 58]

$$\delta_{occ} \approx k^2 \left(\frac{I_p(t_{occ})}{I} \right). \quad (2.10)$$

Una de las ventajas de usar este método es la gran variedad de características del sistema que se pueden llegar a conocer. En primer lugar, con una simple observación de la curva de luz con los tránsitos, ya se puede tener una idea de la relación entre el radio del planeta y el de la estrella y del tamaño de la órbita. Cuanto más grande sea el exoplaneta en comparación con su estrella madre, más profundo será el tránsito producido, puesto que más luz ocultará al pasar por delante de esta. Por otra parte, dependiendo de la distancia de la órbita a la estrella, más o menos tiempo va a tardar el planeta en orbitarla [45]. Más aún, a partir del espectro, por ejemplo, se puede deducir el tamaño de la estrella, por lo que se tendría también el del planeta, usando la relación conocida k [47].

La masa del planeta no es una cantidad que pueda conocerse a partir de esta técnica si únicamente hay un exoplaneta orbitando⁷. Sin embargo, sí que puede calcularse con

⁷En el caso de haber más de un exoplaneta sí que podría conocerse la masa (Ver apartado 2.1.8., *Variación en el tiempo de tránsito*).

el método de velocidad radial que se tratará en la siguiente sección. Así, empleando los tránsitos y la velocidad radial se puede llegar a obtener la densidad del exoplaneta, lo que permite analizar su composición [47].

Otra propiedad posible de estudiar con el método de los tránsitos es la atmósfera del planeta. Cuando se produce el tránsito, el espectro que se obtiene presenta alteraciones, que pueden compararse con el espectro estelar fuera del tránsito. La radiación de la estrella cuando el planeta la está transitando, atraviesa la atmósfera del planeta antes de llegar a la Tierra, la cual absorbe en diferentes frecuencias según su composición. [47]. Este método, la espectroscopía de transmisión, nos permite conocer las especies presentes en la atmósfera planetaria a partir de las absorciones que se observan en el espectro estelar cuando el planeta transita.

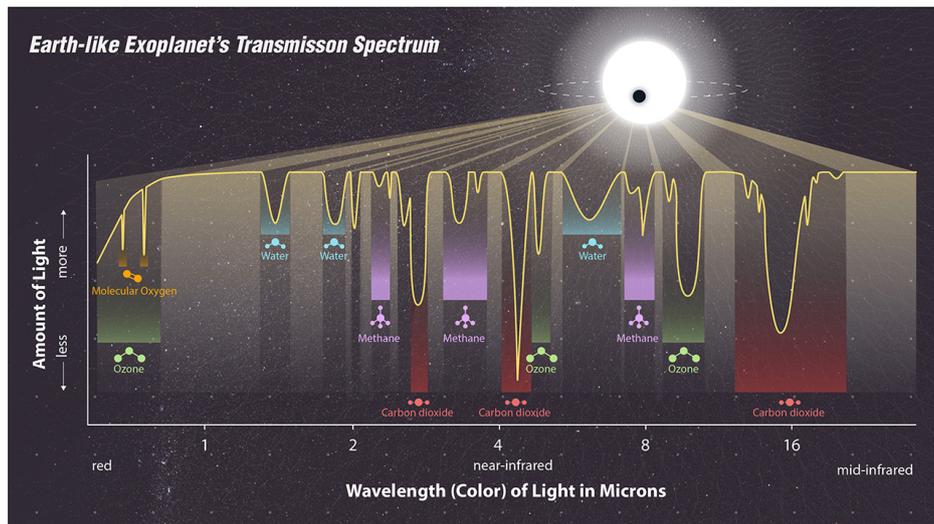


Figura 2.14: Esquema de espectroscopía de transmisión para un planeta como la Tierra. Fuente: [49]. Créditos: NASA, ESA, CSA, STScI, Joseph Olmsted (STScI).

Con el método de detección por tránsitos, a diferencia de otros, se pueden identificar varios exoplanetas en un sistema. Cada uno de ellos describirá una órbita con un periodo diferente y tendrá un tamaño concreto, por lo que los tránsitos en la curva de luz serán diferenciables [45].

Por otro lado, aún siendo tan exitoso, este método tiene ciertas desventajas a tener en cuenta. En primer lugar, la disposición del sistema en *edge-on* hace que no todos los exoplanetas puedan determinarse a través de su tránsito. Por otra parte, para poder confirmar la existencia del exoplaneta, se han de ver varios tránsitos regulares

y periódicos. Los planetas suelen tener periodos entre meses y años, pero el tránsito dura una pequeña fracción de este tiempo (desde horas hasta días), es por esto que los tiempos de observación deben de ser extensos para asegurar el hallazgo. En último lugar, se sabe que puede haber disminuciones del brillo estelar causado por otros factores, por lo que es fácil que esta técnica lleve a algunos falsos positivos [47].

2.1.8. Otros métodos de detección

Velocidad radial

La técnica de velocidad radial para la búsqueda de exoplanetas se basa en el movimiento reflejo que experimenta una estrella cuando un planeta orbita a su alrededor. En general, como dedujo Isaac Newton en su momento, dos cuerpos masivos cualesquiera experimentan un campo gravitacional debido a su interacción. Como resultado, los dos orbitan alrededor del centro de masas del sistema que forman, estando siempre en posiciones opuestas uno respecto al otro y describiendo ambos el mismo tipo de órbita (ver Figura 2.15) [11, p.15].

El centro de masas (CM) de los dos cuerpos se sitúa en la línea imaginaria que une sus centros [11, p.15]. Su posición en esta línea se puede calcular a partir de la siguiente ecuación, que ha sido particularizada para una dimensión:

$$x_{CM} = \frac{m_1 \cdot x_1 + m_2 \cdot x_2}{m_1 + m_2}, \quad (2.11)$$

siendo m_1 y m_2 las masas de ambos objetos y x_1 y x_2 sus posiciones respectivamente [50]. Desarrollando la ecuación anterior, se llega a la siguiente relación entre las masas y las distancias:

$$\frac{m_1}{m_2} = \frac{x_2 - x_{CM}}{x_{CM} - x_1}. \quad (2.12)$$

Si nos centramos en el caso en el que m_1 es la masa de una estrella y m_2 la del planeta, se deduce que el centro de masas estará mucho más cerca de la estrella que del planeta, siendo mayor la diferencia entre las distancias cuanto mayor lo sea entre las masas. Así, el cociente entre las masas de los cuerpos es igual al inverso del cociente de sus distancias al centro de masas [11, p. 15].

La variable que se mide para detectar este movimiento orbital de la estrella es la **velocidad radial**, es decir, la componente de la velocidad en la dirección de la línea visual del observador [51]. Esta velocidad se puede medir estudiando el espectro de la estrella. La frecuencia y longitud de onda con la que se observa la radiación de dicha estrella se verá modificada al alejarse o acercarse al observador, según el efecto Doppler.

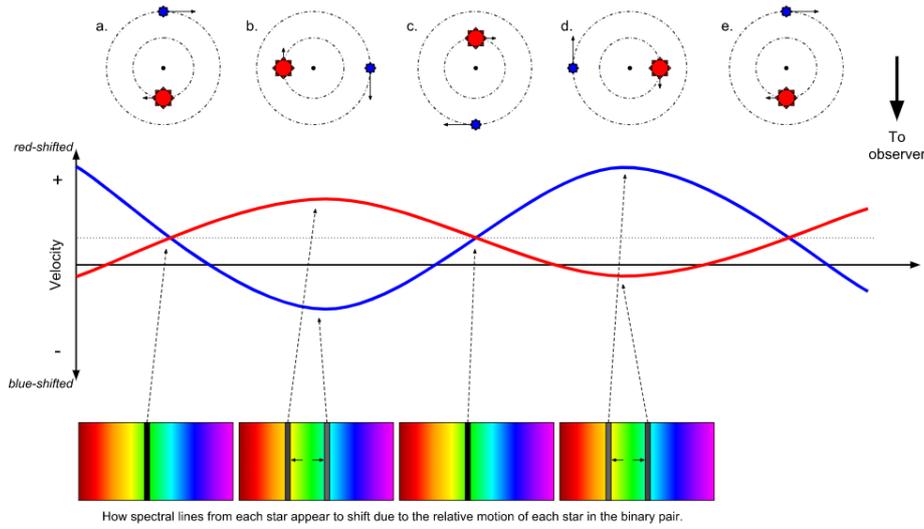


Figura 2.15: [52] Esquema de la variación en el espectro de una estrella debido al efecto Doppler, causado por la presencia de un planeta orbitando la estrella.

La dependencia temporal del movimiento de la estrella en la dirección del observador (eje z) tiene la siguiente expresión:

$$z(t) = r(t)\text{sen}(\omega + \nu(t))\text{sen}(i), \quad (2.13)$$

donde r es la posición radial de la estrella y ν el ángulo de la fase orbital. De esta forma, aplicando la segunda ley de Kepler, la ecuación de la elipse $\left(r = \frac{a(1-\epsilon^2)}{(1+\epsilon\cos\nu)}\right)$ y derivando la expresión anterior, la velocidad radial es [53, p. 17]

$$v_r(t) = \frac{dz(t)}{dt} = \frac{2\pi}{T} \frac{a_S \text{sen}(i)}{\sqrt{1-\epsilon^2}} (\epsilon\cos(\omega) + \cos(\omega + \nu(t))). \quad (2.14)$$

A esta definición se le añade un término aditivo que representa la velocidad radial

constante del centro de masas del sistema, v_0 , de tal forma que se puede escribir:

$$v_r(t) = v_0 + K (\epsilon \cos(\omega) + \cos(\omega + \nu(t))); \text{ con } K = \frac{2\pi a_S \sin(i)}{T \sqrt{1 - \epsilon^2}}, \quad (2.15)$$

siendo K la semi-amplitud de la señal del movimiento de la estrella [53, p. 17]. La forma de dicha señal viene dada por los términos entre paréntesis de la ecuación anterior, de manera que varía principalmente con ϵ y ω , como se puede ver en la siguiente figura [54, p. 6].

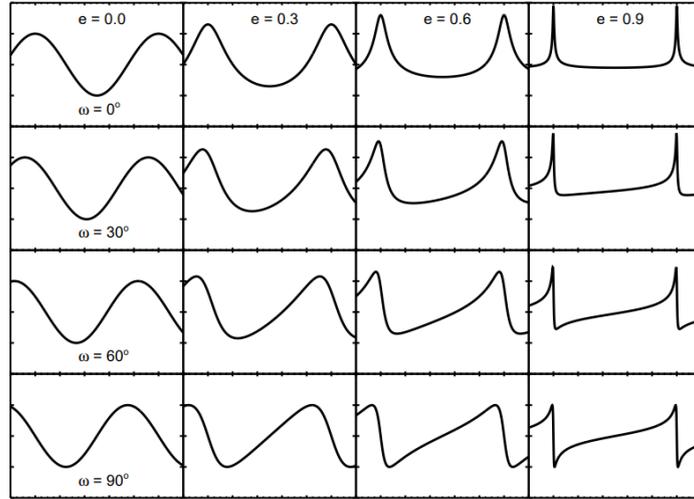


Figura 2.16: [54] Variación de la señal dependiendo de ϵ y ω .

Además, dependiendo de las diferentes amplitudes que presente la gráfica de la velocidad radial, se puede inferir el número de exoplanetas del sistema. Otro dato que se tiene a partir de la velocidad radial es la masa del exoplaneta. Basándose en la definición de K , se puede obtener la siguiente dependencia de a_S con el periodo, la excentricidad y el propio K [53, p. 18]:

$$a_S \sin(i) = \frac{T}{2\pi} \sqrt{1 - \epsilon} K. \quad (2.16)$$

De esta forma, teniendo en cuenta también las siguientes tres relaciones, se obtiene

una expresión para las masas del planeta y de la estrella: [53, p. 18]

$$\left\{ \begin{array}{l} \frac{a^3}{T^2} = \frac{G(M+M_P)}{4\pi^2} \\ a = a_S + a_P \\ M a_S = M_p a_P \end{array} \right\} \Rightarrow \frac{(M_P \text{sen}(i))^3}{(M_P + M)^2} = \frac{T}{2\pi G} K^3 (1 - \epsilon^2)^{\frac{3}{2}}. \quad (2.17)$$

Típicamente, si $M_p \ll M$, se llega a la siguiente relación: [53, p. 18]

$$M_P \text{sen}(i) \approx \left(\frac{T}{2\pi G} \right)^{1/3} K M^{2/3} \sqrt{1 - \epsilon^2}. \quad (2.18)$$

El hecho de que la inclinación no se pueda medir con la técnica de velocidad radial tiene como consecuencia que en numerosas ocasiones no se conozca la masa exacta del planeta. En general, se dice que la masa calculada de esta forma es la mínima del planeta, puesto que únicamente cuando $i = 90^\circ$ la expresión de la ecuación 2.18 se corresponderá con la masa exacta de este [54, pp. 6-8].

Dependiendo de la inclinación del plano orbital con respecto al de referencia, los cambios en la velocidad radial pueden no ser apreciables, siendo completamente imperceptibles cuando los planos son paralelos entre sí, de tal forma que la velocidad radial de la estrella en la dirección del observador es nula. Esto hace que, considerando todos los ángulos de inclinación aleatorios de las órbitas, solo se visualicen aproximadamente un 64 % de los efectos en la velocidad radial que podrían verse si ambos planos estuvieran superpuestos [11, pp. 28-29].

El éxito de este método puede variar notablemente dependiendo también de la masa del exoplaneta. Cuanta más masa tenga, mayor va a ser su efecto sobre el movimiento de la estrella. Siendo la masa el único parámetro variable, si un planeta es 10 veces más masivo que otro, provocará efectos sobre la velocidad radial 10 veces mayores. Por otra parte, cuanto menor sea la distancia entre el planeta y la estrella, mayores serán las variaciones en la velocidad radial. Si ahora se compara un planeta 10 veces más cercano a la estrella que otro, siendo el resto de características del sistema idénticas, entonces los efectos serán tres veces mayores [11, p. 29].

Otro obstáculo que puede desfavorecer el uso de este método para buscar exoplanetas es la presencia de moléculas en la atmósfera terrestre que introducen variaciones

en el espectro. Además, la emisión de radiación de las estrellas es muy inconstante, variando la cantidad en diferentes partes de su superficie y en tiempos aleatorios [11, p. 30].

Astrometría

“La astrometría es la especialización de la Astronomía que estudia la posición de los astros en el cielo, con el fin de establecer las coordenadas celestes y sus variaciones en el tiempo y reconstruir los movimientos de las estrellas” [55].

Esta técnica se basa, al igual que la de velocidad radial, en el movimiento reflejo de una estrella causado por la presencia de un exoplaneta orbitando a su alrededor. La idea de este método es visualizar los cambios en la posición aparente de la estrella para, posteriormente, investigar la existencia de un posible exoplaneta [56].

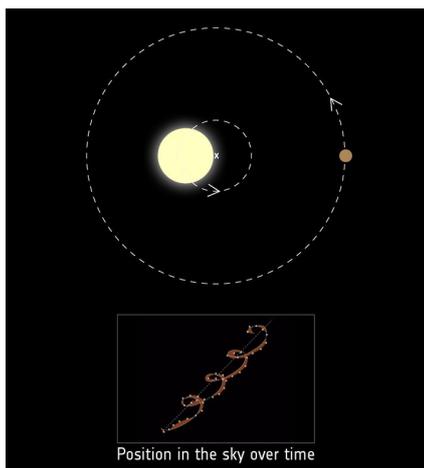


Figura 2.17: [57] Detección de exoplanetas con Astrometría.

Con el fin de obtener estas representaciones, se toman las posiciones de la estrella que se quiere estudiar y de algunas otras a su alrededor [56]. Gracias a misiones como Gaia⁸, este método proporcionará abundantes detecciones en el futuro.

⁸La misión Gaia tiene como objetivo principal obtener un mapa tridimensional de la Vía Láctea. Comenzó el 19 de diciembre de 2013, con un tiempo de duración estimado de cinco años. Sin embargo, su buen rendimiento y las reservas de combustible, han llevado a alargar la misión hasta 2025 [58].

Micro lente gravitacional

Dentro de la teoría general de la relatividad de Einstein, este descubrió las lentes gravitacionales. La definición o el entendimiento de la gravedad como una propiedad geométrica del espacio-tiempo lleva a la idea de que cualquier objeto masivo puede influir en la trayectoria de la luz [59].

De esta forma, si se tiene una fuente de luz distante y entre el observador y dicha fuente se cruza un cuerpo de gran masa (ya sea otra estrella, un planeta, una galaxia...), entonces la trayectoria de la radiación que llega al observador se verá alterada por el efecto de la gravedad sobre ella. El objeto causante de esta flexión recibe el nombre de *lente* [60].

La consecuencia de este suceso es un enfoque de la luz que provoca una distorsión de las imágenes. Dependiendo del tipo de lente, se produce un número diferente de imágenes. Si la lente es puntual y única, entonces se tendrán dos imágenes; por el contrario, si se tiene una lente binaria, se producirán entre tres y cinco. Tanto el número como la forma y magnitud de las imágenes puede variar según la posición entre la lente y la fuente [60].

A continuación nos centraremos en el caso particular en el cual tanto la lente como la fuente son una estrella cada una, es decir, en la subcategoría de *microlente gravitacional* [60].

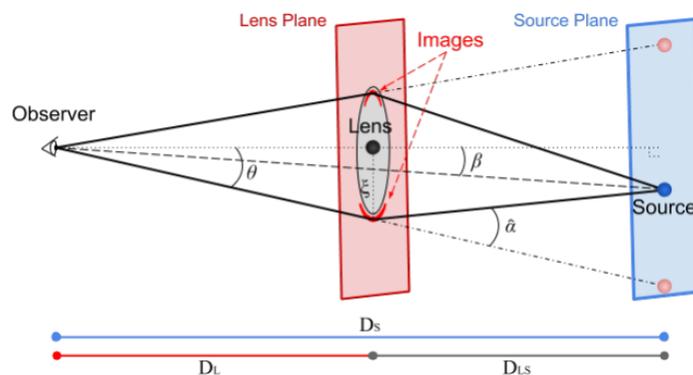


Figura 2.18: [60, p. 5] Esquema de la óptica básica en microlentes gravitacionales.

El ángulo que es desviado el haz de luz, *ángulo de Einstein*, puede calcularse con la siguiente expresión:

$$\hat{\alpha} = \frac{4GM_L}{c^2\xi}, \quad (2.19)$$

donde M_L es la masa de la lente, ξ la distancia entre el rayo y la lente, c la velocidad de la luz en el vacío y G la constante gravitacional [60].

Para definir la posición de la fuente con respecto al eje óptico lente-observador, β , a partir del ángulo, θ , formado por el mismo eje y la imagen que ve el observador, se plantea la disposición de la Figura 2.18. Se supone que la fuente está situada a una distancia D_{LS} de la lente, y esta a su vez a una distancia D_L del observador. La separación entre el observador y la fuente será D_S . Así, se tiene la *ecuación de la lente*: [60]

$$\beta = \theta - \hat{\alpha} \frac{D_{LS}}{D_S}. \quad (2.20)$$

Sustituyendo $\hat{\alpha}$ y $\xi = \theta D_L$ en la ecuación anterior [60]:

$$\beta = \theta - \frac{D_{LS}}{D_S D_L} \frac{4GM_L}{c^2 \theta}. \quad (2.21)$$

En la figura 2.19 se puede observar, en primer lugar, la imagen aparente que visualiza el observador debido a la flexión en la trayectoria del haz causada por una estrella. Cuando la lente se acerca al eje entre el observador y la fuente de radiación se tienen dos imágenes, con muy poca distancia angular entre ellas. Por último, debido a la simetría rotacional del sistema, cuando $\beta = 0$ se crea el *anillo de Einstein* [61].

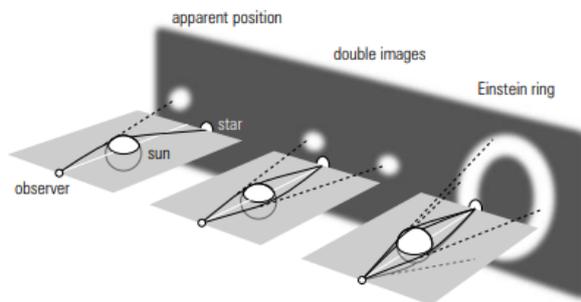


Figura 2.19: [61] Imágenes observadas debido a la flexión en la trayectoria del haz causada por una estrella.

