



Universidad de Valladolid

FACULTAD DE CIENCIAS

TRABAJO FIN DE GRADO
Grado en Física

Clasificación automática de imágenes de cielo mediante Inteligencia Artificial

Autor:

Carolina Calvo Herrero

Tutores:

Dr. Roberto Román Díez

Dr. Juan Carlos Antuña Sánchez

Facultad de Ciencias
Valladolid, febrero de 2023

Dedicado a mis abuelos, Isabel y Mariano.

Resumen

Debido a los desafíos que ha supuesto a lo largo de la historia la predicción de la nubosidad, el desarrollo de la tecnología de inteligencia artificial es cada día más necesario para los meteorólogos. El objetivo de este trabajo es la clasificación de imágenes de cielo mediante inteligencia artificial. Para ello se ha realizado el etiquetado de imágenes procedentes de una cámara de cielo según el número de octas que se observan. El número de octas nos proporcionará la nubosidad de la imagen. A continuación utilizaremos dichas imágenes para el entranamiento de una red neuronal que nos permita predecir el estado del cielo. En el presente trabajo se han realizado 18 modelos diferentes en el entrenamiento de la red neuronal para que aprenda a clasificar imágenes. Tras esto se procede a la validación del conjunto de imágenes, denominado “conjunto de validación”, para analizar diferentes métricas de estudio y realizar las correspondientes valoraciones de los resultados y su viabilidad. Hemos obtenido buenos resultados para diferentes métricas alcanzando tasas de acierto, al permitir errores de una octa, del 92%.

Abstract

Due to the challenges that cloud prediction has posed throughout history, the development of artificial intelligence technology is increasingly necessary for meteorologists. The goal of this work is the classification of sky images using artificial intelligence. To achieve this, images from a sky camera have been labeled according to the number of oktas observed, which provides the cloudiness of the image. We will then use these images to train a neural network to predict the state of the sky. In this study, 18 different models were used to train the neural network to classify images. After training, a set of validation images, called the "validation set," was used to analyze different study metrics and evaluate the results and their feasibility. We have achieved good results for different metrics, with accuracy rates of 92%, allowing errors of one okta

Índice

1. Capítulo 1: Introducción	7
1.1. Nubes	7
1.1.1. El ciclo hidrológico	7
1.1.2. Clasificación de las nubes	8
1.1.3. La atmósfera y la vida	12
1.2. El tiempo atmosférico	13
1.3. Equipos para medir la cubierta nubosa	14
1.4. Inteligencia Artificial	15
1.4.1. Redes neuronales convolucionales	16
1.5. Objetivos	19
2. Capítulo 2: Instrumentación y Metodología	20
2.1. Cámara de cielo	20
2.2. Etiquetado de las imágenes de cielo	21
2.3. Entrenamiento de la red neuronal	21
3. Capítulo 3: Resultados	24
3.1. Error absoluto medio	25
3.2. Error medio	26
3.3. Desviación estándar	27
3.4. Raíz del error cuadrático medio	28
3.5. Distribución de las diferencias	29
3.6. Matrices de confusión	32
4. Capítulo 4: Conclusiones	35

Índice de figuras

1.	Ciclo del agua en la atmósfera [21]	8
2.	Los cirros están hechos de hielo puro y se producen en lo alto de la troposfera [23].	9
3.	Cirrostratus [23]	9
4.	Cirrocumulus [23].	9
5.	Altoestratos sobre Kentucky [23].	9
6.	Foto de la presencia de altocúmulos [10].	10
7.	Nimbostratus [23].	10
8.	Nubes estratos en l puente Golden Gate de San Francisco [10].	11
9.	Stratocumulus [23].	11
10.	Cumulus	11
11.	Cúmulonimbos [10].	11
12.	Los diez géneros de nubes organizados por alturas [12].	12
13.	En el hemisferio Norte, los vientos circulan describiendo un giro en el sentido de las agujas del reloj en los anticiclones, y en el sentido contrario en las borrascas (en el hemisferio Sur es a la inversa).	13
14.	Estructura típica de una red neuronal.	16
15.	Función ReLU.	18
16.	Estructura típica de una CNN para la detección de perros, gatos y pájaros en imágenes.	18
17.	Cámara de cielo OMEA 3C [27].	20
18.	Imágenes tomadas con la cámara de cielo.	21
19.	Elección de tipo de proyecto con "Teachable Machine 2.0".	22
20.	Selección de clases y parámetros en "Teachable Machine 2.0".	22
21.	Histogramas de los modelos 1-6, representando la frecuencia de las diferencias entre los valores predichos y reales.	29
22.	Histogramas de los modelos 7-16, representando la frecuencia de las diferencias entre los valores predichos y reales.	30
23.	Histogramas de los modelos 17 y 18, representando la frecuencia de las diferencias entre los valores predichos y reales.	31
24.	Matrices de confusión de los modelos 1-6.	32
25.	Matrices de confusión de los modelos 7-16.	33
26.	Matrices de confusión de los modelos 17 y 18.	34

Índice de tablas

1.	Clasificación de las nubes según su altura y latitud	11
2.	Número de imágenes de validación para cada etiqueta según el número de octas.	24
3.	Notación de los distintos modelos.	24
4.	Error absoluto medio en octas.	25
5.	Error medio en octas.	26
6.	Desviación estándar en octas.	27
7.	Raíz del error cuadrático medio en octas.	28
8.	Tasas de acierto de cada modelo.	34

1. Capítulo 1: Introducción

Las nubes son importantes para la humanidad por varias razones: regulan el clima y el ciclo de agua en la Tierra, son un indicador del estado de la atmósfera y la meteorología, inspiran arte, literatura y cultura, son un factor importante para la aviación, son una fuente potencial de energía renovable, etc. En resumen, las nubes son un componente crucial del sistema climático y tienen un impacto significativo en la vida humana y la economía [21].