El ángulo de definición del radio del anillo de Einstein a partir de la ecuación 2.21, con $\beta = 0$, es [60]

$$\theta_E = \sqrt{\frac{D_{LS}}{D_S D_L} \frac{4GM_L}{c^2}}. \quad (2.22)$$

Una vez se ha comprendido la consecuencia de las microlentes gravitacionales, es el momento de profundizar más en el caso que interesa en este trabajo, es decir, el estudio de este fenómeno cuando la estrella que actúa como lente tiene un exoplaneta

orbitando a su alrededor. La base del fenómeno es exactamente la misma. Sin embargo, como cabría esperar, el paso de la masa del exoplaneta también tendrá consecuencias en la trayectoria de la radiación.

El enfoque causado por la lente puntual se traduce en un aumento de brillo con su posterior disminución, que dura aproximadamente un mes o varias semanas. Como se puede observar en la Figura 2.20, al cruzarse entre la fuente y el observador un exoplaneta que orbita la estrella lente, se produce un incremento añadido durante algunas horas o días en el brillo medido. Esto es debido a que la presencia de otro objeto masivo provoca una mayor alteración en los haces [57].

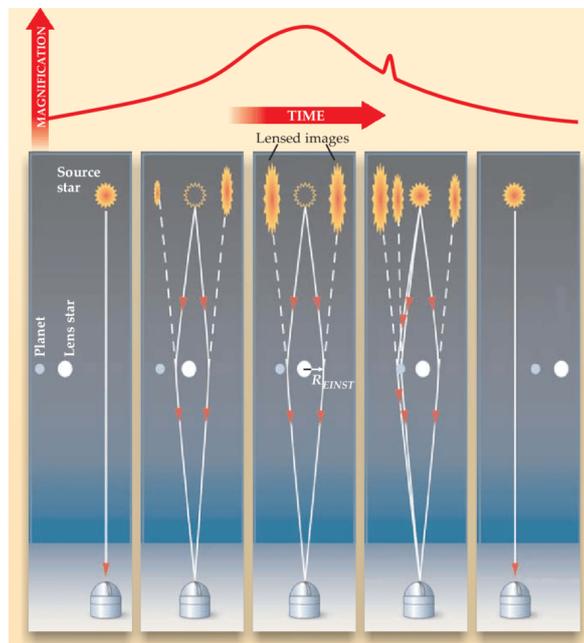


Figura 2.20: Variación del brillo de una estrella al pasar entre la estrella y el observador otra estrella (la lente) y un exoplaneta que la orbita. Fuente: [62]. Adaptada de la web de David Bennett.

Imagen directa

Como el propio nombre indica, la detección de exoplanetas mediante imágenes directas consiste en conocer su existencia a partir de fotografías. El principal problema para llevar a cabo este método es la radiación de la estrella alrededor de la que orbita el planeta, que domina completamente [63].

Para poder atenuar la luz que llega de la estrella, y lograr diferenciar así la radiación proveniente del exoplaneta, destacan dos métodos. El primero de ellos se denomina

coronografía y se basa en un añadido en la construcción de telescopios terrestres para impedir la llegada de la luz de la estrella a su detector [63].

Por otra parte, se está estudiando la posibilidad de emplear también las denominadas *sombras estelares*. Fundamentalmente, cuando se trata de telescopios terrestres, consiste en naves espaciales colocadas de tal forma que obstruyen la llegada de la luz de la estrella a dicho telescopio [63].

Un tercer método para poder distinguir un exoplaneta de forma directa se basa en estudiar la radiación infrarroja. La razón entre el brillo en longitud de onda infrarroja de una estrella y un planeta puede ser de 1 millón (hasta 100.000 en los mejores casos), que, comparado con mil millones en el caso de la luz visible, es mucho más posible de diferenciar [64][11, p. 65]. Aunque la atmósfera terrestre absorbe longitudes de onda de radiación infrarroja, algunas frecuencias llegan hasta telescopios terrestres como el Mauna Kea Observatory en Hawaii, situado a 4205 metros [11, p. 65].

Los planetas que mejor se observan con esta técnica son aquellos que son jóvenes, puesto que emiten más radiación infrarroja al estar más calientes; además, serán más visibles cuanto más masivos. Asimismo, cuanto mayor sea la distancia entre estos y la estrella alrededor de la que orbitan, mejor se diferenciarán ambas señales de luz, permitiendo concluir más fácilmente la existencia del exoplaneta.[11, p. 66] En cuanto a la posición de la órbita con respecto a la dirección visual del observador, es ventajoso que sea *face-on*, para que el planeta no cruce el eje observador-estrella [64].

Aún siendo un método difícil de llevar a cabo, una virtud que posee es el mínimo porcentaje de falsos positivos que se pueden obtener empleándolo. Además, permite obtener información sobre su composición a partir del espectro que refleja su atmósfera [64]. Aunque no es una técnica que haya destacado en la historia de los descubrimientos de exoplanetas, el desarrollo tecnológico de los últimos años facilita en gran medida la toma de imágenes directas [63].

Un ejemplo de planetas descubiertos con este método se puede observar en la Figura 2.21. Se observan los cuatro exoplanetas orbitando alrededor de la estrella HR 8799, la cual se señala con un símbolo rodeada de un círculo negro. Esta sombra es creada para bloquear la luz de la estrella y facilitar la observación de los planetas, como se ha explicado anteriormente [65].

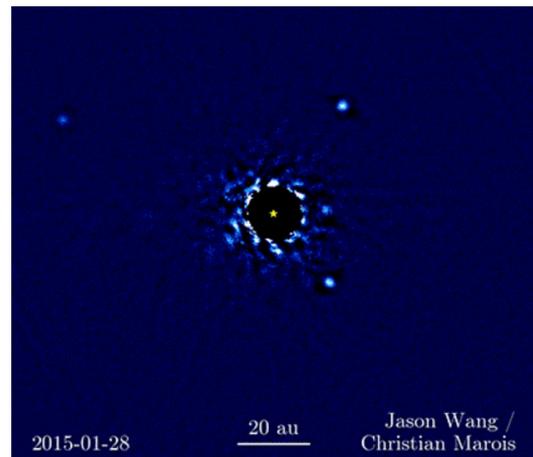


Figura 2.21: Sistema planetario entorno a la estrella HR 8799, detectados con el método de imagen directa. Fuente: [65]. Créditos: Jason Wang y Christian Marois.

Variaciones en el tiempo

Muchos de los eventos que se dan en sistemas estelares y planetarios ocurren de forma regular y periódica, como los tránsitos de los planetas o los tiempos de pulsación de una estrella. Cuando se detectan cambios en las señales que describen estos sucesos, en general es otro cuerpo (no observado directamente) el que los está causando, como un exoplaneta. A continuación se ven los cuatro de los tipos de variaciones temporales más comunes [25].

- Variaciones en el **tiempo de tránsito**: como se vio en la sección anterior, cuando existe un exoplaneta que transita entre la estrella y la dirección del observador, el tránsito ocurre de forma periódica según este orbita. Cuando esta disminución de la radiación que llega al telescopio se da antes o después de lo previsto, significa que el sistema no está formado por un único exoplaneta, sino que hay otro u otros provocando efectos gravitacionales sobre el que se observa [11, pp. 112-113] [25].

La resonancia en los planetas favorece su detección. Cuando dos planetas poseen una relación entre sus periodos orbitales de números enteros (entorno a una estrella en común), se da el fenómeno de resonancia [25]. Cuando dos planetas

se alinean entonces los efectos gravitatorios alcanzan su máximo, produciendo los mayores efectos sobre las variaciones temporales de tránsito [25].

Una ventaja que presenta este método es la posibilidad de calcular la masa de los exoplanetas. Si el sistema estuviera formado por un solo planeta, para el cálculo de la masa se tendría que recurrir a la técnica de velocidad radial, como ya se vio anteriormente. Sin embargo, cuando es un sistema multiplanetario y se producen variaciones temporales en el tránsito, se puede calcular la masa sin falta de más métodos [11, p. 113][25].

- Variaciones en el **tiempo de eclipse**: la idea de este método es la misma que la del anterior. Sin embargo, en este caso el sistema está constituido por dos estrellas binarias eclipsantes, es decir, dos estrellas que se orbitan entre sí en un plano orientado hacia el observador. De esta forma, al igual que ocurre cuando un planeta transita entre una estrella y la Tierra, se producen eclipses cuando una pasa delante de la otra [25] [66].

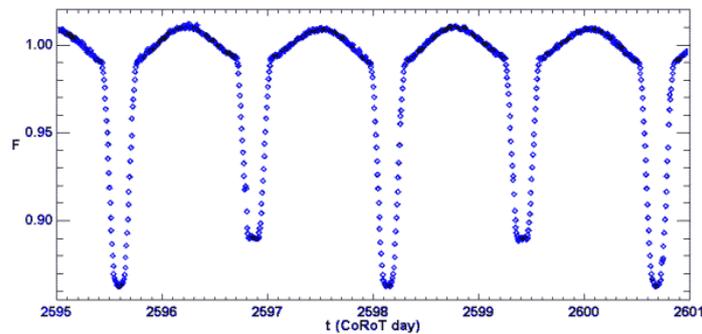


Figura 2.22: [67] Curva de luz de una binaria eclipsante observada por el satélite CoRoT. Crédito: CNES.

Estos eclipses deberían de ser periódicos, como se vio con los tránsitos de exoplanetas. Cuando se presenta una variación en los tiempos de las binarias eclipsantes, puede ser que un planeta las esté orbitando, produciendo estas alteraciones [25].

Cabe destacar que la masa de las estrellas es mucho mayor que la de los planetas que influyen en las variaciones que se estudian. Por esto, los planetas muy masivos son los que presentan consecuencias más detectables en el sistema [25].

- Variaciones en el **tiempo de radiaciones de un Púlsar**: un Púlsar es un tipo

de estrella de neutrones que se caracteriza principalmente por su extraordinaria densidad, por su elevada magnetización y por rotar rápidamente. Típicamente se forman a partir de explosiones supernovas. La gran mayoría han sido revelados por su radiación en longitudes de onda radio. El hecho de que el Púlsar rote y emita radiación a lo largo del eje magnético en forma de cono hace que las señales recibidas por estos cuerpos sean pulsos [68][69].

Como los otros dos fenómenos que se han tratado en este apartado, los pulsos son regulares. De esta forma, si hay alguna variación, puede existir un exoplaneta que esté alterando el sistema [25].

A partir de este método se pueden estudiar las órbitas y la masa de los planetas. Además, la precisión de esta técnica es muy alta, de forma que podrían descubrirse exoplanetas cuya masa fuera hasta un décimo de la terrestre [70]. Sin embargo, los púlsares no solo emiten frecuencias de radio, sino que también pueden emitir rayos X y rayos gamma. Esto hace que los exoplanetas así descubiertos no sean adecuados para la existencia de vida, al menos en los términos que se conoce en nuestro planeta [25].

- **Variaciones en el tiempo de pulsaciones:** por último, también existen exoplanetas orbitando estrellas variables pulsantes. Estas estrellas se caracterizan por estar constantemente contrayéndose y extendiéndose. Esto es debido a no estar en equilibrio hidrostático, lo que hace que busquen el equilibrio entre las fuerzas de la gravedad y la presión de radiación [71].

Cuando estas pulsaciones se dan de forma muy regular, entonces puede estudiarse si hay variaciones debidas a planetas [25].

Modulaciones del brillo orbital

Cuando un planeta orbita una estrella, este puede reflejar la luz estelar, de tal forma que el brillo captado por los telescopios sufre variaciones [11, p. 87]. Estas modificaciones en la radiación no son muy grandes por lo general, sin embargo, como las fases orbitales del exoplaneta son periódicas, los cambios provocados por su reflejo también lo serán, lo que facilita su distinción [72, p. 8].

Cuanto mayor sea el planeta y más cerca orbite de la estrella, más fácil será su detección a partir de este método. Además, cuanto mayor sea el albedo de la superficie, más aumentará la radiación estelar reflejada [11, p. 87].

Una ventaja a destacar de esta técnica es que el planeta no tiene por qué transitar entre el observador y la estrella, pues con que describa una órbita pequeña a su alrededor ya será detectable el reflejo de luz que produce.[73]

Cinemática del disco

Los discos protoplanetarios son discos de acreción integrados por gas y polvo que se forman alrededor de una estrella joven en proceso de formación (protoestrella) y que pueden llevar a la creación de planetas [74, p. 170].

El objetivo al observar estos discos es

llegar a concluir el proceso por el cual se forman los planetas. A la derecha (Figura 2.23) se puede ver la imagen tomada por ALMA; en ella se reconocen estructuras como anillos y brechas en el gas y polvo, cuyo estudio puede dar lugar a comprender pasos importantes en el desarrollo de los exoplanetas [75].

El telescopio ALMA (*Atacama Large Millimeter Array*), está compuesto por 66 antenas y situado en el Observatorio Europeo Austral, en Chile; estudia objetos fríos del Universo [76].

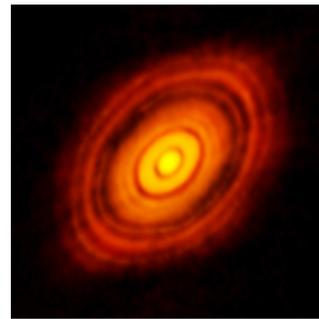


Figura 2.23: [75] Imagen tomada por ALMA del disco entorno a la estrella HL Tau.

Polarización de la luz estelar

La luz estelar por sí sola no presenta ninguna polarización. Sin embargo, cuando esta es reflejada por la atmósfera de un exoplaneta, sí que la presenta. El estudio de esta polarización, que dependerá del ángulo con el que la luz incida en el planeta, es decir, de la fase orbital, permite saber si existe un exoplaneta orbitando la estrella [11, pp. 89-90][72, p. 8].

Hasta la fecha, este método no ha tenido como consecuencia el descubrimiento

de ningún nuevo exoplaneta. Aún así, sí que se ha estudiado la polarización causada por exoplanetas revelados con otros métodos, como es el caso de HD 189733 b [11, p. 90][72, p. 8].

2.2. Redes Neuronales

Una red neuronal es un modelo computacional constituido por neuronas o nodos organizados en diferentes capas y conectados entre sí [77].

Se diferencian tres tipos de capas, según la función principal de cada una de estas [77]:

- **Capa de entrada:** los nodos que forman esta capa reciben los datos sobre los que la red trabajará; está constituida por tantos nodos como atributos tienen los datos.
- **Capas ocultas:** puede ser una o más. Reciben información de los nodos anteriores, y la envían una vez han trabajado sobre ella a capas posteriores.
- **Capa de salida:** a partir de los nodos de esta última capa se obtiene la información deseada (en este caso, la clasificación del dato de entrada). El número de nodos que la forman es igual a las diferentes clases posibles.

El objetivo principal de una red neuronal es que identifique correctamente la etiqueta o clase de la muestra sobre la que se aplique. El procedimiento es el siguiente: se introduce en la red una muestra con sus atributos correspondientes (*input*), la información irá pasando de capa en capa entre las neuronas conectadas entre sí, hasta llegar a la última capa y obtener la clasificación realizada por la red (*output*) [77].

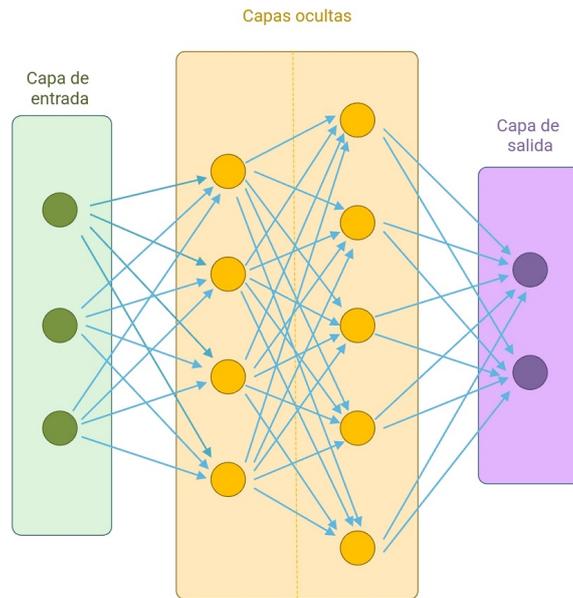


Figura 2.24: Esquema básico de una red neuronal. Imagen realizada usando *Paint* y la aplicación *FlowdiaDiagrams*.

Para conseguir el objetivo, la red se somete a un **entrenamiento** con el conjunto de datos *train*. Durante este proceso se irán modificando ciertos parámetros de la red hasta que se obtenga la combinación con la que se tiene una mayor precisión en la clasificación. Además, cada neurona podrá activarse o no, de tal forma que pasará la información a la siguiente capa o no dependiendo de valor de activación de la función empleada. Algunos de los parámetros son los pesos y los sesgos. Estos primeros se emplean para ponderar la información de cada nodo emisor para cada una de las neuronas que recibe la información, existe un valor diferente para cada conexión. Por otra parte, el sesgo (o *bias*) es característico de cada neurona, este valor se añade a la suma ponderada (ver Ec. 2.24 y Ec. 2.38) de la información que llega a la neurona k e influye en su activación [77].

2.2.1. Entrenamiento de la red

Para llevar a cabo el entrenamiento de la red, se diferencian dos tipos de datos [77]:

- **Datos de entrenamiento** (*training set*): con este conjunto de datos se entrena la red. Para ello, es necesario que las muestras estén etiquetadas, de tal forma

que se pueda estudiar el error que se está cometiendo.

- **Datos test** (*test set*): una vez la red haya sido entrenada, se aplicará el modelo sobre los datos test para conocer los resultados sobre muestras con las que la red no haya tratado nunca. De esta forma, se puede tener una idea de como sería la precisión real de los resultados del modelo creado sobre datos completamente desconocidos.

Se denomina **forward pass** al paso completo de la información de una muestra por todas las capas de la red. Cuando este proceso se haya cumplido para todos los datos de entrenamiento, entonces se habrá completado lo que se conoce como una **epoch**. Al final de cada *epoch* se estudia el error cometido comparando los resultados de la predicción con las etiquetas reales. Teniendo en cuenta las mejoras que se deben hacer, se inicia otra *epoch* con los mismos datos de entrenamiento pero los parámetros actualizados [77].

En general, en cada *epoch* se pasan todos los datos de entrenamiento a través de la red de uno en uno. Sin embargo, existe un hiperparámetro denominado **batch size** que permite pasar subconjuntos de datos a la red de forma simultánea. La ventaja de emplear este método es que computacionalmente puede disminuir el tiempo de entrenamiento de la red; sin embargo, se debe tener en cuenta que como consecuencia puede no obtenerse una optimización tan buena del modelo y disminuir su calidad [77].

Es interesante conocer cómo aprende la red para saber la modificación que tiene que realizar, por ejemplo, sobre los pesos y sesgos. Para ello se emplea un **algoritmo de optimización**, ya que lo que se está buscando en realidad es optimizar el modelo. En general, se emplea el optimizador denominado **descenso de gradiente estocástico (SGD)**. El problema es, en particular, de minimización, y la función que se quiere minimizar es la pérdida o el error que se está cometiendo en la predicción (**function loss**) [77].

En el entrenamiento, el cálculo del gradiente se lleva a cabo mediante el proceso conocido como **backpropagation**. Para explicarlo, se hará referencia únicamente a una muestra (un solo *forward pass*). La notación que se empleará es la siguiente [77]:

- $L \equiv$ número de capas de la red neuronal.
- $l \equiv$ índice de una capa.
- $j \equiv$ índice de un nodo de la capa l .
- $k \equiv$ índice de un nodo de la capa $l-1$.
- $y_j \equiv$ el valor de un nodo j en la capa de salida (L) para una única muestra de entrenamiento.
- $C_0 \equiv$ la función de pérdida o error de la red neuronal para una única muestra de entrenamiento.
- $W_j^{(l)} \equiv$ el vector de pesos de todas las conexiones entre los nodos de la capa $l-1$ con el nodo j de la capa l .
- $w_{jk}^{(l)} \equiv$ el peso de la conexión entre el nodo k de la capa $l-1$ y el nodo j de la capa l .
- $b_j^{(l)}$: sesgo o *bias* del nodo j en la capa l .
- $z_j^{(l)} \equiv$ el *input* del nodo j en la capa l .
- $g^{(l)} \equiv$ la función de activación empleada en la capa l .
- $a_j^{(l)} \equiv$ el valor de activación de salida del nodo j en la capa l .