El análisis de nubes es una tarea importante en meteorología y climatología, ya que permite conocer la cubierta de nubes y su impacto en el clima y en la producción de energía. Con el avance en inteligencia artificial, se ha abierto la posibilidad de utilizar técnicas de reconocimiento de objetos para identificar nubes en imágenes tomadas por una cámara de cielo. Este proyecto tiene como objetivo principal analizar la viabilidad de utilizar inteligencia artificial para reconocer la cubierta de nubes a partir de imágenes obtenidas por una cámara de cielo. Con este estudio se pretende contribuir al desarrollo de soluciones más eficientes y precisas en el análisis de nubes [22].

1.1. Nubes

Las nubes son formaciones características de la troposfera (capa de la atmósfera más cercana a la Tierra), que están ligadas al ciclo del agua de nuestro planeta. Su proceso de formación es muy simple: como consecuencia de la evaporación en la superficie, se produce continuamente el ascenso de vapor de agua a la atmósfera. Conforme asciende, el vapor se enfría, con lo que se produce la condensación del agua en forma de pequeñas gotas líquidas o su solidificación como diminutos cristales de hielo [11].

Este enfriamiento puede suceder también cuando en la atmósfera colisionan frentes fríos y cálidos: el súbito enfriamiento del vapor de agua contenido en la masa de aire caliente provoca el desarrollo de nubes. En ambos casos, la condensación del agua supone una importante liberación de energía.

Normalmente, la presencia de partículas de polvo, polen, etc., en la atmósfera ayuda a la formación de las nubes, ya que dichas partículas actúan como núcleos de condensación del agua.

Las partículas de agua o los cristales de hielo de las nubes se mantienen suspendidos en la atmósfera debido a su poco peso. Si, como consecuencia de cambios en la temperatura, vientos, choque con obstáculos naturales (montañas), etc., se produce la unión de estas partículas, estas aumentan de peso y, finalmente, la gravedad hace que se produzca la precipitación. Esta puede ser en forma de agua, nieve o de granizo, en función de las condiciones atmosféricas de ese momento [11].

1.1.1. El ciclo hidrológico

El agua está en constante movimiento en la naturaleza. La lluvia y la corriente de los ríos son muestras de este movimiento. El agua pasa continuamente de unos lugares a otros del planeta: de la atmósfera cae a la tierra, a los ríos y a los mares, y de todos estos, por evaporación, vuelve a la atmósfera. Este movimiento continuo del agua de unos puntos a otros es lo que se denomina ciclo hidrológico.

El movimiento del agua de los ríos hacia el mar es efecto de la gravedad: se produce a causa de

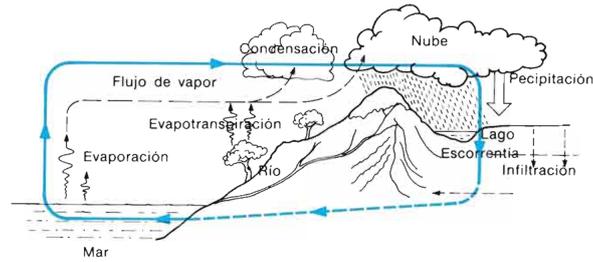


Figura 1: Ciclo del agua en la atmósfera [21]

la pendiente del terreno. El paso del agua a los glaciares y las nieves de las montañas a los ríos se debe a un cambio de estado, la fusión; y el paso del agua de mares, lagos y ríos a la atmósfera, a la evaporación. También se evapora el agua que liberan las hojas de los árboles. Tal y cómo se ve en la figura 1, el vapor de agua se enfría en la atmósfera y se condensa, formando gotitas de agua. Cuando estas tienen un cierto tamaño, caen como precipitaciones: lluvia, nieve o granizo. Así pasa el agua de la atmósfera a la superficie terrestre [11].

Estos sencillos cambios físicos, activados por la energía de la radiación solar y la gravedad, hacen que el agua circule constantemente por el planeta.

1.1.2. Clasificación de las nubes

Mirando al cielo durante días diferentes es posible comprobar que existen distintos tipos de nubes. Se pueden distinguir por la forma, el color, el aspecto (deshilachado, algodonoso, globoso), el color (claro u oscuro) y la altitud a la que aparecen.

La clasificación de Luke Howard de 1803 es un sistema para clasificar las nubes según su altura, forma y densidad. Esta clasificación se utiliza aún en la actualidad y divide las nubes en tres grandes grupos: cumulos, estratos y cirros.

El Atlas internacional de Nubes de la OMM, distingue 10 géneros, 14 especies, 9 variedades y 9 rasgos suplementarios [11]:

1. Género: es la característica principal de la nube. Existen 10 tipos de géneros de nubes: *Cirrus*, *Cirrocumulus*, *Cirrostratus*, *Alto cumulus*, *Altostratus*, *Nimbostratus*, *Stratocumulus*, *Stratus*, *Cumulus*, y *Cumulonimbus*.
2. Especie: clasifica la forma y estructura interna de la nube. Existen 14 especies: *fibratus*, *uncinus*, *spissatus*, *castellanus*, *floccus*, *stratiformis*, *nebulosus*, *lenticularis*, *fractus*, *humilis*, *mediocris*, *congestus*, *calvus* y *capillatus*.
3. Variedad: define la disposición y transparencia de la nube. Hay 9 tipos: *intortus*, *vertebratus*, *undulatus*, *radiatus*, *lacunosus*, *duplicatus*, *translucidus*, *perlucidus* y *opacus*.
4. Rasgos suplementarios y nubes accesorias: son nubes que se unen a la parte principal de una nube. Hay 9 rasgos suplementarios y nubes accesorias: *incus*, *mamma*, *virga*, *praecipitatio*, *arcus*, *tuba*, *pileus*, *velum* y *pannus*.

El género clasifica las nubes por su altura, forma y densidad. A continuación se explica cada uno en detalle:



Figura 2: Los cirros están hechos de hielo puro y se producen en lo alto de la troposfera [23].



Figura 3: Cirrostratus [23]



Figura 4: Cirrocumulus [23].



Figura 5: Altoestratos sobre Kentucky [23].

1. **Cirrus:** nubes de color blanco y aspecto alargado y "deshilachado", que se forman en capas muy altas (por encima de 9000m). Suelen estar formados por cristales de hielo. Véase Figura 2.
2. **Cirrostratus:** se caracteriza por ser una nube de alta altura y densidad uniforme, con forma de velo que cubre todo el cielo. Están compuestas por cristales de hielo y son de color blanco o gris pálido. A menudo, estas nubes son la prelude de un mal tiempo y su presencia suele indicar que la lluvia o la nieve están en camino. Véase Figura 3.
3. **Cirrocumulus:** es una nube de alta altura y está compuesta por cristales de hielo. Se caracteriza por su aspecto de pequeñas olas o mosaicos en el cielo, con una textura lisa y uniforme. El color de estas nubes es blanco o gris pálido. A menudo se las ve en formaciones paralelas o en patrones regulares y pueden ser un indicativo de un cambio en el tiempo, especialmente si están asociados con otros tipos de nubes. Véase Figura 4.
4. **Altoestratos:** están situadas a una altura intermedia, entre los cirros y los estratos, y son nubes de baja densidad y uniformes. Se caracterizan por tener un aspecto uniforme y extenderse por todo el cielo, pero sin cubrirlo completamente. El color de estas nubes suele ser gris o blanco pálido. A menudo son una señal de un tiempo estable y sin lluvia, pero también pueden ser un indicativo de un cambio en el tiempo en combinación con otros tipos de nubes. Véase Figura 5.
5. **Alto cumulus:** están situadas a una altura intermedia, entre los cirros y los cumulus, y se caracterizan por su aspecto de masas blancas o grises en forma de flecos o anillos. Pueden



Figura 6: Foto de la presencia de altocúmulos [10].



Figura 7: Nimbostratus [23].

tener diferentes formas y tamaños y suelen aparecer en grupos o en patrones regulares. El color de estas nubes suele ser blanco o gris pálido. A menudo son un indicativo de un cambio en el tiempo, especialmente si están asociados con otros tipos de nubes. Los esponjosos altocúmulos ondulados tiñen el cielo de rojo por encima de las hileras de árboles. Estas nubes de nivel medio, que se forman por encima de los 2.000 metros, a menudo anuncian que se aproxima una tormenta (véase Figura 6).

6. **Nimbostratus:** nubes grises, a veces muy bajas, que cubren el cielo y originan precipitaciones. Se agrega la palabra "nimbo" al nombre de una nube para indicar que se trata de una nube que suele provocar lluvias. Véase Figura 7.
7. **Stratus:** nubes dispuestas en varias capas, que cubren normalmente todo el cielo y suelen estar a baja altitud. Se presentan habitualmente en invierno y, a veces, van asociados a nieblas. En la Figura 10 podemos ver estas nubes de bajo nivel. Se encuentran por debajo de los 2000 metros y con frecuencia traen lluvia o nieve. Las nubes que tocan el suelo se conocen como niebla. Véase Figura 8.
8. **Stratocumulus:** se encuentra en la baja altura y se caracteriza por su aspecto de masas bajas y uniformes con bordes definidos. Pueden tener diferentes formas y tamaños y suelen aparecer en grupos o en patrones regulares. El color de estas nubes suele ser gris oscuro o blanco pálido. A menudo son un indicativo de tiempo nublado o lluvioso, pero también pueden ser una señal de un tiempo estable sin lluvia. Véase Figura 9.
9. **Cumulus:** nubes de aspecto algodonoso, blancas o blanquecinas, con forma de cúpula en su parte superior. No suelen producir precipitaciones. Véase Figura 10.
10. **Cumulonimbus:** es una nube de tormenta, que se caracteriza por su aspecto de una gran masa vertical y densa. Estas nubes suelen ser muy oscuras y están formadas por agua y vapor de agua. Son responsables de los fenómenos meteorológicos más intensos, como los rayos, los truenos, las tormentas fuertes y las tormentas tropicales. En la Figura 11 podemos ver su característica parte superior con forma de yunque que se produce debido a los fuertes vientos laterales que se encuentran a tales altitudes.



Figura 8: Nubes estratos en l puente Golden Gate de San Francisco [10].



Figura 9: Stratocumulus [23].



Figura 10: Cumulus



Figura 11: Cúmulonimbos [10].

Si, en cambio, se agrupan las nubes por su altitud, utilizando la clasificación de Genova, es posible distinguir nubes bajas, medias y altas. Las nubes bajas se encuentran en altitudes inferiores a 2000 m ,constituidas por agua líquida. Incluyen nubes como los cumulus y los stratus. Las nubes medias se encuentran entre 2000 m y 7000 m de altitud, formadas por una mezcla de hielo y agua. Incluyen nubes como los alto cumulus y los altostratus. Las nubes altas se encuentran por encima de los 7000 m de altitud y están constituidas por cristales de hielo. Incluyen nubes como los cirrus y los cirrostratus.

A continuación se muestra una tabla con la estimación de alturas de los diferentes pisos de nubes en función de la latitud y la época estacional.