El valor de la función loss en un *forward pass* se calcula según la ecuación siguiente:

$$C_0 = \sum_{j=0}^{r-1} \left(a_j^{(L)} - y_j \right)^2, \quad (2.23)$$

es decir, sumando sobre todos los nodos de la última capa los cuadrados de la diferencia entre la predicción (el valor de activación de salida de cada neurona) y el valor real que debería de tener el nodo [77].

Por otra parte, la entrada del nodo j en la capa l toma el siguiente valor [77]:

$$z_j^{(l)} = \sum_{k=0}^{r-1} w_{jk}^{(l)} a_k^{(l-1)}. \quad (2.24)$$

Cada neurona j recibirá la información emitida por los nodos de la capa anterior con los que esté conectada; estos valores de activación de cada nodo k ($a_k^{(l-1)}$) se sumarán de forma ponderada según los pesos de las conexiones correspondientes ($w_{jk}^{(l)}$). A su vez, el nodo j devolverá el resultado de aplicar la función de activación ($g^{(l)}$) sobre $z_j^{(l)}$ [77]:

$$a_j^{(l)} = g^{(l)} \left(z_j^{(l)} \right). \quad (2.25)$$

Teniendo en cuenta las expresiones que se acaban de definir, el error correspondiente a un nodo j depende de los siguientes parámetros [77]:

$$C_{0j} = C_{0j} \left(a_j^{(L)} \left(z_j^{(L)} \left(w_j^{(L)} \right) \right) \right), \quad (2.26)$$

y la pérdida debida a todos los nodos de la última capa se puede escribir como sigue [77]:

$$C_0 = \sum_{j=0}^{r-1} C_{0j}. \quad (2.27)$$

De esta forma, la derivada de C_0 con respecto a un peso concreto de una conexión con la última capa, $w_{jk}^{(L)}$, se calcula con derivadas parciales siguiendo la regla de la cadena [77]

$$\frac{\partial C_0}{\partial w_{jk}^{(L)}} = \left(\frac{\partial C_0}{\partial a_j^{(L)}} \right) \left(\frac{\partial a_j^{(L)}}{\partial z_j^{(L)}} \right) \left(\frac{\partial z_j^{(L)}}{\partial w_{jk}^{(L)}} \right). \quad (2.28)$$

Se concluye así que para una sola muestra, desarrollando la ecuación 2.28, la derivada del error con respecto a un peso $w_{jk}^{(L)}$ es [77]:

$$\frac{\partial C_0}{\partial w_{jk}^{(L)}} = 2 \left(a_j^{(L)} - y_j \right) \left(g'^{(L)} \left(z_j^{(L)} \right) \right) \left(a_k^{(L-1)} \right). \quad (2.29)$$

Si se computa el gradiente de la función *loss* con respecto al mismo peso genérico de la capa L una vez se ha completado un *epoch*, es decir, una vez ha pasado por la red el conjunto completo de datos de entrenamiento, el valor del error total tiene la siguiente expresión:

$$\frac{\partial C}{\partial w_{jk}^{(L)}} = \frac{1}{n} \sum_{i=0}^{n-1} \frac{\partial C_i}{\partial w_{jk}^{(L)}}, \quad (2.30)$$

siendo n el número de muestras de entrenamiento [77].

En el caso de la capa L , como se muestra en la ecuación (2.28), al calcular la derivada de la función *loss* con respecto a un peso $w_{jk}^{(L)}$, se tiene que tener en cuenta la siguiente relación de dependencia: en primer lugar, el error es función de la salida de activación del nodo concreto con respecto al que se esté calculando ($a_j^{(L)}$), esta a su vez depende del input de este nodo ($z_j^{(L)}$), que al mismo tiempo es función del peso en cuestión [77].

Sin embargo, cuando se trata de un peso de las capas ocultas, las relaciones varían. La función de pérdida no depende directamente de los pesos de estas capas, sino que depende de estos a través de los valores de los *outputs* de todos los nodos de la capa final [77],

$$\frac{\partial C_0}{\partial w_{jk}^{(L-1)}} = \underbrace{\left(\frac{\partial C_0}{\partial a_j^{(L-1)}} \right)}_{\text{término 1}} \underbrace{\left(\frac{\partial a_j^{(L-1)}}{\partial z_j^{(L-1)}} \right)}_{\text{término 2}} \underbrace{\left(\frac{\partial z_j^{(L-1)}}{\partial w_{jk}^{(L-1)}} \right)}_{\text{término 3}}. \quad (2.31)$$

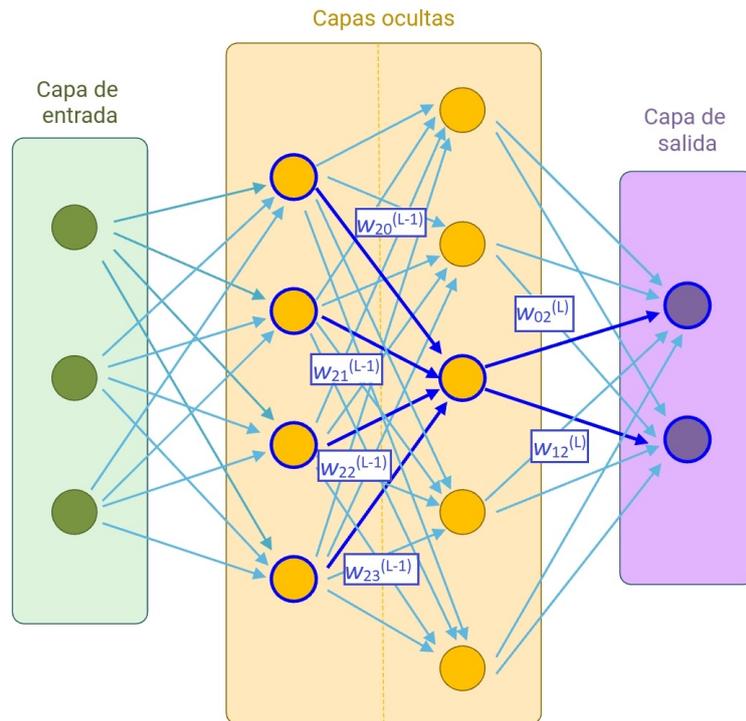


Figura 2.25: Esquema básico de una red neuronal sobre el que se indican los pesos de ciertas conexiones, para ver la dependencia del término 1 de la ecuación 2.31. Imagen realizada usando *Paint* y la aplicación *FlowdiaDiagrams*.

Los términos 2 y 3 de la ecuación 2.31 son exactamente iguales que los correspondientes en la capa L . El factor que tiene especial interés es el término 1 de dicha ecuación. La función C_0 depende de todas las salidas de activación de la última capa, estas son funciones de los inputs de cada neurona, que dependen de los *outputs* de la capa anterior $L - 1$. Se concluye entonces que el cálculo del término 1 es

$$\frac{\partial C_0}{\partial a_j^{(L-1)}} = \sum_{m=0}^{r-1} \left(\left(\frac{\partial C_0}{\partial a_m^{(L)}} \right) \left(\frac{\partial a_m^{(L)}}{\partial z_m^{(L)}} \right) \left(\frac{\partial z_m^{(L)}}{\partial a_j^{(L-1)}} \right) \right), \quad (2.32)$$

con m la variable que recorre todas las neuronas de la última capa [77].

Los dos primeros factores son idénticos a los correspondientes en la ecuación 2.28. En cuanto al último, se puede hacer la siguiente simplificación:

$$\frac{\partial z_m^{(L)}}{\partial a_j^{(L-1)}} \stackrel{\text{def. 2.24}}{=} \frac{\partial}{\partial a_j^{(L-1)}} \sum_{k=0}^{r-1} w_{mk}^{(L)} a_k^{(L-1)} = w_{mj}^{(L)}, \quad (2.33)$$

para lo que se ha empleado en la última igualdad que la derivada de una suma es la suma de las derivadas, y que es nula cuando $k \neq j$ [77].

Se concluye:

$$\frac{\partial C_0}{\partial a_j^{(L-1)}} = \sum_{m=0}^{r-1} 2 (a_m^{(L)} - y_m) (g'^{(L)}(z_m^{(L)})) (w_{mj}^{(L)}). \quad (2.34)$$

Al igual que ocurría para la derivada con respecto a un peso, cuando se estudian n muestras de entrenamiento se calcula la media [77]

$$\frac{\partial C}{\partial a_j^{(L-1)}} = \frac{1}{n} \sum_{i=0}^{n-1} \frac{\partial C_i}{\partial a_j^{(L-1)}}. \quad (2.35)$$

Se ha visto entonces que, para la obtención del gradiente, el procedimiento consiste en el cálculo de una serie de derivadas parciales sobre parámetros desde la última capa hasta la de la conexión con el peso correspondiente, de donde deriva el término *backpropagation* [77].

Así, se llega a conocer la dirección en la que se deben de modificar los pesos para llegar al mínimo del error. Este valor (para cada peso) se multiplica por la **tasa de**

aprendizaje (*learning rate*), que es un parámetro en el intervalo [0.01, 0.0001]. Dependiendo del valor asignado al *learning rate* el cambio que se produce en la dirección para minimizar el error será mayor o menor [77].

Cada peso se actualizará como sigue:

$$\text{peso actualizado} = \text{peso sin actualizar} - (lr \cdot \text{gradiente}) \quad (2.36)$$

siendo *lr* el *learning rate*.

El proceso se hace de forma reiterada hasta alcanzar el número de *epochs* determinado para el entrenamiento, o hasta llegar a un gradiente nulo, lo que significaría que hay un 100% de aciertos con el modelo correspondiente, puesto que no se está cometiendo ningún error [77].

Durante la actualización de los pesos nos podemos encontrar con dos problemas opuestos que llevan a un entrenamiento fallido: **gradiente de desaparición** o *vanishing gradient* y **gradiente de explosión** o *exploding gradient*. El primero de ellos se produce cuando el gradiente de la función *loss* con respecto a los pesos es demasiado pequeño (menor que 1), de manera que el cambio en los pesos será mínimo. Además, este valor es disminuido aún más al multiplicarlo por la tasa de aprendizaje. La consecuencia es la permanencia en un estado de entrenamiento parado, sin llegar a la optimización buscada [77].

Por otra parte, lo que se conoce como explosión del gradiente puede entenderse como un obstáculo en el entrenamiento opuesto al de desaparición. En este caso el gradiente es un número grande (mayor que 1) y los cambios en los pesos serán muy radicales [77].

Los pesos más afectados por ambos problemas son los correspondientes a las conexiones entre neuronas que constituyen las primeras capas de la red. Esto es debido a que, como se ha visto al tratar el *backpropagation*, cuantas más capas sigan a la de la conexión correspondiente, más factores intervendrán en el producto del que resulta el gradiente. De esta forma, si los productos son menores que uno, el valor final será más próximo a 0 cuantos más factores se multipliquen. Del mismo modo, si son números grandes, el producto de ellos será aún mayor [77].

Una solución común para la inestabilidad del gradiente es inicializar los pesos de forma no aleatoria. En general, en un entrenamiento de una red neuronal, los pesos se establecen al inicio aleatoriamente según una distribución normal de media 0 y desviación típica 1. Para entender el objetivo en la determinación de estos pesos iniciales, ha de hacerse referencia a la reproductividad de la variable normal [77]:

$$X_i \sim N\left(\mu_i, \sqrt{\sigma_i^2}\right) \Rightarrow Y = \sum_i X_i \sim N\left(\sum_i \mu_i, \sqrt{\sum_i \sigma_i^2}\right). \quad (2.37)$$

Recordando la definición del *input* de un nodo (def. 2.24)

$$z_j^{(l)} = \sum_{k=0}^{n-1} w_{jk}^{(l)} a_k^{(l-1)},$$

y que se acaba de comentar que los pesos iniciales siguen una distribución $N(0, 1)$, se tiene que la desviación típica de la variable que define este *input* será mucho mayor que 1. Luego, la entrada de una neurona puede tomar valores muy divergentes entre sí. En todos los tipos de funciones de activación (ver figura Fig. 2.27), valores extremos y opuestos tienen imágenes contrarias y extremas, por lo que esta gran desviación típica lleva consigo unos resultados muy dispares en el entrenamiento. El objetivo entonces es disminuir la desviación típica de la distribución normal que caracteriza los pesos, de tal forma que sea menor que uno [77].

Como se ha mencionado previamente, durante el entrenamiento también se actualizan los sesgos de cada neurona. Cuando una neurona tiene un sesgo no nulo, el *input* correspondiente sobre el que se aplicará la función activación es [77]:

$$z_j^{(l)} = \sum_{k=0}^{n-1} \left(w_{jk}^{(l)} a_k^{(l-1)}\right) + b_j^{(l)}. \quad (2.38)$$

La agregación de este sumando al valor de entrada puede influir completamente en la activación o no de una neurona. Si se quiere variar el umbral a partir del cual se desea que la neurona emita información (esté activada), entonces el sesgo correspondiente es

el siguiente:

$$z_j^{(l)} = \sum_{k=0}^{n-1} \left(w_{jk}^{(l)} a_k^{(l-1)} \right) \underbrace{- \text{umbral}}_{\text{sesgo}}. \quad (2.39)$$

Para verlo claro se presenta un ejemplo con la función activación ReLU (Ver figura 2.27):

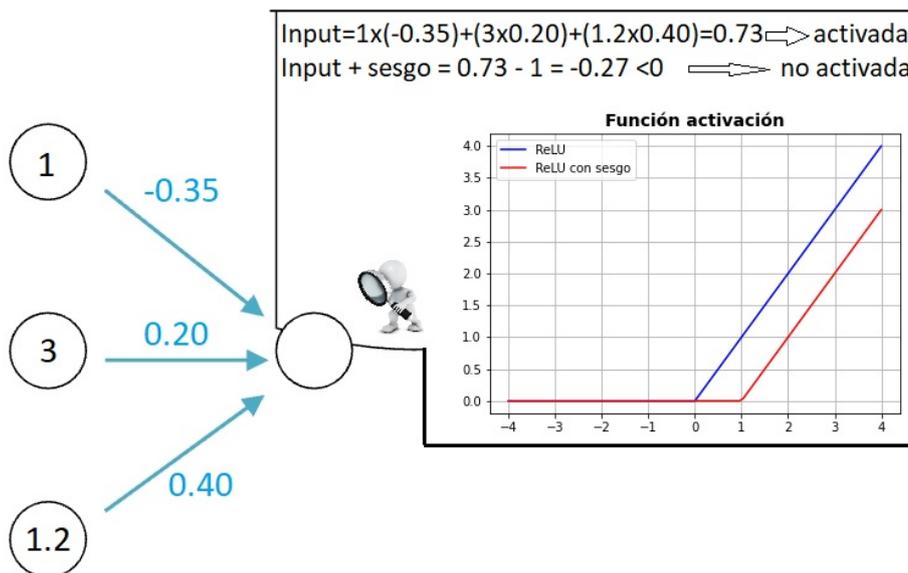


Figura 2.26: Explicación esquemática funcionamiento del sesgo con una función de activación ReLU. Imagen realizada con la ayuda de *Python*, *Paint*, la aplicación *FlowdiaDiagrams* y con la figura de la fuente [78].

2.2.2. *Overfitting* y *Underfitting*

Existen dos problemas muy usuales en el entrenamiento de la red: *overfitting* y *underfitting*. Puede ocurrir que la precisión del modelo sea muy alta sobre los datos de entrenamiento, pero no ocurra lo mismo sobre los datos test. Esto es debido a que en ocasiones la red ha entrenado de forma que ha interiorizado rasgos demasiado específicos de los datos de entrenamiento y no es capaz de generalizarlo a otras muestras. El caso contrario es el de *underfitting*, que ocurre cuando la red ni siquiera es capaz de predecir de forma correcta las etiquetas del conjunto de entrenamiento [77].

Existen soluciones para ambos casos. En primer lugar, para no llegar a un sobreajuste se tienen las siguientes alternativas [77]:

- Minimizar el número de *epochs*. Así, la red no entrenará tantas veces sobre los

mismos datos, no dándole la posibilidad de que ajuste los parámetros demasiado.

- Añadir más datos al conjunto de entrenamiento. Al hacer esto, además de tener un mayor número de muestras sobre las que entrenar, también se está aumentando la variedad de características, de tal forma que el rango de variables en los datos alcanzado por la red sea mayor. Se puede incrementar el número de datos de entrenamiento reuniendo más en el estudio o modificando los que ya se tienen. Esta segunda opción consiste en variar los rasgos o características de las muestras de las se dispone de forma razonada; por ejemplo, si se tratase de imágenes, se podrían modificar rotándolas, recortándolas, cambiando los colores o haciendo zoom.
- **Dropout**: cuantas más neuronas tenga la red, más capacitada estará para ajustar los parámetros y hacer una mejor predicción. Lo que se busca con el *dropout* es anular la actividad de ciertos nodos, de tal forma que no puedan intervenir en el entrenamiento del modelo.
- **Regularización**: otra técnica para disminuir el sobre ajuste es reducir el valor de los pesos de las conexiones de la red. Para hacer esto, se busca aumentar el resultado de la función *loss*, para que al calcular el gradiente se sigan actualizando los pesos de forma que disminuyan sus valores. Así, ninguna característica será demasiado determinante en la clasificación y será más probable que la red generalice bien los resultados. Destacan dos tipos de regularización [79] [80, pp. 146-149]:

- **L2**: En este caso el término que se suma a la función de pérdida depende de la norma de las matrices de pesos de cada capa [79] [80, pp. 146-149]:

$$loss + \left(\sum_{l=1}^n \|W^{[l]}\|^2 \right) \frac{\lambda}{2m} \quad (2.40)$$

donde:

- $n \equiv$ número de capas.
- $W^{[l]} \equiv$ matriz de pesos de la capa l .

- $m \equiv$ número de *inputs*.
- $\lambda \equiv$ parámetro de regularización. Dependiendo de cómo de grande sea su valor, más o menos aumentará el resultado de la función de coste.
- **L1:** En este caso la norma no se eleva al cuadrado [79] [80, pp. 146-149]:

$$loss + \left(\sum_{l=1}^n \|W^{[l]}\| \right) \frac{\lambda}{2m} \quad (2.41)$$

La diferencia entre ambos tipos es principalmente que en el segundo los pesos pueden reducirse hasta anularse y en el primero, aunque puedan ser casi cero, nunca se anulan por completo. Esto se traduce en la práctica a que con $L1$ el modelo puede dejar de considerar ciertas características de los datos para la clasificación, pero con $L2$ todas intervendrán, por poco que sea [79] [80, pp. 146-149].

Construir una red, entrenarla, y finalmente observar que está sobreajustando al aplicarla sobre el conjunto de datos test, es un problema que ralentiza mucho el procedimiento para obtener un modelo que pueda aplicarse. Para evitar en la medida de lo posible encontrarse con este error una vez finalizado todo el entrenamiento, se introduce otro conjunto de datos: **datos de validación** (*validation set*). Este conjunto se diferencia tanto de los de entrenamiento como de los de test. Al final de cada *epoch* se aplicará el modelo sobre este nuevo grupo con la finalidad de ir contrastando el error y la precisión en el *training set* frente al error y la precisión en unos datos sobre los que la red no ha entrenado, ni ha tratado en ningún momento. Todos los datos de este conjunto estarán, al igual que los de entrenamiento, etiquetados, lo que permite el cálculo del error y la precisión [77] [80, p. 43].

Por otra parte, las soluciones que previenen del *underfitting* se basan mayoritariamente en proceder de forma contraria al caso anterior [77]:

- Maximizar el número de *epochs* para permitirle entrenar más sobre el conjunto de entrenamiento.
- Construir un modelo más complejo aumentando por ejemplo el número de neuronas o de capas.

- Relacionado con el punto anterior, es conveniente también reducir el *dropout*, puesto que si lo que se busca en este caso es un modelo más complejo, disminuir el número de neuronas activadas no disminuiría la flexibilidad de la red.
- Añadir más datos al conjunto de entrenamiento, aportando más ejemplos sobre los que poder entrenar.

2.2.3. Parámetros variables en una red neuronal

El modelo que se ha introducido es la base de una red neuronal, sin embargo, hay una variedad infinita de redes. Además de poder variar el número de capas y de nodos, también pueden emplearse diferentes tipos de capas, de funciones de activación, de funciones de pérdida para medir la precisión del modelo... A continuación se introducen los tipos más destacados de cada uno.

En primer lugar, el tipo de capa empleado depende del objetivo que se quiere perseguir con la red neuronal:

- **Capas convolucionales:** son tipos de capas que se usan fundamentalmente para el reconocimiento de imágenes. Más adelante se desarrollarán las redes neuronales convolucionales.
- **Capas recurrentes:** la función de estas capas es identificar patrones en secuencias, como en una serie temporal o en textos de lenguaje natural.
- **Capas densas:** se caracteriza por ser una capa en la que todas las neuronas están conectadas con todas las de la capa anterior y las de la posterior.
- **Capas *Pooling*:** estas capas logran reducir el tamaño de las salidas, reduciendo la cantidad de datos a través de la media o del máximo.
- **Capas de normalización.**

En cuanto a las funciones de activación, destacan tres tipos:

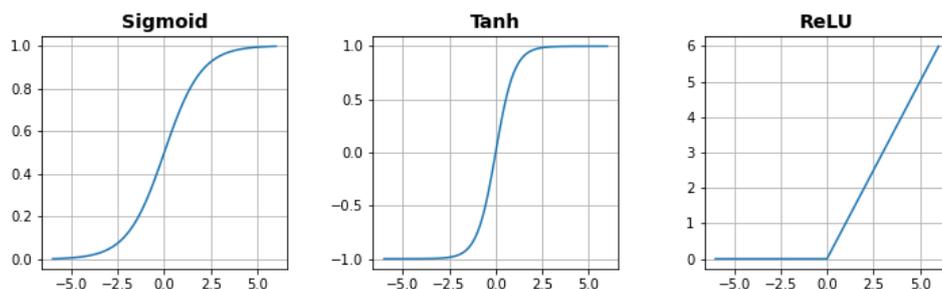


Figura 2.27: Representación de tres tipos de funciones de activación: Sigmoid, Tanh y ReLU. Imagen realizada con *Python*.

- **Función de activación Sigmoid:** Como se puede observar en la figura, esta función diferencia tres clases de valores. Cuando el número es muy negativo, la función devuelve un valor muy cercano a cero; en caso contrario, si es muy positivo, la salida será muy cercana a uno. Si el número de entrada es casi nulo, se transformará en un valor entre el cero y el uno [80, pp. 70-71] [77]. La neurona que se active dependiendo de esta función, estará más activada cuanto más positivos sean los números de entrada [77].
- **Función de activación ReLU (*Rectified linear unit*):** esta función devuelve el mismo número si este es positivo, o cero si es negativo. En este caso, si el *input* de la función es negativo o nulo, entonces la neurona no se activará [80, pp. 70-71] [77].
- **Tangente hiperbólica:** como el propio nombre indica, devuelve la tangente hiperbólica del número sobre el que se aplique [80, pp. 70-71] [77].

La función *Sigmoid* suele emplearse para los nodos de la última capa, ya que sus valores están comprendidos entre cero y uno, lo que facilita la interpretación de la salida en una clasificación binaria. Por otra parte, las otras dos funciones se emplean en capas intermedias, siendo más usual la función ReLU por ofrecer un entrenamiento de la red más rápido.

A continuación se introducen las diferentes funciones que pueden emplearse para estudiar el error o la precisión del modelo:

- **Error medio cuadrático:** se basa en el cálculo de la diferencia entre la clase o etiqueta real de la entrada, elevado al cuadrado. Cuando no se está estudiando

únicamente un dato, sino un conjunto de ellos, entonces el resultado sería la media de todos ellos [80, p. 86].

$$\text{error medio cuadrático} = (\textit{output} - \text{valor real})^2$$

- **Error medio absoluto:** la idea es muy similar a la del error anterior, diferenciándose en que en este caso se calcula el valor absoluto de la diferencia en vez de el cuadrado [80, p. 86].

$$\text{error medio absoluto} = |\textit{output} - \text{valor real}|$$

- **Error medio porcentual absoluto:** calcula el valor absoluto de la diferencia relativa al valor real; posteriormente se calcula la media sobre todos los datos y se multiplica por 100, para obtener el porcentaje [80, p. 86].

$$\text{error medio absoluto (\%)} = |(\textit{output} - \text{valor real})/\text{valor real}|$$

- **Entropía cruzada binaria:** se emplea en problemas de clasificación binaria, en los cuales la salida de la red neuronal es la probabilidad de pertenecer a una clase determinada (entre cero y uno) [80, p. 86].
- **Entropía cruzada categórica:** la idea se basa en la anterior, se diferencia en que en este caso se emplea para una clasificación entre un mayor número de clases [80, p. 86].

2.2.4. Normalización de los datos

Tener escalas de datos muy diferentes puede afectar negativamente en el entrenamiento de la red. Sin embargo, es una realidad muy habitual, puesto que las características que definen las muestras pueden tomar valores muy extremos o también estar definidas en escalas totalmente dispares unas de otras. Esto aumentaría notablemente el tiempo de entrenamiento. Además, valores muy altos llevarían como ya se ha visto

a la *explosión* del gradiente. Una solución para este último caso es la regularización, tratada en un apartado anterior.

En primer lugar, para evitar problemas por no tener datos en la misma escala, se han de procesar estos mismos datos antes de aplicar la red sobre ellos. Los procesos más comunes son el de normalización y el de estandarización. El primero de ellos consiste en transformar los datos en el intervalo $[0,1]$. El segundo se basa en restar a todos los datos la media de todos ellos y dividirlos entre la desviación estándar, de tal forma que el nuevo conjunto tenga como media 0 y como desviación estándar 1. En la práctica habitualmente se refiere a ambos como *normalización*.

Aún así, puede darse otro problema que lleve a una inestabilidad del gradiente, pues se pueden presentar pesos numéricamente muy elevados por destacar mucho más una propiedad entre todas ellas, llevando consigo una salida de la neurona muy grande. En este caso se debe de aplicar también una estandarización a la salida de la función de activación, proceso que se denomina *Batch Normalization*. Este tratamiento consta de tres partes:

1. Estandarizar: $z = \frac{\text{output} - \text{media}}{\text{desviación estándar}}$.
2. Multiplicar por un parámetro: $z \cdot a$.
3. Sumar por otro parámetro: $z \cdot a + b$.

Los parámetros a y b serán también optimizados durante el entrenamiento. Se realiza para cada batch.

2.2.5. *Fine-tuning*

Es muy común en todos los ámbitos de la vida, sea cual sea el problema que se quiere solucionar o la investigación que se esté realizando, emplear conocimiento ya adquirido en otras situaciones que pueden ser más o menos parecidas a la que se desee tratar en ese momento. El *Fine-tuning* está relacionado con esta idea, es decir, con lo que se entiende por *transfer learning*.

Al crear una red neuronal para que sea capaz de realizar una clasificación determinada, puede existir otro modelo que sirva como base y evite tener que realizar un

entrenamiento de la red desde el inicio. Por ejemplo, si se tiene un modelo que diferencia gatos en imágenes, y se busca ahora crear otro que reconozca perros, son muchas las características que pueden compartir ambos.

En este sentido, variables como el número de capas, los tamaños de los filtros, el orden de los tipos de capas o el número de neuronas en las capas densas pueden ser determinadas a partir del modelo que ya se tiene entrenado. Por otra parte, se recuerda que cuanto más inicial es la capa de la red, más superficiales son los objetos o las figuras que reconoce, por lo que los pesos de las primeras capas ocultas del modelo original también pueden emplearse en la variante del que se busque construir.

Así, el *Fine-tuning* es el método que describe este *reciclaje*. Permite dejar invariantes los pesos originales cuyo éxito ya se tiene garantizado y seguir el entrenamiento a partir de variables ya ajustadas con modelos entrenados con fines similares [77].

2.2.6. Redes Neuronales Convolucionales

Las Redes Neuronales Convolucionales (CNNs) son un tipo de redes neuronales cuya principal función es realizar la clasificación de imágenes. Los seres humanos estamos capacitados para diferenciar patrones en una imagen y saber identificar objetos. El objetivo de las CNNs es otorgar la misma habilidad a un ordenador, proceso que se conoce como *computer vision*.

Siguen el mismo esquema de estructura visto en el modelo presentado, con la capa de entrada, la de salida y las capas ocultas, y pueden tener diferentes tipos de capas; sin embargo, se caracterizan por presentar capas convolucionales. El procedimiento es también igual que el que se ha visto hasta ahora; los datos de entrada llegan a la capa, se procesa la información y se envía la salida a la siguiente. En este tipo de capas las entradas reciben el nombre de *input channels* y las salidas el de *output channels*. Por otra parte, las transformaciones de la información en las capas convolucionales también recibirán un nombre concreto; en *deep learning*, se denominan operaciones convolucionales, mientras que, en términos matemáticos, se refiere a ellas como correlaciones cruzadas [77].

Cada capa convolucional tiene un número determinado de filtros que se debe de especificar. Cada uno de estos filtros tiene como función reconocer una forma deter-

minada. Los primeros se centran en identificar los patrones más básicos; como rectas, curvas, esquinas... A medida que la información va llegando a capas más profundas, los filtros tienen como objetivo detectar patrones cada vez más complejos, hasta alcanzar la figura completa que se buscaba identificar [77].

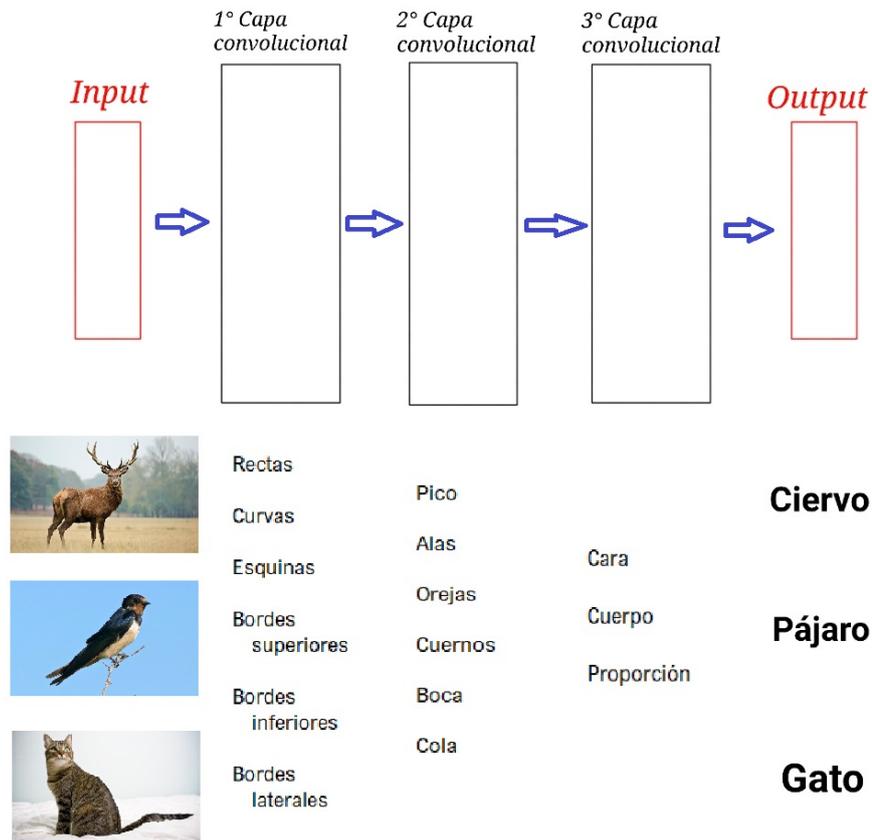


Figura 2.28: Esquema de rasgos reconocidos por los filtros al profundizar en más capas convolucionales. Figura realizada con la aplicación *Paint*, empleando imágenes de las siguientes fuentes: [81] [82] [83].

Tanto los filtros como las imágenes se representan con matrices. En el caso de una imagen, cada elemento de la matriz representa un pixel de esta. Por otra parte, el tamaño de los filtros puede variar y debe de ser concretado por el programador; están formados por los siguientes elementos: 0,1, -1. Dependiendo de la situación de estos elementos en la matriz del filtro, este identificará una forma u otra. Es importante destacar también que para las imágenes en blanco y negro únicamente se necesita un canal para el estudio del color; sin embargo, si la imagen es a color, entonces se emplearán tres canales, uno para cada color primario (rojo, verde y azul).

Para entender mejor el funcionamiento de estas redes, nos centraremos a continuación en una única capa con un solo filtro.

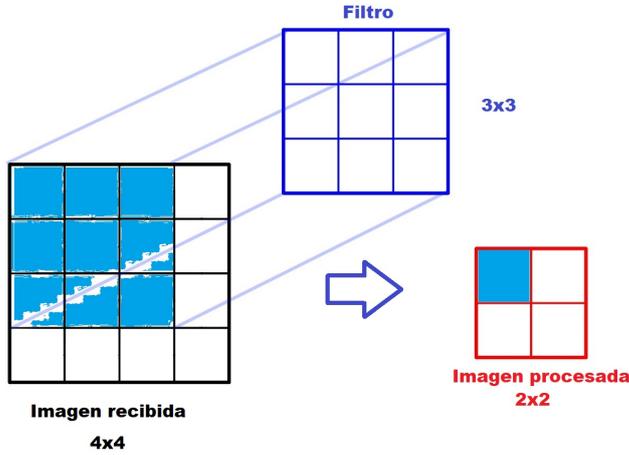


Figura 2.29: Aplicación de un filtro sobre una imagen de entrada. Imagen realizada con las aplicaciones *Write* y *Paint*.

La operación de convolución es un producto interno que generaliza el producto escalar, puede ser llamado Producto interno de Frobenius o suma del producto de Hadamard. En el ejemplo presentado, la operación del filtro sobre el primer conjunto de píxeles se define como sigue:

$$I \cdot F = i_{11} \cdot f_{11} + i_{12} \cdot f_{12} + \dots + i_{jk} \cdot f_{jk} + \dots + i_{33} \cdot f_{33} \quad (2.42)$$

siendo i el índice de la fila y j el de la columna de las matrices definidas a continuación:

$$I = \begin{pmatrix} i_{11} & i_{12} & i_{13} \\ i_{21} & i_{22} & i_{23} \\ i_{31} & i_{32} & i_{33} \end{pmatrix} \text{ y } F = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \quad (2.43)$$

En general, si el filtro que se aplica sobre la imagen tiene un tamaño $f \times f$ y la dimensión de la imagen es $n \times n$, entonces:

$$\text{tamaño del output} = (n - f + 1) \times (n - f + 1) \quad (2.44)$$

Pongamos de ejemplo un filtro 3x3 (representado en la figura de color azul) y una imagen 4x4 representada de color negro. El filtro realizará la operación de convolución sobre una submatriz de la imagen de su mismo tamaño (3x3) y se desplazará repitiendo el proceso. En este caso, el filtro recubre la imagen con cuatro posiciones diferentes posibles, por lo que la imagen procesada (representada de color rojo) es 2x2.

Debido a este proceso, es habitual que la imagen se vaya reduciendo a medida que pasa por capas convolucionales. Esto puede tener efectos negativos en la predicción de la red, puesto que se estarán perdiendo muchos datos importantes, especialmente si una parte esencial del patrón está situada en los bordes, puesto que sobre estos no se hacen tantas convoluciones como en los píxeles del centro.

Una técnica que permite resolver este problema es el *zero padding*. La idea de este método es añadir dos filas y dos columnas de valores nulos en los extremos de la matriz que representa la imagen. De esta forma, los bordes de la imagen se estudiarán más detalladamente y no se perderá información. Además, el tamaño de la imagen procesada será igual que la de la recibida por la capa [77].

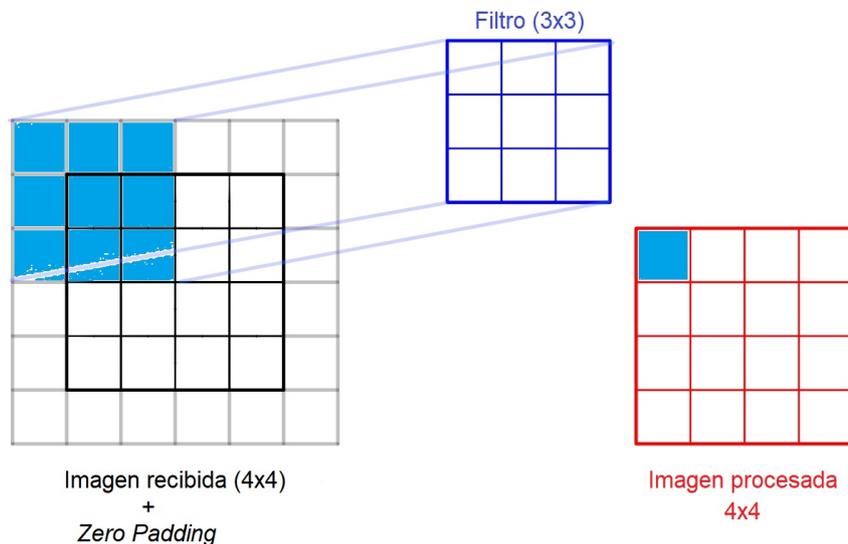


Figura 2.30: Esquema del funcionamiento del *zero padding*. Imagen realizada con las aplicaciones *Write* y *Paint*.

Hasta aquí se ha visto el funcionamiento básico de una red neuronal convolucional. A este modelo se pueden añadir diferentes capas con varias funciones, contruyendo así redes más complejas y con características más precisas. A continuación se explican algunas de las más usuales.

Como ya se ha mencionado previamente, existe un tipo de capa denominado *pooling*. Es habitual situar una de estas capas tras una convolucional. La más común es la llamada *Max Pooling*, cuya función es disminuir el tamaño de la imagen eligiendo los píxeles más activos entre subconjuntos. La idea es intentar optimizar la computación

quedándose únicamente con la información más relevante. Por ejemplo, en la Figura 2.28 los paisajes en los que están los animales o los fondos de cada una de esas imágenes son totalmente despreciables a la hora de predecir de qué animal se trata [80, pp. 212-213].

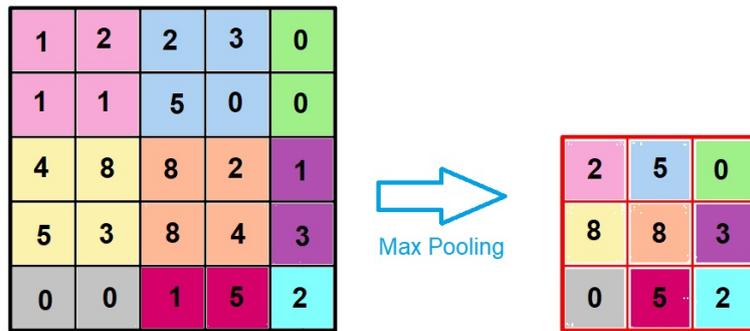


Figura 2.31: Esquema del funcionamiento de *Max Pooling*. Imagen realizada con las aplicaciones *Write* y *Paint*.

Existe otro tipo de capa *pooling* llamada *Average Pooling*. En vez de calcular el máximo del subconjunto de píxeles, calcula la media de todos los valores que lo forman [80, pp. 212-213]. Entendiendo ambas funciones como un filtro, además de definir el tamaño de los subconjuntos sobre los que se quiere aplicar (el tamaño del filtro), se ha de concretar el *stride*⁹, que es el número de píxeles que se desplaza el filtro entre cada operación. En el ejemplo de la figura anterior el tamaño del filtro es 2×2 y el *stride* es 2.

Son varias las ventajas de emplear este tipo de capas. En primer lugar, se reduce el coste computacional, debido a que se emite una señal con menos información. Sin embargo, es fundamental el hecho de que en este proceso no se pierden datos importantes, puesto que las funciones están definidas de tal forma que se guardan los píxeles más activados en el caso de *Max Pooling*, o se mantiene un promedio de los valores. Por último, también es importante que, al estar disminuyendo el detalle de la información,

⁹El *stride* no solo se emplea en los filtros de *pooling*, sino que puede emplearse para caracterizar cualquier proceso de filtrado de una red convolucional.

se puede reducir así también el *overfitting*.

Otra etapa a destacar en una CNN es el *Flattening*, que se lleva a cabo en el paso previo de introducir los datos en una red ANN. Como se ha visto anteriormente, las redes neuronales artificiales trabajan con información dispuesta en vectores, por lo que se ha de pasar de una matriz a un vector colocando una fila detrás de otra, desde la primera hasta la última (trasponiéndolo, puesto que se trabaja con vectores columna) [80, pp. 214-215].