Tabla 1: Clasificación de las nubes según su altura y latitud

Nivel	Región Polar	Región Templada	Región Tropical
Nubes altas	3-8 km	3-13 km	6-18 km
Nubes medias	2-4 km	2-7 km	2-8 km
Nubes bajas	Superficie-2 km	Superficie-2 km	Superficie-2 km

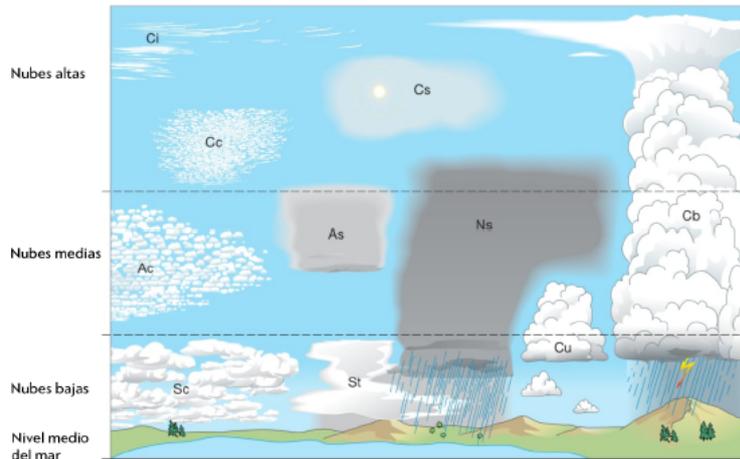


Figura 12: Los diez géneros de nubes organizados por alturas [12].

1.1.3. La atmósfera y la vida

El Sol emite radiación dentro del espectro electromagnético. Además de la luz visible, las más importantes para la Tierra son la radiación ultravioleta y la infrarroja. La radiación ultravioleta es muy energética, tanto que si llegara en su totalidad a la superficie de la Tierra, podría causar daño a los seres vivos. Se considera que este tipo de radiación es cancerígena. La Tierra por otro lado emite radiación infrarroja y el balance entre la radiación que llega del Sol y la emitida por la Tierra es la que determina la temperatura del planeta [24].

La Tierra se puede asemejar a un gran invernadero en el que la cubierta es la atmósfera. Determinados gases atmosféricos, como el dióxido de carbono, son los responsables de que la atmósfera actúe dejando pasar la radiación del Sol, que calienta la superficie, e impidiendo que abandone parte de la radiación infrarroja emitida por la Tierra. Así la presencia de la atmósfera suaviza la temperatura del planeta (la media es de unos 15°C) y favorece el mantenimiento de la vida.

El efecto invernadero es un fenómeno natural, y una de las claves de que el planeta sea apto para albergar vida. No así el incremento de este efecto, grave problema medioambiental que se debe al aumento de la concentración de los “gases de invernadero” en la atmósfera [25]. Esta concentración anormal de estos gases hace que el efecto invernadero se intensifique, lo cual puede conducir, en muy poco tiempo, a un calentamiento global. Este calentamiento sería peligroso para la vida en el planeta y podría tener importantes consecuencias para la población humana.

La capa de ozono estratosférica ejerce la importante función de repeler la radiación ultravioleta. Esta radiación es incompatible con la vida: si pudiera llegar libremente a la Tierra, probablemente la vida no se hubiera formado tal y como la conocemos, o hubiera sido difícil que evolucionase y se extendiese [26].

Las actividades industriales, los transportes y la minería vierten gran cantidad de sustancias que pueden contaminar la atmósfera y afectar negativamente a la salud de las personas. La principal fuente de emisión de contaminantes atmosféricos son los automóviles, las calefacciones domésticas y las industrias. Tanto para nuestra salud como para la de los demás seres vivos, y para el clima, los contaminantes suponen un problema muy grave. La contaminación se ha convertido en un problema global cuyas consecuencias afectan a todo el planeta.

1.2. El tiempo atmosférico

Nubes, viento, tormentas, precipitaciones son manifestaciones de la dinámica de la atmósfera. No son exclusivas de la Tierra: la gran tormenta de Júpiter es un fenómeno atmosférico que ha durado más de cien años.

Los fenómenos atmosféricos se producen por las diferencias de temperatura y presión en distintos puntos de la atmósfera. La presión atmosférica es un factor determinante para la aparición de muchos de estos fenómenos. En cualquier mapa del tiempo se pueden delimitar zonas de altas presiones y de bajas presiones relativas.

Los anticiclones (señalados con una A en los mapas meteorológicos) son zonas de la atmósfera donde la presión es alta (ver Figura 13). Suele ir asociados a situaciones de buen tiempo o de niebla.

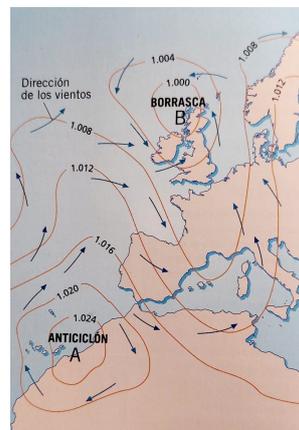


Figura 13: En el hemisferio Norte, los vientos circulan describiendo un giro en el sentido de las agujas del reloj en los anticiclones, y en el sentido contrario en las borrascas (en el hemisferio Sur es a la inversa).

En cambio, las borrascas (B) son zonas de la atmósfera donde la presión es más baja y el aire tiende a ascender. Suelen ir asociadas a situaciones de inestabilidad, en las que hay frecuentes precipitaciones de lluvia o nieve. Normalmente se desplazan de oeste a este.