2.2.7. Redes Convolucionales 1D

Las redes convolucionales vistas en el apartado anterior son redes convolucionales 2D (de dos dimensiones). Como se ha explicado, los datos de entrada, generalmente imágenes, se analizan procesando la información en dos ejes. Sin embargo, hay casos en los que la señal que se desea estudiar únicamente varía en una dirección. Cuando se da tal situación, se emplean **redes convolucionales 1D** (de una dimensión) [84].

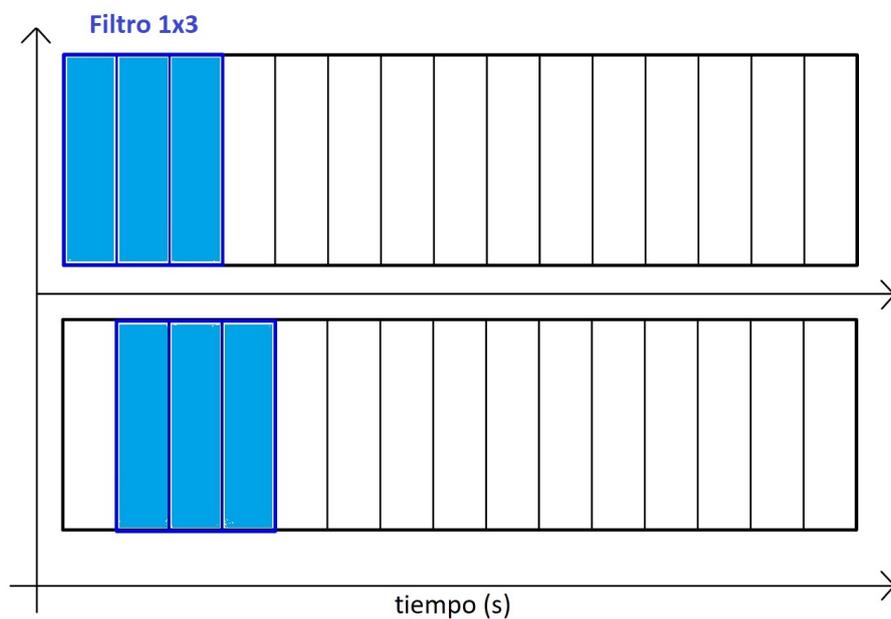


Figura 2.32: Esquema representativo de una red convolucional 1D construída por un filtro 3x1, siendo aplicada sobre una serie temporal con un stride igual a 1. Imagen realizada con las aplicaciones *Write* y *Paint*.

Típicamente con este método se estudian las series temporales, en las que la información se guarda a lo largo del tiempo. Otros ejemplos son todos los que se traten de

datos sensoriales, como los audios o los textos, puesto que una forma de tratarlos es también viéndolos como series temporales [84].

Capítulo 3

Metodología

3.1. Generación de curvas de luz artificiales

Para entrenar la red neuronal que se quiere construir se necesitan datos artificiales que representen los datos reales sobre los que se quiere aplicar el modelo una vez optimizado. Al ser el objetivo clasificar exoplanetas, es decir, que la respuesta de la red sea si **sí** hay un exoplaneta en la curva de luz o si, por lo contrario, **no** lo hay, se deben de crear curvas con y sin tránsito.

3.1.1. Curvas de luz con tránsito

La idea principal para la producción de este tipo de datos es superponer un tránsito con un periodo determinado sobre un flujo normalizado. Está claro que la parte más laboriosa de este apartado es crear los tránsitos. Para ello, se empleará el paquete *Batman* de *Python* [85].

A continuación se analizarán los parámetros que requiere la función de *Batman* para crear los tránsitos. Estos serán los necesarios para caracterizar tanto la estrella como el exoplaneta. Las estrellas que se emplearán en el estudio son las de la Secuencia Principal. En muchos casos, para determinar el valor de algunos de estos parámetros, se emplearán como base los estudios realizados por las misiones Kepler y TESS en la detección de exoplanetas. El proceso a seguir es el siguiente:

- En primer lugar, se define el **radio estelar**, R , de forma aleatoria en un intervalo

de $[0.3,2]R_{\odot}$.

- A partir del radio se calcula la **masa estelar**, M , correspondiente. Existen diferentes relaciones en función de la estrella que se estudie. La ecuación empírica que se emplea en este trabajo para relacionar estas dos variables para las estrellas de la Secuencia Principal es la siguiente [41]:

$$\frac{M}{M_{\odot}} = \left(\frac{R}{R_{\odot}} \right)^{\frac{1}{0.75}} \quad (3.1)$$

- Por otra parte, el **periodo orbital**, T , del planeta se crea también de forma aleatoria entre 2 y 30 días, de forma que se está en el rango de tiempo de estudio de la mayoría de los sectores de la misión TESS.
- Una vez se tienen los datos necesarios, se define el **semi-eje** de la órbita a partir de la Tercera Ley de Kepler:

$$a = \sqrt[3]{\frac{GMT^2}{4\pi^2}}, \quad (3.2)$$

siendo a el semi-eje mayor de la elipse en unidades astronómicas (UA) ($a=1UA$ para la Tierra), y G la constante gravitacional [86].

En este caso se puede emplear esta aproximación por ser la masa estelar dominante. Pero si se quiere ser más exacto, la ley debería de escribirse:

$$a = \sqrt[3]{\frac{G(M+m)T^2}{4\pi^2}}, \quad (3.3)$$

siendo m la masa del exoplaneta [86].

- La **inclinación orbital** es aproximadamente 90° , por esto, se le asigna un valor aleatorio entre 88° y 92° . Es decir, se desea crear exoplanetas en una situación lo más cerca posible de *edge-on* (ver Figura 2.11).
- La **excentricidad** puede tomar valores entre 0 y 1. En este caso, para simplificar el problema, se estudian órbitas aproximadamente circulares, por lo que la excentricidad se definirá en el intervalo entre 0 y 0.1.

- La relación entre el planeta y la estrella se determina a partir del cociente de ambos radios, rp . En esta definición se diferencia entre las estrellas de menor radio (enanas M de la Secuencia Principal) y las de mayor (el resto de estrellas de la Secuencia Principal).

- Si el radio estelar $\in [0.3, 0.7]R_{\odot}$, entonces:

$$rp \in [0.01, 0.04]. \quad (3.4)$$

- Si, en cambio, el radio estelar $\in [0.7, 2]R_{\odot}$:

$$rp \in [0.01, 0.1]. \quad (3.5)$$

Basándose en el estudio que realiza TESS¹, la duración de la toma de datos es de 30 días (como ya se mencionó en la definición del periodo orbital) y la cadencia de 30 minutos, cadencia empleada tanto en la misión extendida K2 como en las *Full Frame Image* de TESS. Es decir, se tendrán un total de 1440 puntos correspondientes a cada curva de luz.

Por otra parte, el ruido en los datos se introduce según una distribución gaussiana. La desviación típica se determina directamente a partir de la magnitud aparente de la estrella. Esta relación se fundamenta de nuevo en el estudio de la misión Kepler, en el cual se observó la precisión durante media hora dependiendo de la magnitud aparente de las estrellas [87]. A partir de estos datos y haciendo una regresión lineal para aumentar el intervalo de magnitudes aparentes se tiene la siguiente gráfica:

¹Se recuerda que aproximadamente los días que tarda TESS en estudiar cada región del cielo son 30.

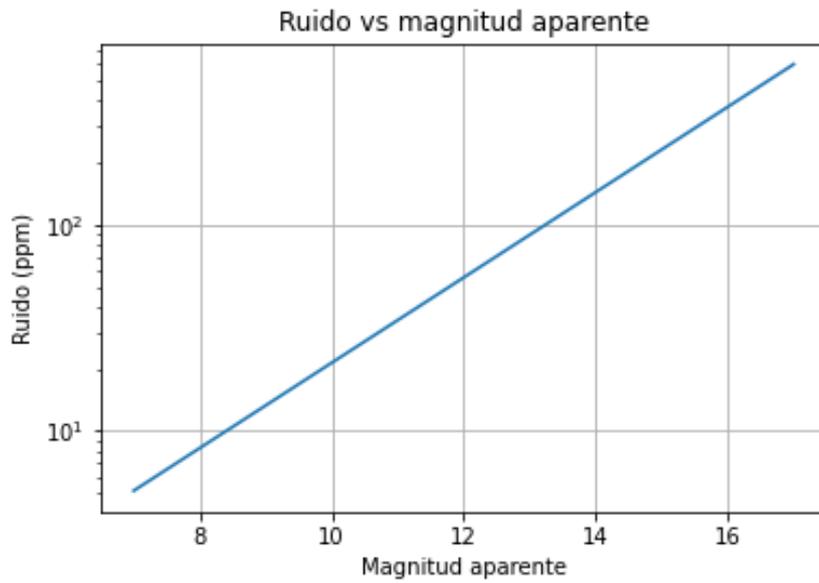


Figura 3.1: Relación entre el ruido de los datos en la curva de luz y la magnitud aparente de la estrella. Regresión lineal realizada a partir de la Figura 5 de [87].

Por otra parte, el flujo debe de estar normalizado a 1 para facilitar posteriormente el entrenamiento de la red (ver sección 2.2.4.), por lo que la media de la distribución gaussiana que define el ruido será 1.

A continuación se muestran tres ejemplos de curvas con tránsitos creadas siguiendo el procedimiento que se acaba de explicar. Se observa que en la producción de las curvas se está cubriendo todo el rango de variedades de tránsitos, desde los más fáciles de detectar (ver Figura 3.2) hasta los menos identificables visualmente (ver Figura 3.4).

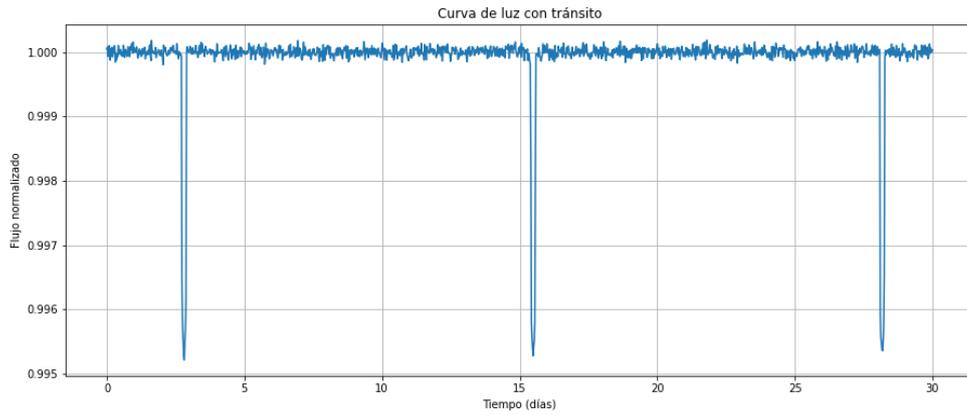


Figura 3.2: Curva de luz de una estrella que presenta tránsitos correspondientes a un exoplaneta. Las características del sistema estrella-exoplaneta son las siguientes: $M = 1.135M_{\odot}$, $R = 1.100R_{\odot}$, $rp = 0.064$, semi-eje = $2.380R_{\odot}$, inclinación orbital = 89.526° , excentricidad= 0.093, $T = 12.682$ días. La estrella tiene una magnitud aparente = 12.300 y un ruido gaussiano = 63.954ppm.

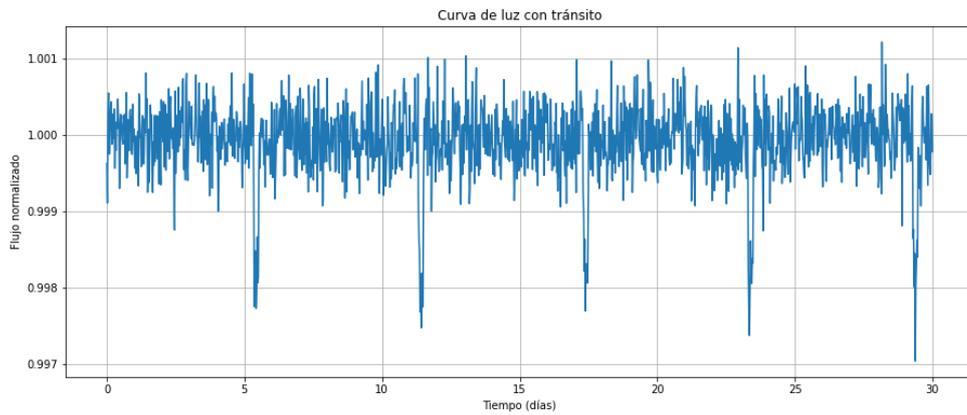


Figura 3.3: Curva de luz de una estrella que presenta tránsitos correspondientes a un exoplaneta. Las características del sistema estrella-exoplaneta son las siguientes: $M = 2.317M_{\odot}$, $R = 1.878R_{\odot}$, $rp = 0.041$, semi-eje = $18.313R_{\odot}$, inclinación orbital = 90.544° , excentricidad= 0.029, $T = 5.990$ días. La estrella tiene una magnitud aparente = 16.000 y un ruido gaussiano = 372.648ppm.

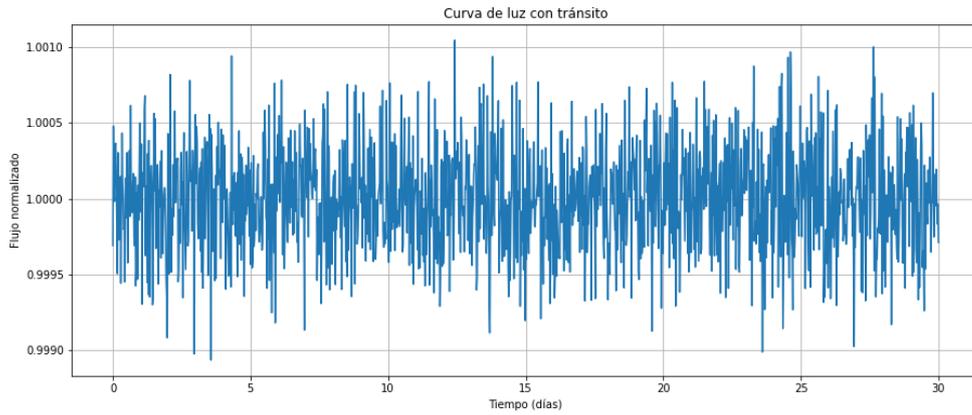


Figura 3.4: Curva de luz de una estrella que presenta tránsitos correspondientes a un exoplaneta. Las características del sistema estrella-exoplaneta son las siguientes: $M = 0.586M_{\odot}$, $R = 0.670R_{\odot}$, $rp = 0.054$, semi-eje = $29.530R_{\odot}$, inclinación orbital = 91.810° , excentricidad = 0.095, $T = 24.390$ días. La estrella tiene una magnitud aparente = 15.800 y un ruido gaussiano = 338.785ppm.

3.1.2. Curvas de luz sin tránsito

Un caso mucho más sencillo es la producción de curvas de luz sin tránsito. Para generar estos datos simplemente se tiene que recrear el flujo de una estrella, por lo que se sigue la parte final del proceso explicado en el apartado anterior. Únicamente se define de forma aleatoria la magnitud aparente en un intervalo entre 11 y 17, y se generan valores según la distribución gaussiana con media 1 y desviación típica la correspondiente según la gráfica 3.1. Se muestra a continuación un ejemplo de estas curvas:

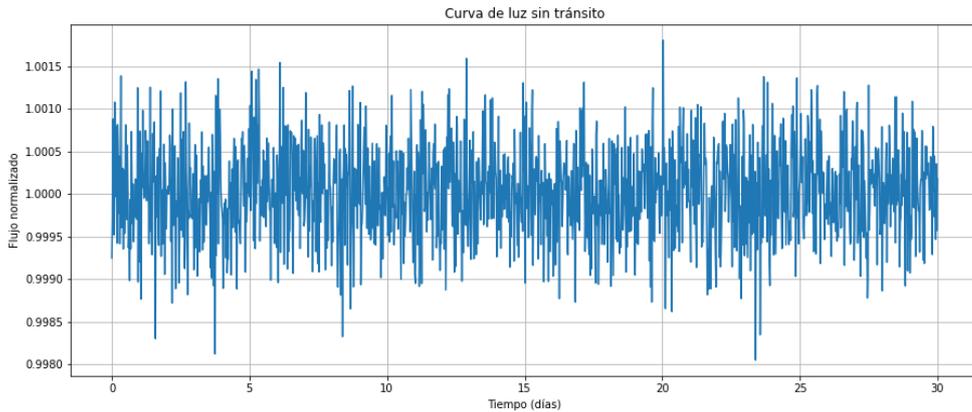


Figura 3.5: Curva de luz correspondiente a una estrella sin tránsito, con una magnitud aparente = 16.800 y un ruido gaussiano = 545.504ppm.

Como se puede observar, la diferencia entre la Figura 3.5 y la Figura 3.4 es mínima, de tal forma que se confirma que en los datos de entrada de la red hay tránsitos que para nada son obvios.

3.2. Definición Red Neuronal para la detección de exoplanetas

Estamos ante un problema en el que los datos (curvas de luz de las estrellas) varían en una sola dimensión, en el tiempo. Por ello, según lo visto en la sección *Redes Convolucionales 1D* (2.2.7), se construye una red convolucional en una dimensión.

Los parámetros que se estudian en la construcción de esta red neuronal son los cuatro más importantes: número de capas, tamaño del filtro, número de *strides* y número de capas convolucionales en cada filtro. Para cada uno de estos parámetros se definen dos valores diferentes, por lo que en total se tienen 16 combinaciones para contrastar.

El **número de capas** será 3 o 5, elegidos así para evitar un sobreajuste en la red. Por otra parte, el **tamaño del filtro** será 7 u 11. En este caso, se han determinado estos valores teniendo en cuenta que la duración de los tránsitos es de horas o pocos días. Por ello, se debe de definir el filtro de tal manera que no sea demasiado grande como para que los tránsitos más pequeños pasen desapercibidos, ni demasiado pequeño

como para que no se identifique el cambio en el brillo.

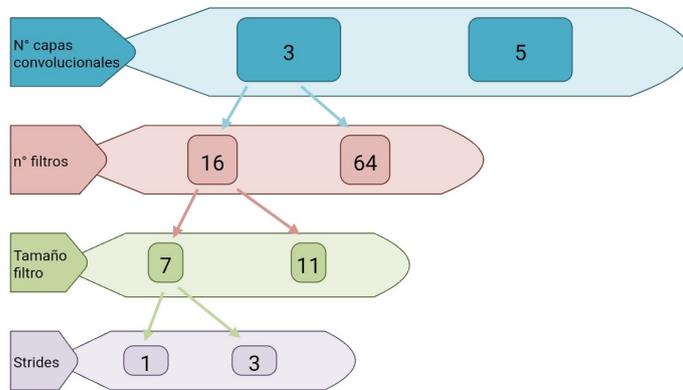


Figura 3.6: Esquema de los datos que toman los parámetros en el entrenamiento de la red. Imagen que son una buena aproximación para el entrenamiento de la red.

En cuanto al **número de strides**, será 1 o 3, pues no tiene sentido tomar un número mucho mayor. Cada nivel de la red tendrá 16 o 64 **capas convolucionales**. Este parámetro suele tomarse como una potencia de dos, y los números determinados se encuentran en la media, por lo que son una buena aproximación para el entrenamiento de la red.

Además, en este caso obviamente no hay que diferenciar colores en los datos de entrada (como ocurriría por ejemplo en el tratamiento de imágenes), por lo que no se necesitarán los tres canales correspondientes a los colores primarios.

Por último, se añade la capa de *Flattening*, transformando los datos en vectores para que puedan ser la entrada de la capa posterior, la capa *Dense*. Esta capa estará formada por 128 neuronas y la función de activación de cada una de ellas será la función *ReLU*. La última capa será también de tipo *Dense*, pero formada por 2 neuronas, con la función de activación *Softmax*, dando lugar así a las salidas correspondientes a la clasificación binaria.

Para el entrenamiento de la red se producen 1000000 de curvas de luz en total, la mitad con tránsito y la otra mitad sin tránsito. De todas ellas, el 80% se emplearán como entrenamiento y el otro 20% constituirá el conjunto de datos test. Un 10% de los datos de entrenamiento serán empleados como datos de validación, para evitar así el *overfitting* y poder analizar los resultados de la red en cada *epoch*.

Capítulo 4

Resultados y discusión

A continuación se presentan los resultados obtenidos para cada combinación de parámetros (explicadas en el capítulo anterior). Para la comparación del rendimiento de cada modelo, se contrastan los valores de *train accuracy* y *test accuracy*, así como los de *train loss* y *test loss* (los valores de entrenamiento considerados serán los de la última *epoch*). Además, se analizan también las clasificaciones obtenidas en cada caso, a partir de la matriz de confusión¹.

A partir de las tablas 4.1, 4.2 y 4.3, se observa que todas las combinaciones presentadas dan como resultado un modelo con un porcentaje de aciertos de aproximadamente el 90 %, a excepción de cuatro de ellos. Las cuatro combinaciones de parámetros con los que la red no entrena bien se presentan en las tablas en color rojo. En estos últimos casos, tanto el *train accuracy* como el *test accuracy* no superan el valor de *accuracy* del 0.5. Además, el valor de la función coste en ambos conjuntos de datos es aproximadamente del 0.7.

¹La matriz de confusión es una herramienta que permite conocer el rendimiento de las clasificaciones realizadas. Recorriendo las filas de arriba a abajo y de izquierda a derecha, los valores que se muestran son los siguientes: verdaderos negativos, falsos positivos, falsos negativos y verdaderos positivos.

Nº capas	Nº filtros/capa	Tamaño filtro	Strides	Train Accuracy	Train Loss
3	16	7	1	0.93798614	0.17750624
3	16	7	3	0.91173470	0.23790595
3	16	11	1	0.94234306	0.16927752
3	16	11	3	0.91139030	0.23805850
3	64	7	1	0.93309164	0.19275445
3	64	7	3	0.91337085	0.23410206
3	64	11	1	0.50036250	0.69318370
3	64	11	3	0.91262640	0.23554425
5	16	7	1	0.93741950	0.18231127
5	16	7	3	0.90640557	0.24785838
5	16	11	1	0.93130140	0.19448373
5	16	11	3	0.90595555	0.24869800
5	64	7	1	0.50033474	0.69318600
5	64	7	3	0.90475970	0.25183980
5	64	11	1	0.50069310	0.69318634
5	64	11	3	0.50006250	0.69318986

Tabla 4.1: Valores de *train accuracy* y *train loss* en la última *epoch* de cada uno de los modelos definidos con la combinación determinada.

Nº capas	Nº filtros/capa	Tamaño filtro	Strides	Test Accuracy	Test Loss
3	16	7	1	0.89356500	0.36440566
3	16	7	3	0.91024000	0.24088362
3	16	11	1	0.88597000	0.38888480
3	16	11	3	0.90178000	0.26284054
3	64	7	1	0.88992500	0.34994644
3	64	7	3	0.90384500	0.25144103
3	64	11	1	0.50000000	0.69322060
3	64	11	3	0.91003500	0.24064787
5	16	7	1	0.89187500	0.32709040
5	16	7	3	0.89982500	0.26830822
5	16	11	1	0.89205000	0.30301230
5	16	11	3	0.90613000	0.24813998
5	64	7	1	0.50000000	0.69315920
5	64	7	3	0.90691500	0.24887106
5	64	11	1	0.50000000	0.69317630
5	64	11	3	0.50000000	0.69317140

Tabla 4.2: Valores de *test accuracy* y *test loss* de cada uno de los modelos definidos con la combinación determinada.

Para entender lo que está sucediendo en los modelos que no llevan a una buena clasificación, se estudian las matrices de confusión correspondientes (ver Tabla 4.3). Para el modelo definido con los parámetros $(3,64,11,1)^2$, se observa que tanto los falsos

²Se hace referencia a los parámetros de cada modelo escribiéndolos en una tupla de 4 elementos en el siguiente orden: número de capas, número de filtros por capa, tamaño del filtro y *strides*.

positivos como los verdaderos positivos son 100000; es decir, los 200000 datos del conjunto test se están clasificando como curvas con tránsito. Ocurre lo mismo en el modelo (5,64,11,1).

Por el contrario, para las redes de parámetros (5,64,7,1) y (5,64,11,3), lo que ocurre es que todos se clasifican como estrellas sin exoplanetas en tránsito. Es decir, tanto los verdaderos negativos como los falsos negativos son 100000.

Ninguna de las dos situaciones anteriores es deseada, puesto que sea cual sea la curva de luz registrada, la red determinará siempre el mismo tipo de sistema planetario: o bien todas las estrellas presentan un planeta o ninguna lo hace.

Nº capas	Nº filtros/capa	Tamaño filtro	Strides	Matriz de confusión
3	16	7	1	$\begin{pmatrix} 96080 & 3920 \\ 17367 & 82633 \end{pmatrix}$
3	16	7	3	$\begin{pmatrix} 99463 & 537 \\ 17415 & 82585 \end{pmatrix}$
3	16	11	1	$\begin{pmatrix} 93998 & 6002 \\ 16804 & 83196 \end{pmatrix}$
3	16	11	3	$\begin{pmatrix} 94927 & 5073 \\ 14571 & 85429 \end{pmatrix}$
3	64	7	1	$\begin{pmatrix} 94464 & 5536 \\ 16479 & 83521 \end{pmatrix}$
3	64	7	3	$\begin{pmatrix} 96978 & 3022 \\ 16209 & 83791 \end{pmatrix}$
3	64	11	3	$\begin{pmatrix} 99280 & 720 \\ 17273 & 82727 \end{pmatrix}$
3	64	11	1	$\begin{pmatrix} 0 & 100000 \\ 0 & 100000 \end{pmatrix}$
5	16	7	1	$\begin{pmatrix} 96305 & 3695 \\ 17930 & 82070 \end{pmatrix}$
5	16	7	3	$\begin{pmatrix} 95799 & 4201 \\ 15834 & 84166 \end{pmatrix}$
5	16	11	1	$\begin{pmatrix} 95537 & 4463 \\ 17127 & 82873 \end{pmatrix}$
5	16	11	3	$\begin{pmatrix} 99422 & 578 \\ 18196 & 81804 \end{pmatrix}$
5	64	7	1	$\begin{pmatrix} 100000 & 0 \\ 100000 & 0 \end{pmatrix}$
5	64	7	3	$\begin{pmatrix} 98440 & 1560 \\ 17057 & 82943 \end{pmatrix}$
5	64	11	1	$\begin{pmatrix} 0 & 100000 \\ 0 & 100000 \end{pmatrix}$
5	64	11	3	$\begin{pmatrix} 100000 & 0 \\ 100000 & 0 \end{pmatrix}$

Tabla 4.3: Matrices de confusión de cada uno de los modelos definidos con la combinación determinada.

Para observar lo que ocurre durante el entrenamiento de las cuatro redes neuronales

con las que no se consiguen buenos resultados, se presentan a continuación las gráficas 4.1 y 4.2. En ellas se puede observar la evolución de *train accuracy* y *val accuracy* a medida que aumentan los *epochs*.

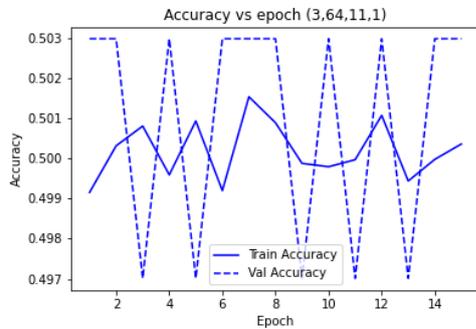


Figura 4.1: Representación de la evolución de *train accuracy* y *val accuracy* con los *epochs*. Modelo con parámetros (3,64,11,1).

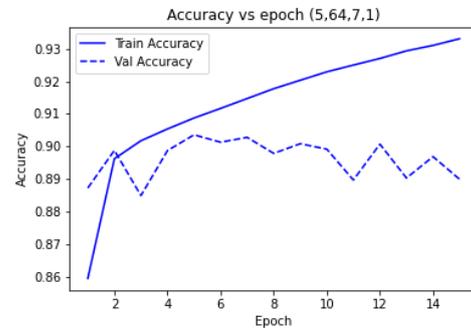


Figura 4.2: Representación de la evolución de *train accuracy* y *val accuracy* con los *epochs*. Modelo con parámetros (5,64,7,1).

En ninguno de los dos casos la red está entrenando. En la representación de la izquierda (ver Figura 4.1), los resultados son continuamente erróneos, variando entorno a 0.5 en todas las pruebas. Por otro lado, en la de la derecha (ver Figura 4.2) lo que ocurre es que la red se está aprendiendo las características de los datos de entrenamiento, pero no está aprendiendo a generalizarlos. Por esto, se observa una evolución adecuada en los valores de *train accuracy*, mientras que los de *val accuracy* permanecen prácticamente invariantes.

Entre todos los modelos que funcionan se tienen prácticamente los mismos resultados, todos ellos exitosos. Para poder estudiar uno más detenidamente, nos centraremos en el correspondiente a la combinación (3,16,7,3).

Observando las Figuras 4.3 y 4.4, se observa que en este caso la precisión va aumentando con cada *epoch* y el valor de la función coste disminuyendo, por lo que la red sí que está entrenando de forma adecuada. Se han empleado 15 *epoch* en el entrenamiento, sin embargo, a partir de las gráficas se puede ver que desde la décima *epoch*, la red comienza a mejorar los resultados únicamente en el conjunto *training set*. Esto es un hecho que se quiere evitar, como ya se ha comentado en varias ocasiones, por lo que se concluye que lo ideal sería entrenar la red hasta 10 *epochs*.

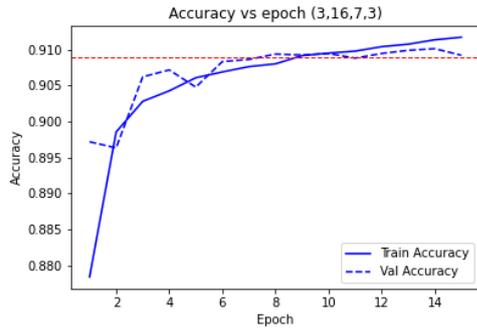


Figura 4.3: Representación de la evolución de *train accuracy* y *val accuracy* con los *epochs*. Modelo con parámetros (3,16,7,3). Se representa en rojo la precisión en la que se estabiliza el modelo.

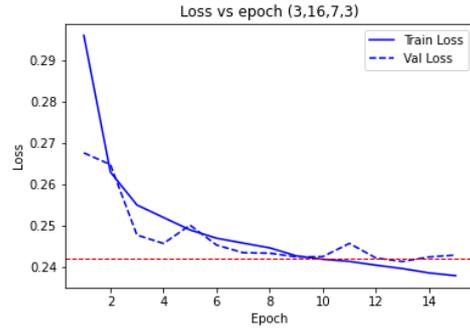


Figura 4.4: Representación de la evolución de *train loss* y *val loss* con los *epochs*. Modelo con parámetros (3,16,7,3). Se representa en rojo el valor de la función coste en el que se estabiliza el modelo.

Por último, es interesante estudiar las características de los exoplanetas que acierta la red y las de los que no consigue detectar. Los dos parámetros más influyentes dentro de los estudiados, son el periodo orbital y el valor del cociente rp , por lo que son los que se analizan.

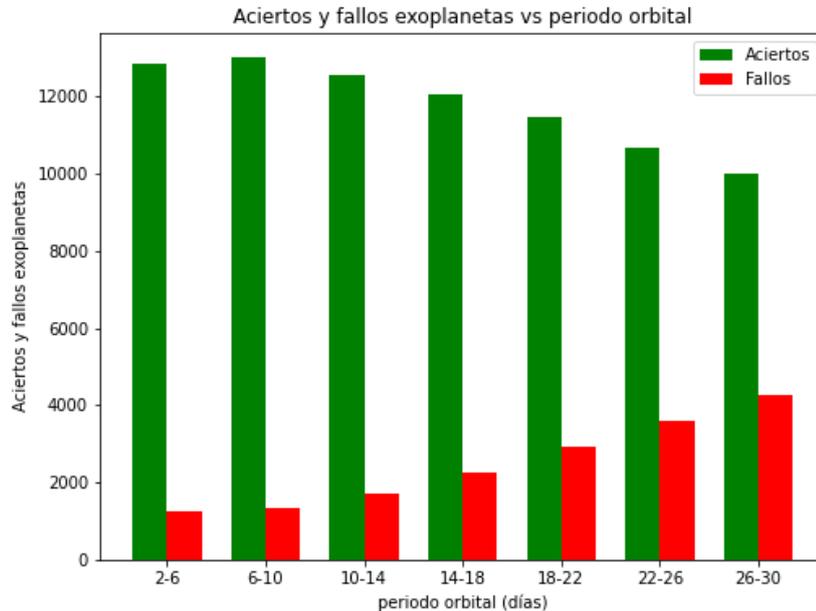


Figura 4.5: Representación de los aciertos y fallos en la detección de exoplanetas con el modelo (3,16,7,3), en función del periodo orbital.

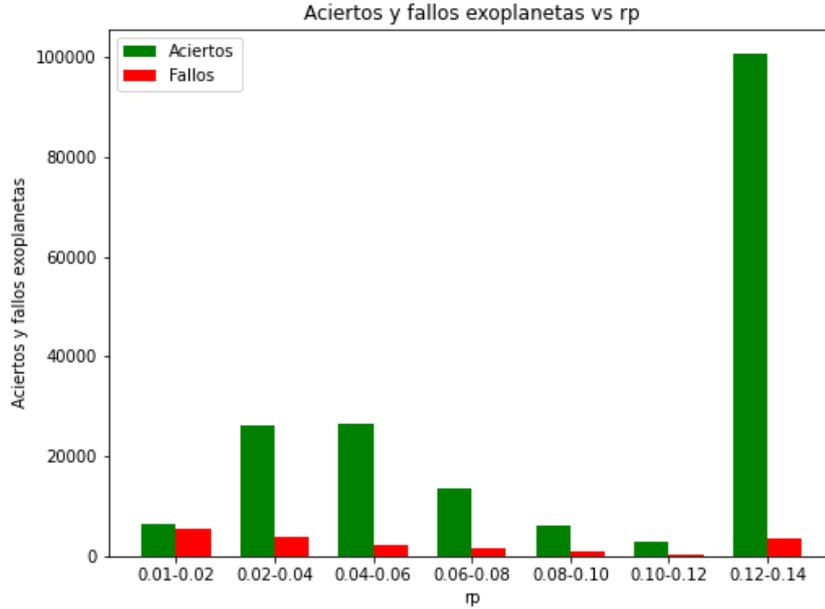


Figura 4.6: Representación de los aciertos y fallos en la detección de exoplanetas con el modelo (3,16,7,3), en función del cociente rp .

Al observar la figura 4.5, se comprueba que la creación de curvas de luz ha sido con un periodo orbital de los exoplanetas completamente aleatorio. Además, se confirma el buen funcionamiento de la red neuronal, puesto que la gran mayoría de clasificaciones son correctas. Para periodos bajos, se tiene un mayor número de tránsitos en la curva de luz que estudia la red, por lo que es más fácil detectar la presencia de un exoplaneta. Sin embargo, a medida que el periodo orbital aumenta, la dificultad va aumentando también. Al llegar a un periodo de 26-30 días, se observa un peor resultado en la clasificación; esto es lógico, puesto que se recuerda que las curvas de luz que se emplean en este estudio son de 30 días, por lo que en estos casos habrá un único tránsito. Aún así, en estas situaciones tan extremas, la red sigue teniendo un rendimiento satisfactorio.

En el caso de los aciertos y fallos con respecto al rp (ver Figura 4.6), se debe hacer un estudio más profundo. A la hora de realizar esta representación, se obtienen unos resultados que claramente muestran que los sistemas estrella-exoplaneta no han sido creados con un rp aleatorio. La idea inicial del trabajo era que este parámetro estuviera comprendido entre 0.01 y 0.1; sin embargo, un error en la programación llevó a que estos valores fueran divididos entre el radio de la estrella en unidades solares

(que se recuerda que estaba definido en un intervalo 0.3-2). Así, en muchos casos se ha obtenido un rp mayor que lo deseado. A pesar de no tener la misma fracción de planetas para cada tamaño, se observa en este caso también que el rendimiento de la red es muy bueno para todos ellos. El único intervalo que no resulta con un número de aciertos tan elevado (con respecto a los fallos) es el menor, es decir, 0.01-0.02, lo que es lógico, ya que un planeta con estas características provoca un tránsito mucho más pequeño en la curva de luz.

Capítulo 5

Conclusión

En este trabajo se han estudiado dos campos que están en auge en la actualidad: los exoplanetas y las redes neuronales; y se han aplicado estas últimas a la búsqueda e identificación de sistemas planetarios con diferentes características.

El ser humano, aunque haya ido evolucionando la perspectiva de estudio del Universo, siempre se ha preguntado por la existencia de otros planetas. Conocer la disposición de los sistemas planetarios en el Universo, así como la posibilidad de que algún exoplaneta sea habitable, son campos de estudio permanentes.

Existen diferentes métodos de detección de planetas fuera del Sistema Solar, tales como: el método de detección por tránsitos, el método de velocidad radial, la astrometría, la microlente gravitacional, la toma de imágenes directas, y las variaciones en el tiempo de diferentes fenómenos.

Desde el año 2009, en el que se lanzó el telescopio Kepler, el número de planetas que se ha identificado ha ido aumentando cada vez de forma más rápida, habiendo hoy en día un total de 5054 exoplanetas confirmados. Tanto esta misión como su sucesora, la misión TESS, se basan en el método de tránsito para su detección. Este hecho, y la gran cantidad de características que se pueden conocer a partir de dicho método (con ayuda también de la velocidad radial) hace que sea el más destacado.

Así, se ha desarrollado un programa que permite la creación de curvas de luz con y sin tránsitos. Las características de las estrellas y los planetas que las orbitan han sido descritas basándose en las regiones estudiadas por Kepler y TESS, según se indicó en el apartado de *Metodología*.

El 80 % de los datos creados son empleados en el entrenamiento de la red neuronal, y el 20 % restante forma el conjunto de datos necesarios para la evaluación del rendimiento del modelo final. Se comparan un total de 16 redes neuronales diferentes, definidas con todas las combinaciones posibles de los siguientes parámetros:

- Número de capas: 3 o 5.
- Tamaño del filtro: 7 u 11.
- Número de *strides*: 1 o 3.
- Capas convolucionales: 16 o 64.

La gran mayoría de las redes, una vez entrenadas, presentan un rendimiento muy satisfactorio, aproximadamente del 90 % de aciertos en todos los casos. Teniendo en cuenta que las curvas de luz empleadas han sido lo más realistas posibles, y los buenos resultados obtenidos, se concluye que se podría seguir investigando por este camino de cara a desarrollar más la red neuronal.

Cabe destacar también que el hecho de haber realizado un análisis de los resultados dependiendo de las características de los exoplanetas creados, ha ayudado a detectar al final un error en el programa (el parámetro *rp* no ha recorrido todos los valores deseados en un principio). Esto lleva a resaltar la necesidad e importancia de los análisis de datos para identificar este tipo de problemas que, de otra forma, pasarían desapercibidos.

Bibliografía

- [1] Manel Montes y Jorge Munnshe. *Las primeras sondas*. National Geographic, 2020.
- [2] Rafael Bachiller. 1995. *El 'boom' de los planetas extrasolares*. URL: <https://www.elmundo.es/elmundo/2009/11/17/ciencia/1258473431.html> (visitado 03-03-2022).
- [3] NASA. *Historic Timeline. Explore-Exoplanet Exploration*. URL: <https://exoplanets.nasa.gov/alien-worlds/historic-timeline/> (visitado 04-03-2022).
- [4] NASA *Exoplanet Archive*. URL: https://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html (visitado 11-07-2022).
- [5] Ethem Alpaydin. *Introduction to Machine Learning*. 3.^a ed. Adaptive Computation and Machine Learning. Cambridge, MA: MIT Press, 2014.
- [6] Donald Michie y col., eds. *Machine Learning, Neural and Statistical Classification*. USA: Ellis Horwood, 1995.
- [7] Peter Harrington. *Machine Learning in Action*. USA: Manning Publications Co., 2012. ISBN: 1617290181.
- [8] Enrique Castaños. *El Universo: del geocentrismo al heliocentrismo*. URL: <https://cienciadelux.com/2015/10/20/el-universo-del-geocentrismo-al-heliocentrismo/> (visitado 03-03-2022).
- [9] São Paulo. *Quién fue Giordano Bruno, el místico "visionario" quemado en la hoguera hace 418 años*. URL: <https://www.bbc.com/mundo/noticias-43097025> (visitado 03-03-2022).
- [10] John Wenz. *How the first exoplanets were discovered*. URL: <https://astronomy.com/news/2019/10/how-the-first-exoplanets-were-discovered#:~:text=In%5C%201992%5C%2C%5C%20astronomers%5C%20discovered%5C%20the,form%5C%20they'd%5C%20really%5C%20anticipated.&text=The%5C%20first%5C%20exoplanets%5C%20ever%5C%20discovered,exoplanets%5C%20around%5C%20sun%5C%2Dlike%5C%20stars.> (visitado 03-03-2022).
- [11] Donald Goldsmith. *Exoplanets: hidden worlds and the quest for extraterrestrial life*. Harvard University Press, 2018.
- [12] Daniel Marín. *El misterio de los júpiteres calientes o qué pasaría si los planetas gigantes pudieran formarse cerca de su estrella*. URL: <https://danielmarin.naukas.com/2015/12/02/el-misterio-de-los-jupiteres-calientes/> (visitado 15-06-2022).