Una de las formas más fiables de predecir el tiempo a corto plazo (desde unas horas a unos dos o tres días) es la observación continua de la atmósfera. En la actualidad, este estudio es posible gracias a la red de satélites artificiales que envían continuamente fotografías de la atmósfera. El análisis de estas fotografías permite seguir la evolución de grandes frentes de nubes, tormentas, zonas de claros, etc. Combinando estas imágenes con los datos de presión en diferentes puntos de la atmósfera, es posible construir el mapa meteorológico de la situación actual y los de las previsiones para las horas o días siguientes.

1.3. Equipos para medir la cubierta nubosa

Esta sección muestra algunos instrumentos de medición de cobertura nubosa siguiendo el trabajo de Tapakis y Charalambides [14]. Hay dos tipos de equipos: los de base terrestre y los satélites. Los primeros se usan para tomar datos locales, mientras que los satélites miden áreas extensas. Los equipos terrestres tienen una visión limitada, pero ofrecen medidas precisas sobre una ubicación fija. Por otro lado, los satélites tienen baja resolución temporal y pueden tener errores, pero cubren amplias áreas.

Los equipos terrestres se agrupan en cuatro categorías. La primera incluye los piranómetros y radiómetros, que miden la irradiancia y un rango espectral específico. Se basan en la idea de que cualquier cuerpo con temperatura superior a 0 absoluto emite radiación electromagnética, cuyo espectro se puede usar para identificar objetos como el Sol, nubes, aire, tierra, mar, hielo, etc. La segunda categoría es la de las cámaras de cielo que fotografían el cielo en el espectro visible, infrarrojo o ultravioleta. La tercera incluye radares que miden la retrodispersión de una señal electromagnética emitida por el propio radar. Miden la diferencia de tiempo y la pérdida de intensidad de la señal para determinar la posición de las nubes. Los ceilómetros, parte de este tercer grupo, miden la altura de la base de las nubes. La última categoría es la radiosonda, incluida en globos meteorológicos para monitorear las nubes. Los globos, llenos de helio, alcanzan altitudes superiores a 30km. La radiosonda incluye sensores que miden temperatura, presión del aire y humedad relativa. Durante su ascenso, detecta nubes proporcionando información sobre su altura de base, espesor y altura superior. Las radiosondas son útiles porque pueden detectar múltiples capas de nubes en cada nivel de la atmósfera y el movimiento horizontal del globo indica la velocidad y dirección del viento que mueve las nubes [13].

Las cámaras de cielo son un tipo destacado de equipos de tierra debido a su facilidad de instalación, mantenimiento y bajo costo. Se pueden colocar en muchos lugares y cubrir una amplia área. Son utilizadas para fotografiar el cielo y suelen estar compuestas de un sensor CCD o CMOS y una lente de ojo de pez.

Las cámaras de cielo se utilizan comúnmente para detectar la cubierta de nubes [14], pero también han sido empleadas en tareas más avanzadas, como la predicción de la irradiancia solar [15], la deducción de la radiancia del cielo y medidas de luminiscencia [16], y la medición de propiedades de los aerosoles [17]. Aunque son menos precisas que otros fotómetros, pueden obtener un mapeado hemisférico de la radiancia del cielo en un corto periodo de tiempo. Además, los sensores de las cámaras pueden ajustarse para obtener imágenes de alta resolución. Estas propiedades hacen que las cámaras de cielo sean un instrumento valioso que merece ser estudiado para ser combinado con algoritmos de inteligencia artificial (IA).

1.4. Inteligencia Artificial

La inteligencia artificial, IA, ha mejorado el reconocimiento de objetos en imágenes [18]. Este trabajo se enfoca en evaluar la aplicación de la IA para identificar nubes en imágenes tomadas por cámaras de cielo. El objetivo principal es analizar la viabilidad de usar imágenes de cámaras de cielo para reconocer la cubierta de nubes mediante IA.

Los estudios que dieron origen a lo que hoy conocemos como Inteligencia Artificial moderna surgieron en la década de 1940. Sin embargo, no fue hasta 1950 con la publicación del artículo de Alan Turing “Computing Machinery and Intelligence” cuando estos estudios comenzaron a tener una gran influencia. En este artículo, Alan Turing especulaba sobre la posibilidad de que una máquina pudiera imitar el comportamiento humano y propuso el famoso Test de Turing para medir la inteligencia de una máquina y compararla con la de un humano [2].

La inteligencia artificial simbólica se basa en la idea de crear inteligencias artificiales que comprendan el mundo a través de representaciones simbólicas, similares a como los humanos lo hacen. Se desarrollaron algoritmos basados en la lógica para expresar cada situación en un programa y se utilizaron para crear “sistemas expertos” en la década de 1980. Sin embargo, estos sistemas tienen limitaciones como un alto costo, necesidad de actualizaciones constantes y pueden ser menos precisos con más reglas. Además, no mejoran con el tiempo o se adaptan a nuevas situaciones debido a su naturaleza basada en reglas estrictas.

Actualmente, el enfoque dominante en el campo de la inteligencia artificial es el aprendizaje automático (machine learning, ML). En lugar de intentar crear algoritmos basados en la lógica para imitar el pensamiento humano, el aprendizaje automático se enfoca en descubrir las reglas que explican el comportamiento deseado a partir de datos reales.

Redes neuronales Convolucionales (CNNs, por sus siglas en inglés) son un tipo de red neuronal utilizadas para el procesamiento de imágenes. Estas redes se caracterizan por tener capas de conexión que se aplican de manera local a una imagen, lo que permite que la red aprenda características específicas de una imagen. Estas características se pueden utilizar para clasificar imágenes, detectar objetos, generar imágenes y más.

Una de las principales ventajas de las CNNs es su capacidad para aprender características de imágenes de manera automática, lo que las hace especialmente útiles para tareas de visión artificial. En comparación con otras redes neuronales, las CNNs tienen menos parámetros, lo que las hace más fáciles de entrenar y menos propensas a sobreajuste.