- [13] Jordi Pereyra. *Los exoplanetas*. National Geographic, 2018.
- [14] Erik Gregersen. *MOST. Canadian orbiting telescope*. URL: <https://www.britannica.com/topic/MOST-telescope> (visitado 23-05-2022).
- [15] Malcolm Fridlund. *El satélite CoRoT: búsqueda de planetas parecidos a la Tierra*. URL: <https://www.scienceinschool.org/es/article/2010/corot-es/> (visitado 23-05-2022).
- [16] NASA Exoplanet archive. *Exoplanet plots. Pre-generated Exoplanet Plots*. URL: <https://exoplanetarchive.ipac.caltech.edu/exoplanetplots/> (visitado 12-06-2022).
- [17] NASA/JPL/Caltech. *Technology Overview*. URL: <https://exoplanets.nasa.gov/exep/technology/technology-overview/> (visitado 12-06-2022).
- [18] *Telescopio Kepler: Descubrimientos, funcionamiento y más*. URL: <https://misistemasolar.com/telescopio-kepler/> (visitado 23-04-2022).
- [19] Steve B. Howell y col. “The K2 Mission: Characterization and Early Results”. *Publications of the Astronomical Society of the Pacific* 126.938 (abr. de 2014), págs. 398-408. DOI: 10.1086/676406. URL: <https://doi.org/10.1086%5C%2F676406>.
- [20] Michele Johnson. *Kepler and K2. Overview*. URL: https://www.nasa.gov/mission_pages/kepler/overview/index.html (visitado 23-04-2022).
- [21] Daniel Marín. *Misión Kepler K2, prolongando la vida del buscador de exoplanetas*. URL: <https://danielmarin.naukas.com/2014/02/25/mision-kepler-k2-prolongando-la-vida-del-buscador-de-exoplanetas/> (visitado 30-04-2022).
- [22] Daniel Marín. *Los planetas descubiertos por la misión Kepler K2*. URL: <https://danielmarin.naukas.com/2015/11/30/los-planetas-descubiertos-por-la-mision-kepler-k2/> (visitado 30-04-2022).
- [23] *K2 Campaign Fields*. URL: <https://archive.stsci.edu/missions-and-data/k2/campaign-fields> (visitado 23-04-2022).
- [24] Thomas Barclay. *Mission objectives*. URL: <https://heasarc.gsfc.nasa.gov/docs/tess/objectives.html> (visitado 09-07-2022).
- [25] Emily Sandford y col. *Then Planetary Society. Timing Variations*. URL: <https://www.planetary.org/articles/timing-variations> (visitado 05-04-2022).
- [26] Alex Riveiro. *La resonancia orbital*. URL: <https://www.astrobitacora.com/la-resonancia-orbital/> (visitado 06-06-2022).
- [27] George R. Ricker y col. “Transiting Exoplanet Survey Satellite”. *Journal of Astronomical Telescopes, Instruments, and Systems* 1.1 (oct. de 2014), pág. 014003. DOI: 10.1117/1.jatis.1.1.014003. URL: <https://doi.org/10.1117%5C%2F1.jatis.1.1.014003>.
- [28] Patrick Blau. *TESS Orbit Design*. URL: <https://spaceflight101.com/tess/tess-orbit-design/> (visitado 12-06-2022).

- [29] Aomawa Shields y col. “The Habitability of Planets Orbiting M-dwarf Stars”. *Physics Reports* 663 (oct. de 2016). DOI: 10.1016/j.physrep.2016.10.003.
- [30] Carles Paul. *¿Solos ante la Galáxia?* URL: <https://abcienciade.wordpress.com/2019/02/14/solos-ante-la-galaxia/> (visitado 27-04-2022).
- [31] Pat Brennan. *Gas Giant*. URL: <https://exoplanets.nasa.gov/what-is-an-exoplanet/planet-types/gas-giant/> (visitado 02-07-2022).
- [32] Pat Brennan. *Neptune-like*. URL: <https://exoplanets.nasa.gov/what-is-an-exoplanet/planet-types/neptune-like/> (visitado 02-07-2022).
- [33] Pat Brennan. *Super-Earth*. URL: https://exoplanets.nasa.gov/what-is-an-exoplanet/planet-types/super-earth/#otp_what_is_a_super-earth? (visitado 02-07-2022).
- [34] Pat Brennan. *Earth and Super-Earth*. URL: <https://exoplanets.nasa.gov/resources/15/earth-and-super-earth/> (visitado 02-07-2022).
- [35] Pat Brennan. *Terrestrial*. URL: <https://exoplanets.nasa.gov/what-is-an-exoplanet/planet-types/terrestrial/> (visitado 02-07-2022).
- [36] Rick Chen. *Small Planets Come in Two Sizes*. URL: <https://www.nasa.gov/image-feature/ames/small-planets-come-in-two-sizes> (visitado 17-06-2022).
- [37] Pat Brennan. *Terrestrial*. URL: https://exoplanets.nasa.gov/what-is-an-exoplanet/planet-types/terrestrial/#otp_a_strange_gap_in_planet_sizes (visitado 17-06-2022).
- [38] Marie-Eve Naud. *Twice the size of Earth*. URL: <http://www.exoplanetes.umontreal.ca/twice-the-size-of-earth/?lang=en> (visitado 17-06-2022).
- [39] Dan Maoz. *Astrophysics in a nutshell*. Princeton University Press, 2007.
- [40] JEAN TATE. *UNIVERSE TODAY. Space and astronomy news*. URL: <https://www.universetoday.com/52252/main-sequence/> (visitado 03-07-2022).
- [41] Lysiane Patricia Karine Mornas. *Apuntes de la asignatura Astrofísica y Cosmología curso 20-21*.
- [42] Wikipedia. *Nodos de la órbita*. URL: https://es.wikipedia.org/wiki/Nodos_de_la_%5C%C3%5C%B3rbita (visitado 04-04-2022).
- [43] Vishal Upendran. *Weekly literature review 1: Exoplanet detection using deep learninga*. URL: <https://medium.com/@uvishal1995/weekly-literature-review-1-exoplanet-detection-using-deep-learning-4930ee4f8733> (visitado 04-04-2022).
- [44] Pat Brennan. *NASA Exoplanet Exploration. Planets Beyond Our Solar System. What’s a transit?* URL: <https://exoplanets.nasa.gov/faq/31/whats-a-transit/> (visitado 05-04-2022).
- [45] Pat Brennan. *NASA Exoplanet Exploration. Planets Beyond Our Solar System. Transit*. URL: <https://exoplanets.nasa.gov/alien-worlds/ways-to-find-a-planet/#/2> (visitado 05-04-2022).
- [46] Renée Dotson y col. *Exoplanets*. University of Arizona Press, 2011.

- [47] *Down in Front!: The Transit Photometry Method*. URL: <https://www.planetary.org/articles/down-in-front-the-transit-photometry-method> (visitado 05-04-2022).
- [48] John W. Mason. *Exoplanets: Detection, Formation, Properties, Habitability (Springer-Praxis Books in Astronomy and Planetary Sciences)*. 1.^a ed. Springer-Praxis Books in Astronomy and Planetary Sciences. Springer, 2008.
- [49] *Reading an Earth-like Exoplanet's Transmission Spectrum*. URL: <https://webbtelescope.org/contents/media/images/01FEE26XVSM851DHPVCE1KB4S2?Tag=Spectroscopy> (visitado 05-04-2022).
- [50] Teresa Martín Blas y col. *Dinámica*. URL: <https://www2.montes.upm.es/dptos/digfa/cfisica/dinamsist/tierraluna.html> (visitado 04-04-2022).
- [51] SEA. *Velocidad radial*. URL: <https://www.sea-astronomia.es/glosario/velocidad-radial> (visitado 20-04-2022).
- [52] Oliver Toogood. *Mr Toogood's Physics. A Level Physics notes to support my lessons at LCS*. URL: http://www.alevelphysicsnotes.com/astrophysics/doppler_effect.php (visitado 20-05-2022).
- [53] Dr. Sascha Quanz. *Extrasolar Planets / FS 2019*. URL: https://quanz-group.ethz.ch/education/lectures/2019_courses/extrasolar-planets---fs2019.html (visitado 17-04-2022).
- [54] Jason T Wright y col. "Exoplanet detection methods". *arXiv:1210.2471* (2012).
- [55] *Astrometría-Diccionario de Astronomía*. URL: <https://www.astromia.com/glosario/astrometria.htm> (visitado 05-04-2022).
- [56] Pat Brennan. *NASA Exoplanet Exploration. Planets Beyond Our Solar System. Astrometry*. URL: <https://exoplanets.nasa.gov/alien-worlds/ways-to-find-a-planet/#/5> (visitado 05-04-2022).
- [57] ESA. *How to find an exoplanet*. URL: https://www.esa.int/Science_Exploration/Space_Science/Cheops/How_to_find_an_exoplanet (visitado 05-04-2022).
- [58] *La misión Gaia*. URL: <https://gaiaverse.eu/la-mision-gaia/> (visitado 20-05-2022).
- [59] Pat Brennan. *NASA Exoplanet Exploration. Planets Beyond Our Solar System*. URL: <https://exoplanets.nasa.gov/alien-worlds/ways-to-find-a-planet/#/4> (visitado 05-04-2022).
- [60] Yiannis Tsapras. "Microlensing Searches for Exoplanets". *Geosciences* 8.10 (sep. de 2018), pág. 365. DOI: 10.3390/geosciences8100365. URL: <https://doi.org/10.3390%5C%2Fgeosciences8100365>.
- [61] Claus Grupen y col. *Astroparticle Physics*. 1.^a ed. Springer, 2005.
- [62] Stanton Peale. *Microlensing can reveal exoplanets*. URL: https://www.researchgate.net/figure/Microlensing-can-reveal-exoplanets-By-gravitational-bending-of-light-a-foreground_fig1_243391943 (visitado 05-05-2022).

- [63] Pat Brennan. *NASA Exoplanet Exploration. Planets Beyond Our Solar System*. URL: <https://exoplanets.nasa.gov/alien-worlds/ways-to-find-a-planet/#/3> (visitado 05-04-2022).
- [64] Matt Williams. *Universe Today. Space and astronomy news*. URL: <https://www.universetoday.com/140341/what-is-direct-imaging/> (visitado 05-04-2022).
- [65] Marc Kaufman. *A four-planet system in orbit, directly imaged and remarkable*. URL: <https://exoplanets.nasa.gov/news/1404/a-four-planet-system-in-orbit-directly-imaged-and-remarkable/> (visitado 08-05-2022).
- [66] *Sistemas eclipsantes*. URL: <https://www.astromia.com> (visitado 10-04-2022).
- [67] M.Gerbaldi y col. *Binarias eclipsantes o fotométricas*. URL: <https://media4.obspm.fr/public/VAU/masa/binario/eclipses/INTRODUCTION/index.html> (visitado 10-04-2022).
- [68] Antxon Alberdi. *Púlsares: los faros para navegantes cósmicos*. URL: <http://www-revista.iaa.es/36/p%C3%BAlsares%E2%80%81los-faros-para-navegantes-c%C3%B3smicos> (visitado 15-04-2022).
- [69] *Púlsares*. URL: <https://www.astromia.com/universo/pulsares.htm> (visitado 17-04-2022).
- [70] Las Cumbres Observatory. *Pulsar Timing Method*. URL: <https://lco.global/spacebook/exoplanets/pulsar-timing-method/#:~:text=Known%20pulsars%20have%20rates%20ranging,with%20one%20or%5C%20more%20planets.> (visitado 17-04-2022).
- [71] *Pulsating Variable Stars*. URL: https://www.atnf.csiro.au/outreach/education/senior/astrophysics/variable_pulsating.html (visitado 17-04-2022).
- [72] Bence Béky. *Development and Application of Tools to Characterize Transiting Astrophysical Systems*. 2014. DOI: 10.48550/ARXIV.1408.0401. URL: <https://arxiv.org/abs/1408.0401>.
- [73] Joseph Stromberg. *How Do Astronomers Actually Find Exoplanets?* URL: <https://www.smithsonianmag.com/science-nature/how-do-astronomers-actually-find-exoplanets-180950105/> (visitado 05-04-2022).
- [74] Anglada y col. "Protoplanetary disks, jets, and the birth of the stars". *Spanish SKA White Book* (2015).
- [75] *Planet-Forming Disks*. URL: <https://almascience.eso.org/alma-science/planet-forming-disks> (visitado 20-04-2022).
- [76] ALMA. *En busca de nuestros orígenes cósmicos*. URL: <https://www.eso.org/public/spain/teles-instr/alma/> (visitado 10-07-2022).
- [77] Chris y col. *Deep Learning Fundamentals - Classic Edition*. URL: https://deeplizard.com/learn/video/YRhxvVk_sIs (visitado 20-06-2022).
- [78] *Hombre investigando con lupa*. URL: https://www.freepik.es/foto-gratis/hombre-investigando-lupa_918620.htm (visitado 30-05-2022).

- [79] Josh Starmer. *Ridge vs Lasso Regression, Visualized!!!* URL: https://www.youtube.com/watch?v=Xm2C_gTA18c (visitado 10-04-2022).
- [80] Ritesh Bhagwat y col. *Applied Deep Learning with Keras: Solve complex real-life problems with the simplicity of Keras*. Packt Publishing Limited, 2019.
- [81] *Las cuernas de los ciervos pueden frenar cáncer*. URL: <https://www.club-caza.com/actualidad/actualverold2.asp?nn=9096> (visitado 30-05-2022).
- [82] *Golondrina común*. URL: <https://aves.animalesbiologia.com/paseriformes/golondrina-comun-hirundo-rustica> (visitado 30-05-2022).
- [83] *La Vanguardia*. URL: <https://www.lavanguardia.com/vivo/mascotas/20190404/461425413181/comportamiento-animal-gatos-reconocen-nombre-no-obedecen.html> (visitado 30-05-2022).
- [84] Shiva Verma. *Understanding 1D and 3D Convolution Neural Network — Keras*. URL: <https://towardsdatascience.com/understanding-1d-and-3d-convolution-neural-network-keras-9d8f76e29610> (visitado 01-07-2022).
- [85] Laura Kreidberg. *Quickstart-batman 2.4.6 documentation*. URL: <https://lkreidberg.github.io/batman/docs/html/quickstart.html> (visitado 07-07-2022).
- [86] Wikipedia. *Leyes de Kepler*. URL: https://es.wikipedia.org/wiki/Leyes_de_Kepler (visitado 04-04-2022).
- [87] David G. Koch y col. “Kepler Mission Design, Realized Photometric Performance, and Early Science”. *The Astrophysical Journal* 713.2 (mar. de 2010), págs. L79-L86. DOI: 10.1088/2041-8205/713/2/179. URL: <https://doi.org/10.1088/5C%2F2041-8205%5C%2F713%5C%2F2%5C%2F179>.
- [88] J.P. Beaulieu y col. “Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing”. *Nature* 439.7075 (ene. de 2006), págs. 437-440. DOI: 10.1038/nature04441. URL: <https://doi.org/10.1038/5C%2Fnature04441>.
- [89] Pat Brennan. *NASA Exoplanet Exploration. Planets Beyond Our Solar System*. URL: <https://exoplanets.nasa.gov/alien-worlds/ways-to-find-a-planet/#/1> (visitado 05-04-2022).
- [90] *Space-Warping Planets: The Microlensing Method*. URL: <https://www.planetary.org/articles/space-warping-planets-the-microlensing-method> (visitado 06-04-2022).
- [91] *Space-Warping Planets: The Microlensing Method*. URL: <https://roman.gsfc.nasa.gov/exoplanets-microlensing.html> (visitado 05-04-2022).
- [92] Walt Feimer. *NASA Conceptual Image Lab. Gravitational Microlensing Animation*. URL: <https://svs.gsfc.nasa.gov/20242> (visitado 05-04-2022).
- [93] Adam Rogers. *General Relativity and Gravitational Lensing: Part 3-Microlensing*. URL: <https://theamateurrealist.wordpress.com/2014/09/15/general-relativity-and-gravitational-lensing-part-3-microlensing/> (visitado 20-06-2022).

- [94] Michele Johnson. *Mission Manager Update: Kepler's K2 Mission Begins Campaign 13*. URL: <https://blogs.nasa.gov/kepler/2017/03/23/keplers-k2-mission-is-now-in-its-thirteenth-observing-campaign-monitoring-a-patch-of-sky-in-the-direction-of-the-constellation-taurus-in-addition-to-observing-21000-target-stars-measureme/> (visitado 23-04-2022).

Apéndice A: Código creación curvas de luz con y sin tránsito

```
# -*- coding: utf-8 -*-
"""
Creacion definitiva datos

LIDIA SAINZ LEDO
"""

#M dulos necesarios
import batman

import numpy as np
import matplotlib.pyplot as plt
import h5py
import pandas as pd
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
from sklearn.utils import shuffle
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import curve_fit
from sklearn.metrics import confusion_matrix
import seaborn as sn

import zipfile

"""
En primer lugar se va a realizar una regresion lineal a partir de datos obtenidos
del articulo Kepler Mission Design, Realized Photometric Performance, and Early Science.
Asi se sabra la relacion entre el ruido y la magnitud aparente de la estrella.
"""
#Se crean los arrays para saber el ruido que se corresponde para cada magnitud:
x_m=[7,8,9,10,11,12,13]
y_m=[5.4, 8.8, 14, 22, 34, 54,90]
plt.figure()
plt.plot(x_m,y_m)
plt.semilogy()
plt.show()

#Se realiza la regresion lineal:
def func(x,a,b):
    return a**x*b
popt, pcov = curve_fit(func, x_m, y_m)

func=np.vectorize(func)
ruido_m=func([7,8,9,10,11,12,13,14,15,16,17],popt[0], popt[1])
plt.figure()
plt.plot([7,8,9,10,11,12,13,14,15,16,17],ruido_m)
plt.semilogy()
plt.show()

#Vector del que se va a obtener aleatoriamente la magnitud aparente de la estrella.
array_magnitud=np.arange(11,17.1,0.1)
#Ruido para cada valor de la magnitud aparente del array anterior.
valores_ruido_m=func(array_magnitud, popt[0], popt[1])

"""Se fija el numero de datos de cada tipo que se va a crear:"""

ntrain = 800000 #Variable que define el numero de curvas de entrenamiento que se crean (80% de los datos)
ntest = 200000 #Variable que define el numero de curvas de test que se crean (20% de los datos)

"""Se fijan los parametros del array del tiempo:"""

t_ini = 0 #Variable que define el tiempo inicial de las curvas que se crean, en dias
t_fin = 30 #Variable que define el tiempo final de las curvas que se crean, en dias
num_puntos = 1440 #Numero de puntos en la curva de luz.
#Se elige ese numero por ser el tipico de TESS con cadencia de 30 minutos.

"""Se crean las listas donde se guardaran todos los datos de test y train creados:"""

parametros_train = [] #Aqui se guardaran los parametros de los sistemas creados
```

```

train_samples = [] #Aqui se guardaran los flujos creados
train_labels = [] #Aqui se guardara si hay transito 1, o no, 0

parametros_test = [] #Aqui se guardaran los par metros de los sistemas creados
test_samples = [] #Aqui se guardaran los flujos creados
test_labels = [] #Aqui se guardara si hay transito 1, o no, 0

"""Se crean las curvas de luz artificiales (CON TRANSITO):"""

G=6.7*10**(-8) #erg cm g^-2
Solar_radius=7.0*10**10 #cm
Solar_mass = 2.0*10**33 #g
for exo in range(int(round(ntrain/2.))+round(ntest/2.)):
    radius_star=np.random.uniform(0.3, 2) #radio de la estrella en Radios solares
    mass_star=(radius_star)**(1./0.75) #masa de la estrella en Masas solares
    orbital_period=np.random.uniform(2.,30.) #periodo orbital en dias

    time_conjunction=np.random.uniform(0,orbital_period-1) #tiempo de conjunci n
    #El semieje mayor en unidades solares
    semi_axis=((G*(mass_star*Solar_mass)*(orbital_period*86400)**2.)/(4*np.pi**2.))**(1./3)/Solar_radius
    orbital_inclination=np.random.uniform(88.,92.) #inclinacion orbital en grados
    eccentricity=np.random.uniform(0.,0.1) #excentricidad
    magnitud_aparente=float(np.random.choice(array_magnitud,1)) #magnitud aparente
    indice_magnitud=np.where(array_magnitud==magnitud_aparente)
    ruido_magnitud=float(valores_ruido_m[indice_magnitud]) #ruido de esa magnitud aparente
    if radius_star >=0.3 and radius_star <=0.7:
        interval=np.random.uniform(0.01,0.04)
        rp_cociente=interval/radius_star
    elif radius_star >0.7:
        interval=np.random.uniform(0.01,0.1)
        rp_cociente=interval/radius_star

    #Probamos modelo
    params = batman.TransitParams() #objeto para almacenar marametros de transito
    params.t0 = time_conjunction #tiempo inferior de conjuncion
    params.per = orbital_period #periodo orbital
    params.rp = rp_cociente #radio del planeta en unidades del radio solar
    params.a = semi_axis/radius_star #semi eje mayor en unidades del radio solar
    params.inc = orbital_inclination #inclinacion orbital en grados
    params.ecc = eccentricity #excentricidad
    params.w = 90. #longitud del periastro en grados
    params.limb_dark = "nonlinear" #limb darkening (modelo)
    params.u = [0.5, 0.1, 0.1, -0.1] #limb darkening (coeficientes) [u1, u2, u3, u4]

    t=np.linspace(t_ini, t_fin, num_puntos)#tempo en el que calcular la curva de luz
    m = batman.TransitModel(params, t, fac=1e-6) #se inicializa el modelo

    flux = m.light_curve(params) #se calcula la curva de luz
    #Guardo los parametros de cada sistema creado
    #Orden:
    #flujo, masa estrella, radio estrella, rp_cociente, semieje, inclinacion orbital, eccentricity,
    #periodo orbital, time_conjunction, magn. aparente, sigma
    #cuidado con las unidades de sigma
    if exo>=int(round(ntrain/2.)):
        parametros_test.append([mass_star, radius_star, rp_cociente, semi_axis, orbital_inclination, eccentricity,
                                orbital_period, time_conjunction, magnitud_aparente, ruido_magnitud])
    else:
        parametros_train.append([mass_star, radius_star, rp_cociente, semi_axis, orbital_inclination,
                                eccentricity, orbital_period, time_conjunction, magnitud_aparente, ruido_magnitud])

    ruido_flux=np.random.normal(0, ruido_magnitud*10**(-6.), num_puntos)
    flux_total=flux+ruido_flux

    if exo>=int(round(ntrain/2.)):
        test_samples.append(flux_total)
        test_labels.append(1)
    else:
        train_samples.append(flux_total)
        train_labels.append(1)

"""Se crean las curvas de luz artificiales (SIN TRANSITO):"""

for exo in range(int(round(ntrain/2.))+round(ntest/2.)):
    magnitud_aparente=float(np.random.choice(array_magnitud,1))
    indice_magnitud=np.where(array_magnitud==magnitud_aparente)
    ruido_magnitud=float(valores_ruido_m[indice_magnitud])
    #Orden: magn. aparente, sigma
    #(cuidado que sigma no esta en las unidades correctas)!!!
    if exo>=int(round(ntrain/2.)):
        parametros_test.append([0,0,0,0,0,0,0, magnitud_aparente, ruido_magnitud])
    else:
        parametros_train.append([0,0,0,0,0,0,0, magnitud_aparente, ruido_magnitud])

    flux_total=np.random.normal(1, ruido_magnitud*10**(-6.), num_puntos)
    t=np.linspace(t_ini, t_fin, num_puntos)#times at which to calculate light curve

    if exo>=int(round(ntrain/2.)):
        test_samples.append(flux_total)
        test_labels.append(0)
    else:
        train_samples.append(flux_total)
        train_labels.append(0)

```

```

"""Me aseguro de que todo est guardado en arrays y con la forma deseada:
"""

train_samples = np.array(train_samples)
train_labels = np.array(train_labels)
parametros_train=np.array(parametros_train)
parametros_train = np.around(parametros_train,4)
test_samples = np.array(test_samples)
test_labels = np.array(test_labels)
parametros_test=np.array(parametros_test)
parametros_test = np.round(parametros_test,4)
print(parametros_train)
print(np.shape(train_samples))
print(np.shape(train_labels))
print(np.shape(parametros_train))
print(np.shape(test_samples))
print(np.shape(test_labels))
print(np.shape(parametros_test))

"""Mezclo los exoplanetas y los no exoplanetas para no introducir los datos ordenados en la red:"""

parametros_train_shuffle ,train_samples_shuffle ,train_labels_shuffle = shuffle(parametros_train ,
train_samples ,train_labels , random_state=0)
parametros_test_shuffle ,test_samples_shuffle ,test_labels_shuffle = shuffle(parametros_test ,
test_samples ,test_labels , random_state=0)

print(np.shape(parametros_train))
print(np.shape(parametros_train_shuffle))
print(np.shape(parametros_test_shuffle))

"""Guardo los datos en formato h5:"""

#Guardo los train_samples , train_labels y parametros_train en un fichero h5
f1 = h5py.File("train_data.hdf5" , "w")
dset1 = f1.create_dataset("Parametros_train" , (ntrain,10) , dtype='float32' , data=parametros_train_shuffle)
dset2 = f1.create_dataset("Train_samples" , (ntrain,num_puntos) , dtype='float32' , data=train_samples_shuffle)
dset3 = f1.create_dataset("Train_labels" , (ntrain,1) , dtype='i' , data=train_labels_shuffle)
f1.close()

#Guardo los test_samples , test_labels y parametros_test en un fichero h5
f2 = h5py.File("test_data.hdf5" , "w")
dset4 = f2.create_dataset("Parametros_test" , (ntest,10) , dtype='float32' , data=parametros_test_shuffle)
dset5 = f2.create_dataset("Test_samples" , (ntest,num_puntos) , dtype='float32' , data=test_samples_shuffle)
dset6 = f2.create_dataset("Test_labels" , (ntest,1) , dtype='i' , data=test_labels_shuffle)
f2.close()

"""Codigo para acceder a los datos ya guardados en drive:
from google.colab import drive
drive.mount('/content/drive')

f3 = h5py.File('/content/drive/MyDrive/Ficheros/train_data.hdf5' , 'r')

list(f3.keys())

dset = f3['Train_samples']
train_samples = dset[:]

dset = f3['Train_labels']
train_labels = dset[:]

dset = f3['Parametros_train']
parametros_train = dset[:]

f4 = h5py.File('/content/drive/MyDrive/Ficheros/test_data.hdf5' , 'r')

list(f4.keys())

dset = f4['Test_samples']
test_samples = dset[:]

dset = f4['Test_labels']
test_labels = dset[:]

dset = f4['Parametros_test']
parametros_test = dset[:]
"""

"""Me aseguro de que las curvas de entrenamiento se est n creando bien:"""

t=np.linspace(t_ini ,t_fin ,num_puntos)
for i in range(0,20):
    if train_labels[i]==1:
        print("Tr nsito")
        print(train_labels[i])
    else:
        print("No tr nsito")
        print(train_labels[i])
    plt.figure(figsize=(20,4))
    #.xlim(,)
    plt.plot(t ,train_samples[i])
    plt.show()

"""Me aseguro de que las curvas de test se est n creando bien:"""
t=np.linspace(t_ini ,t_fin ,num_puntos)
for i in range(0,20):

```

```
if test_labels[i]==1:
    print("Tr nsito")
    print(test_labels[i])
else:
    print("No tr nsito")
    print(test_labels[i])
plt.figure(figsize=(20,4))
# .xlim(,)
plt.plot(t, test_samples[i])
plt.show()
```

Apéndice B: Código definición y entrenamiento red neuronal 3 capas convolucionales

```
# -*- coding: utf-8 -*-
"""
Red convolucional 3 capas

LIDIA SAINZ LEDO
"""

import numpy as np
import matplotlib.pyplot as plt
import h5py
import pandas as pd
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
from keras.utils import np_utils
#from keras.callbacks import EarlyStopping

from sklearn.utils import shuffle
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import curve_fit
from sklearn.metrics import confusion_matrix
import seaborn as sns

"""Se fija el numero de datos de cada tipo que se va a crear:"""
ntrain = 800000 #Variable que define el numero de curvas de entrenamiento que se crean (80% de los datos)
ntest = 200000 #Variable que define el numero de curvas de test que se crean (20% de los datos)

"""Se fijan los parametros del array del tiempo:"""
t_ini = 0 #Variable que define el tiempo inicial de las curvas que se crean, en dias
t_fin = 30 #Variable que define el tiempo final de las curvas que se crean, en dias
num_puntos = 1440 #Numero de puntos en la curva de luz.
#Se elige ese numero por ser el t pico de TESS con cadencia de 30 minutos.

""" Se leen los datos ya creados y guardados"""
f1=h5py.File('train_data.hdf5','r')
f1.keys()
f2=h5py.File('test_data.hdf5','r')
f2.keys()

dset = f1['Train_samples']
train_samples = dset[:]

dset = f1['Train_labels']
train_labels = dset[:]

dset = f1['Parametros_train']
parametros_train = dset[:]

dset = f2['Test_samples']
test_samples = dset[:]

dset = f2['Test_labels']
test_labels = dset[:]

dset = f2['Parametros_test']
parametros_test = dset[:]

f1.close()
f2.close()

""" Normalizacion y reshape de los datos:"""

train_data = 1- (train_samples - np.amin(train_samples)) / (np.amax(train_samples)-np.amin(train_samples))
#Uso 1- para MaxPooling para que destaquen los transitos
train_labels = np_utils.to_categorical(train_labels, 2)
train_data=train_data.reshape(ntrain,num_puntos,1)
```

```

test_data = 1 - (test_samples - np.amin(test_samples)) / (np.amax(test_samples) - np.amin(test_samples))
#Uso 1- para MinPooling para que destaquen los transitos
test_labels = np_utils.to_categorical(test_labels, 2)
test_data=test_data.reshape(n_test, num_puntos,1)

"""RED CONVOLUCIONAL BUCLE:"""

number_layers_list=[3,5]
filter_size_list=[7,11]
number_strides_list=[1,3]
number_conv_layers_filters_list=[16,64]

"""Se definen las variables donde se guardara la informacion de cada modelo creado y entrenado """
histories=[]
test_loss_list=[]
test_accuracy_list=[]
val_loss_list=[]
val_accuracy_list=[]
train_loss_list=[]
train_accuracy_list=[]
matrix_list=[]
predictions_list=[]

for i in range(len(number_conv_layers_filters_list)):
    for j in range(len(filter_size_list)):
        for k in range(len(number_strides_list)):
            model=tf.keras.models.Sequential([
                tf.keras.layers.Conv1D(number_conv_layers_filters_list[i],(filter_size_list[j]),
                                       number_strides_list[k], activation='relu', input_shape=(num_puntos,1)),
                tf.keras.layers.Conv1D(number_conv_layers_filters_list[i],(filter_size_list[j]),
                                       number_strides_list[k], activation='relu'),
                tf.keras.layers.Conv1D(number_conv_layers_filters_list[i],(filter_size_list[j]),
                                       number_strides_list[k], activation='relu'),
                tf.keras.layers.Flatten(),
                tf.keras.layers.Dense(128, activation='relu'),
                tf.keras.layers.Dense(2, activation='softmax')
            ])
            model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
            model.summary()

            history = model.fit(train_data, train_labels, validation_split=0.1,epochs=15)
            test_loss, test_accuracy = model.evaluate(test_data, test_labels) #Evaluaci n (test) de la red
            prediction = model.predict(test_data)

            hist = pd.DataFrame(history.history)
            last_val_accuracy = hist['val_accuracy'].iloc[-1]
            last_train_accuracy = hist['accuracy'].iloc[-1]
            last_val_loss = hist['val_loss'].iloc[-1]
            last_train_loss = hist['loss'].iloc[-1]

            valor_prediction= np.argmax(prediction, axis=1)
            prediction_cm = confusion_matrix(np.argmax(test_labels, axis=1), valor_prediction)

            histories.append(hist)
            test_loss_list.append(test_loss)
            test_accuracy_list.append(test_accuracy)
            val_loss_list.append(last_val_loss)
            val_accuracy_list.append(last_val_accuracy)
            train_loss_list.append(last_train_loss)
            train_accuracy_list.append(last_train_accuracy)
            matrix_list.append(prediction_cm)
            predictions_list.append(valor_prediction)

""" Se guardan todos los datos en un archivo h5 """
f5 = h5py.File("red_bucle_1.hdf5", "w")
dset1 = f5.create_dataset("histories", (8,15,4), dtype='float32', data=histories)
dset2 = f5.create_dataset("test_loss", (8,), dtype='float32', data=test_loss_list)
dset3 = f5.create_dataset("test_accuracy", (8,), dtype='float32', data=test_accuracy_list)
dset4 = f5.create_dataset("val_loss", (8,), dtype='float32', data=val_loss_list)
dset5 = f5.create_dataset("val_accuracy", (8,), dtype='float32', data=val_accuracy_list)
dset6 = f5.create_dataset("train_loss", (8,), dtype='float32', data=train_loss_list)
dset7 = f5.create_dataset("train_accuracy", (8,), dtype='float32', data=train_accuracy_list)
dset8 = f5.create_dataset("matrix", (8,2,2), dtype='float32', data=matrix_list)
dset9 = f5.create_dataset("predictions", (8,n_test), dtype='float32', data=predictions_list)
f5.close()

```

Apéndice C: Código definición y entrenamiento red neuronal 5 capas convolucionales

```
# -*- coding: utf-8 -*-
"""
Red convolucional con 5 capas

LIDIA SAINZ LEDO
"""

import numpy as np
import matplotlib.pyplot as plt
import h5py
import pandas as pd
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
from keras.utils import np_utils
#from keras.callbacks import EarlyStopping

from sklearn.utils import shuffle
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import curve_fit
from sklearn.metrics import confusion_matrix
import seaborn as sns

"""Se fija el numero de datos de cada tipo que se va a crear:"""
ntrain = 800000 #Variable que define el numero de curvas de entrenamiento que se crean (80% de los datos)
ntest = 200000 #Variable que define el numero de curvas de test que se crean (20% de los datos)

"""Se fijan los parametros del array del tiempo:"""
t_ini = 0 #Variable que define el tiempo inicial de las curvas que se crean, en dias
t_fin = 30 #Variable que define el tiempo final de las curvas que se crean, en dias
num_puntos = 1440 #Numero de puntos en la curva de luz.
#Se elige ese numero por ser el tiempo pico de TESS con cadencia de 30 minutos.

""" Se leen los datos ya creados y guardados"""
f1=h5py.File('train_data.hdf5','r')
f1.keys()
f2=h5py.File('test_data.hdf5','r')
f2.keys()

dset = f1['Train_samples']
train_samples = dset[:]

dset = f1['Train_labels']
train_labels = dset[:]

dset = f1['Parametros_train']
parametros_train = dset[:]

dset = f2['Test_samples']
test_samples = dset[:]

dset = f2['Test_labels']
test_labels = dset[:]

dset = f2['Parametros_test']
parametros_test = dset[:]

f1.close()
f2.close()

"""Normalizacion y reshape de los datos:"""

train_data = 1-(train_samples - np.amin(train_samples)) / (np.amax(train_samples)-np.amin(train_samples))
#Uso 1- para MaxPooling para que destaquen los transitos
train_labels = np_utils.to_categorical(train_labels, 2)
train_data=train_data.reshape(ntrain, num_puntos, 1)
```

```

test_data = 1 - (test_samples - np.amin(test_samples)) / (np.amax(test_samples) - np.amin(test_samples))
#Uso 1- para MinPooling para que destaquen los tr nsitos
test_labels = np_utils.to_categorical(test_labels, 2)
test_data=test_data.reshape(n_test, num_puntos,1)

"""RED CONVOLUCIONAL BUCLE: """

number_layers_list=[3,5]
filter_size_list=[7,11]
number_strides_list=[1,3]
number_conv_layers_filters_list=[16,64]

"""Se definen las variables donde se guardara la informacion de cada modelo creado y entrenado """
histories=[]
test_loss_list=[]
test_accuracy_list=[]
val_loss_list=[]
val_accuracy_list=[]
train_loss_list=[]
train_accuracy_list=[]
matrix_list=[]
predictions_list=[]

for i in range(len(number_conv_layers_filters_list)):
    for j in range(len(filter_size_list)):
        for k in range(len(number_strides_list)):
            model=tf.keras.models.Sequential([
                tf.keras.layers.Conv1D(number_conv_layers_filters_list[i],(filter_size_list[j]),
                    number_strides_list[k], activation='relu', input_shape=(num_puntos,1)),
                tf.keras.layers.Conv1D(number_conv_layers_filters_list[i],(filter_size_list[j]),
                    number_strides_list[k], activation='relu'),
                tf.keras.layers.Conv1D(number_conv_layers_filters_list[i],(filter_size_list[j]),
                    number_strides_list[k], activation='relu'),
                tf.keras.layers.Conv1D(number_conv_layers_filters_list[i],(filter_size_list[j]),
                    number_strides_list[k], activation='relu'),
                tf.keras.layers.Conv1D(number_conv_layers_filters_list[i],(filter_size_list[j]),
                    number_strides_list[k], activation='relu'),
                tf.keras.layers.Flatten(),
                tf.keras.layers.Dense(128, activation='relu'),
                tf.keras.layers.Dense(2, activation='softmax')
            ])
            model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
            model.summary()

            history = model.fit(train_data, train_labels, validation_split=0.1,epochs=15)
            test_loss, test_accuracy = model.evaluate(test_data, test_labels) #Evaluacion (test) de la red
            prediction = model.predict(test_data)

            hist = pd.DataFrame(history.history)
            last_val_accuracy = hist['val_accuracy'].iloc[-1]
            last_train_accuracy = hist['accuracy'].iloc[-1]
            last_val_loss = hist['val_loss'].iloc[-1]
            last_train_loss = hist['loss'].iloc[-1]

            valor_prediction= np.argmax(prediction, axis=1)
            prediction_cm = confusion_matrix(np.argmax(test_labels, axis=1), valor_prediction)

            histories.append(hist)
            test_loss_list.append(test_loss)
            test_accuracy_list.append(test_accuracy)
            val_loss_list.append(last_val_loss)
            val_accuracy_list.append(last_val_accuracy)
            train_loss_list.append(last_train_loss)
            train_accuracy_list.append(last_train_accuracy)
            matrix_list.append(prediction_cm)
            predictions_list.append(valor_prediction)

""" Se guardan todos los datos en un archivo h5 """
f5 = h5py.File("red_bucle_5capas.hdf5", "w")
dset1 = f5.create_dataset("histories", (8,15,4), dtype='float32', data=histories)
dset2 = f5.create_dataset("test_loss", (8,), dtype='float32', data=test_loss_list)
dset3 = f5.create_dataset("test_accuracy", (8,), dtype='float32', data=test_accuracy_list)
dset4 = f5.create_dataset("val_loss", (8,), dtype='float32', data=val_loss_list)
dset5 = f5.create_dataset("val_accuracy", (8,), dtype='float32', data=val_accuracy_list)
dset6 = f5.create_dataset("train_loss", (8,), dtype='float32', data=train_loss_list)
dset7 = f5.create_dataset("train_accuracy", (8,), dtype='float32', data=train_accuracy_list)
dset8 = f5.create_dataset("matrix", (8,2,2), dtype='float32', data=matrix_list)
dset9 = f5.create_dataset("predictions", (8,n_test), dtype='float32', data=predictions_list)
f5.close()

```