Una de las características fundamentales de las redes neuronales convolucionales es su estructura de capas. Estas redes constan de varias capas de conexión, cada una de las cuales se aplica de manera local a una imagen. La primera capa, llamada capa de entrada, es la que recibe la imagen. La última capa, llamada capa de salida, es la que produce la salida de la red. Entre estas dos capas se encuentran varias capas intermedias, llamadas capas ocultas.

Las capas ocultas de una CNN se dividen en dos tipos: capas de conexión y capas de pooling. Las capas de conexión son las que aprenden características de las imágenes. Estas capas se componen de un conjunto de filtros, cada uno de los cuales se aplica de manera local a la imagen. Cada filtro aprende una característica específica de la imagen, como las líneas, los bordes, los colores, etc.

Las capas de pooling son las que reducen la dimensionalidad de la imagen. Estas capas se aplican de manera local a la imagen y reducen la resolución de la misma. Esto ayuda a la red a ser más robusta a las variaciones en la posición y el tamaño de los objetos en la imagen.

1.4.1. Redes neuronales convolucionales

En esta sección nos centraremos en la utilización de la tecnología de inteligencia artificial para analizar las imágenes de cámaras de cielo.

Tradicionalmente, se han utilizado algoritmos para identificar zonas blancas y azules en las imágenes de cielo. Sin embargo, actualmente se están utilizando técnicas de inteligencia artificial para analizar de manera eficiente la nubosidad a partir de imágenes de cámaras de cielo. Esto es importante para clasificar adecuadamente las nubes futuras y también para analizar volúmenes de datos pasados. Se están utilizando redes neuronales convolucionales entrenadas para reconocer el número de octas de la bóveda celeste con una precisión y efectividad adecuadas [4].

Los algoritmos de redes neuronales son fundamentales en el aprendizaje automático, se inspiran en la estructura de las neuronas del cerebro y se componen neuronas artificiales [5].

Las redes neuronales convolucionales utilizan una técnica llamada “convolución” para detectar patrones en las imágenes, permitiendo a la red aprender características específicas de estas. Las CNN se han utilizado con éxito en una variedad de aplicaciones, como la reconocimiento de rostros, reconocimiento de objetos en imágenes y análisis de imágenes médicas.

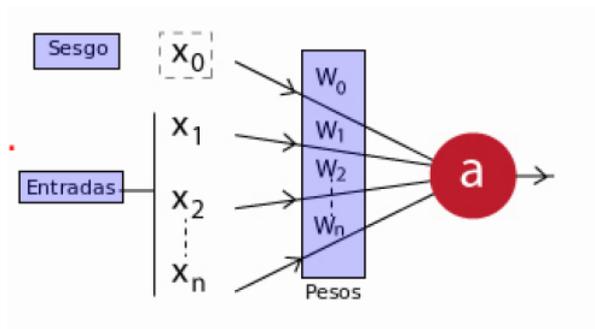


Figura 14: Estructura típica de una red neuronal.

Una neurona artificial es un modelo matemático que simula el comportamiento de una neurona biológica. En una red neuronal artificial, cada neurona recibe entradas, aplica una serie de operaciones matemáticas y produce una salida. La salida de una neurona puede ser utilizada como entrada para otras neuronas en la red.

En una red neuronal artificial, las entradas son generalmente valores numéricos (x_0, x_1, \dots, x_i) , y las operaciones matemáticas se realizan mediante un conjunto de pesos (w_0, \dots, w_i) y una función de activación (f). Los pesos son valores numéricos que determinan la importancia de cada entrada para la salida de la neurona, mientras que la función de activación es una función matemática que limita el rango de salida de la neurona (a). Ver ecuación (1).

$$a = f\left(\sum_{i=0}^n w_i \cdot x_i\right) \quad (1)$$

En resumen, las neuronas artificiales son una representación matemática de las neuronas biológicas, que se utilizan para construir redes neuronales artificiales, las cuales son capaces de aprender y realizar tareas complejas a partir de datos de entrada.

La operación de la convolución consiste en multiplicar una matriz de entrada y una matriz bidimensional de pesos, llamada filtro o kernel. Estos filtros son normalmente más pequeños que la matriz de entrada y permiten realizar la operación sistemáticamente mientras se recorre la imagen. El filtro aprende a reconocer características de las imágenes sin importar su posición, lo que implica invarianza traslacional.

Las redes convolucionales aprovechan esta invarianza para usar los mismos parámetros en el filtro y reducir las imágenes con el pooling.

El resultado final es una matriz bidimensional conocida como mapa de características. Las capas de convolución suelen contener varios filtros diferentes para que las redes neuronales puedan aprender a identificar múltiples características.

Una CNN se compone de varias capas, cada una de las cuales tiene un propósito específico en el procesamiento de imágenes. A continuación se describen algunas de las capas más comunes en una CNN [6]:

1. **Capa de entrada:** esta capa es donde se introduce la imagen original. La imagen se convierte en un tensor (un arreglo multidimensional) antes de ser procesada por la red.
2. **Capa de convolución:** estas capas son responsables de detectar patrones y características en la imagen. Cada capa de Convolución utiliza un conjunto de filtros, que se aplican localmente a la imagen. Los filtros buscan patrones específicos en la imagen, como bordes, texturas, etc.
3. **Capas de pooling:** estas capas son responsables de reducir la dimensionalidad de la imagen. Esto se hace mediante la aplicación de operaciones de agrupamiento, como la media o el máximo. Esto ayuda a la red a detectar características más generales en la imagen.
4. **Capas completamente conectadas:** estas capas son similares a las capas de una red neuronal tradicional. Se utilizan para realizar la clasificación final de la imagen.
5. **Capa de activación:** son responsables de aplicar una función de activación a los valores de salida de cada capa [h]. La función de activación tiene como objetivo agregar no linealidad a la red, lo que permite a la red aprender patrones más complejos en los datos. Algunas funciones de activación más comunes son: tahn, sigmoid, softmax, Leaky ReLU y ReLU. ReLU (Rectified Linear Unit) es una de las funciones de activación más populares en las CNNs. La función ReLU es simple y eficiente, y se define como $f(x) = \max(0, x)$ (véase Figura 15). Esta función cambia los valores negativos de entrada por un cero mientras que los valores positivos permanecen inalterables [7].
6. **Capas de dropout:** es una técnica de regularización utilizada en redes neuronales para reducir el sobreajuste. La idea detrás de esta técnica es eliminar aleatoriamente un porcentaje de las neuronas en cada capa durante el entrenamiento. Al hacer esto, se evita que las neuronas se especialicen en patrones específicos de los datos de entrenamiento, lo que ayuda a la red a generalizar mejor a los datos de prueba.
7. **Capa de salida:** esta capa proporciona la salida de la red, que puede ser una clase o una etiqueta para la imagen.

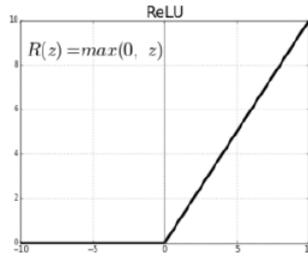


Figura 15: Función ReLU.

Además de estas capas, las CNNs también pueden incluir capas adicionales, como capas de normalización entre otros. Dependiendo del problema, la arquitectura de una CNN puede variar, pero las capas de entrada, convolución, pooling y capas completamente conectadas son comunes en la mayoría de las CNNs. Podemos observar en la Figura 16 un ejemplo de su arquitectura. A través de la función *softmax* se transmiten las probabilidades a la neurona de salida. En la salida se realiza una clasificación *one-hot-encoding*, el cual es un vector de ceros con un uno en la posición que corresponde a la etiqueta de la clasificación [9].

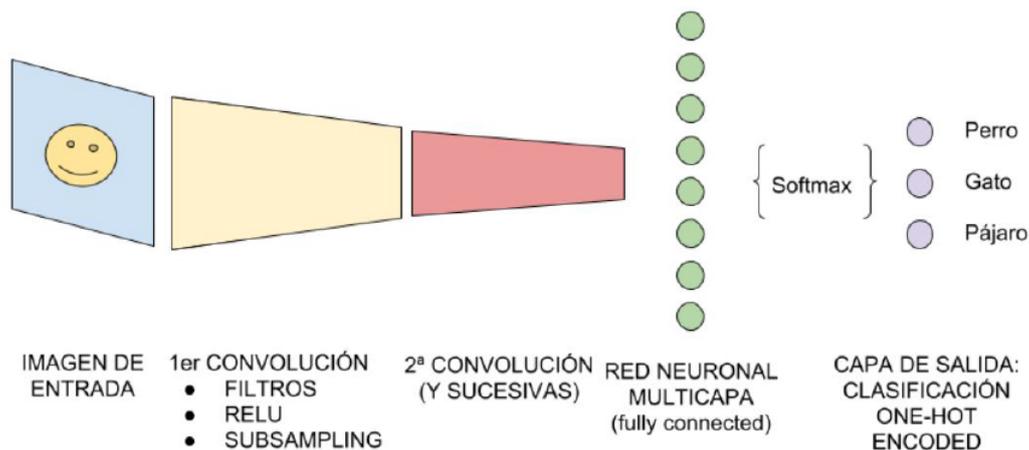


Figura 16: Estructura típica de una CNN para la detección de perros, gatos y pájaros en imágenes.

Para algoritmos supervisados, se utiliza el proceso de Backpropagation para aprender a través de la comparación de la salida generada por la red con la deseada, calcular errores y propagarlos hacia atrás a través de la red para actualizar los pesos de las neuronas y reducir la diferencia entre el resultado esperado y el obtenido [5].

1.5. Objetivos

El objetivo principal de este trabajo es obtener y evaluar una red neuronal convolucional que sea capaz de predecir la cubierta nubosa a partir de imágenes capturadas con una cámara de cielo. Para esto utilizaremos imágenes tomadas tanto de día como de noche que entrenaremos adecuadamente para su después evaluación.

Este documento TFG se ha dividido en 5 capítulos, siendo el primer capítulo la Introducción.

-Capítulo 2: Metodología

En esta sección se describen las propiedades de la cámara de cielo con la que se han obtenido las imágenes. Explicamos cómo se ha realizado el etiquetado de imágenes así como el método de entrenamiento de la red neuronal que hemos utilizado.

-Capítulo 3: Resultados

En este capítulo analizamos las configuraciones obtenidas tras la validación de las imágenes median inteligencia artificial.

-Capítulo 4: Conclusiones

Se detallan las conclusiones obtenidas del trabajo realizado planteando futuras líneas de investigación.

-Capítulo 5: Bibliografía

Bibliografía utilizada a lo largo del Trabajo de Fin de Grado.

2. Capítulo 2: Instrumentación y Metodología

A lo largo de este capítulo se desarrolla la metodología de este Trabajo Fin de Grado. Se explica la instrumentación necesaria para la captura de imágenes. Esta es la cámara de cielo OMEA 3C perteneciente y controlada por el “Grupo de Óptica Atmosférica” de la Universidad de Valladolid [27].

Se detalla el etiquetado de las imágenes así como la configuración para el entrenamiento de la red neuronal y su posterior evaluación. Para facilitar esto, se ha utilizado la herramienta de aprendizaje automático de código abierto creada por Google: “Teachable Machine”. A continuación gracias a “Google Colab” pudimos escribir y ejecutar el código python sin necesidad de instalación de software.

2.1. Cámara de cielo

La cámara de cielo es un instrumento utilizado en meteorología para medir la cobertura nubosa y la posición de las nubes en el cielo. Es un dispositivo de fotografía automática que toma fotografías panorámicas del cielo con una frecuencia regular, generalmente cada 5 o 10 minutos [1].

La cámara empleada para la realización del presente trabajo está ubicada en la terraza de la facultad de Ciencias de la Universidad de Valladolid. Para esta cámara se han clasificado imágenes tanto de día como de noche. Además se han utilizado imágenes de una cámara OMEA 3C instalada en Fuentecaliente (La Palma) durante la erupción del volcán de Cumbre Vieja, otra OMEA 3C instalada en el observatorio meteorológico de Izaña (Tenerife), y por último una cámara SONA 202-NF instalada en Valladolid. Estas tres cámaras solo se han usado para clasificar imágenes de día [13] [4].



Figura 17: Cámara de cielo OMEA 3C [27].

El modelo principal de cámara utilizada para tomar imágenes de noche es el OMEA3C de Alcor System (se puede ver en la Figura 17). Este dispositivo consiste en un sensor CMOS combinado con una lente de pez, los cuales están protegidos dentro de una cúpula de cristal transparente e impermeable [4]. El sensor CMOS empleado es el SONY IMX178 y está programado para capturar imágenes con una resolución de 3096x2080 píxeles y una profundidad de color de 14 bits. Este sensor cuenta con un filtro Bayer que tiene un patrón RGGB, donde la mitad de los píxeles son sensibles al verde, una cuarta parte a rojo y otra cuarta parte al azul [13].

Cada imagen capturada es una combinación HDR de varias imágenes con secuencias multi-exposición, cada 5 minutos por el día y cada 2 minutos por la noche. El número de imágenes incluidas en la combinación HDR varía dependiendo de la altura del sol [4].

2.2. Etiquetado de las imágenes de cielo

Esta es la parte del proyecto en la cual se procedió al etiquetado manual de las imágenes nocturnas correspondientes al mes de enero de 2021. Las fotografías, tanto las nocturnas tomadas en Valladolid, como las diurnas correspondientes a La Palma y Tenerife, se tomaron en intervalos de dos y 5 minutos siendo 3361 el número total de fotos etiquetadas.

Las etiquetas utilizadas fueron las siguientes:

1. **Número de octas:** esto identificará la nubosidad que presenta el cielo. Con un etiquetado del 0 al 8 podremos saber si el cielo está completamente despejado (0 octas) o completamente cubierto de nubes (8 octas). Esta medida depende del observador por tanto al ser subjetiva se asume un error de ± 1 octa. En la Figura 18 podemos apreciar algunos ejemplos.

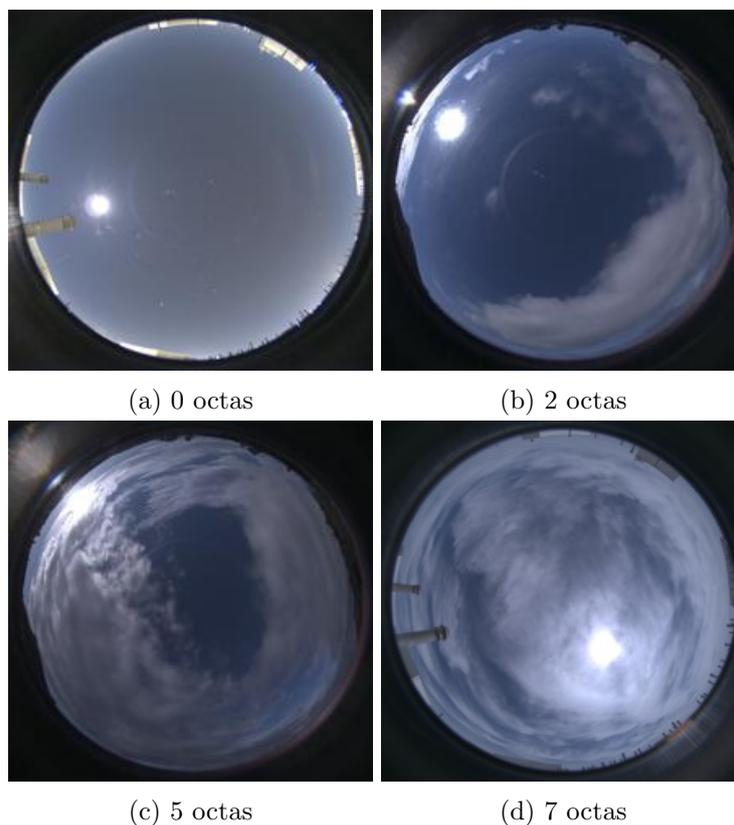


Figura 18: Imágenes tomadas con la cámara de cielo.

2.3. Entrenamiento de la red neuronal

En este apartado se detallan los pasos a seguir para el entrenamiento de la red neuronal. Tras el etiquetado de las imágenes, éstas se dividen en dos conjuntos, entrenamiento y validación. El conjunto de entrenamiento está formado por el 80 % de las imágenes y se utiliza para entrenar

la red neuronal. El restante 20 % de las imágenes son las que llamaremos conjunto de validación y con las cuales realizaremos las pruebas necesarias para cuantificar el modelo entrenado. Las imágenes han sido elegidas de forma que tengamos uniformidad en el número de octas. Además cabe destacar que hemos utilizado mitad de imágenes de noche y mitad imágenes de día del total.

Para entrenar los modelos se utiliza una red neuronal sobre la cual podremos clasificar las imágenes según el número de octas que cubre el cielo.

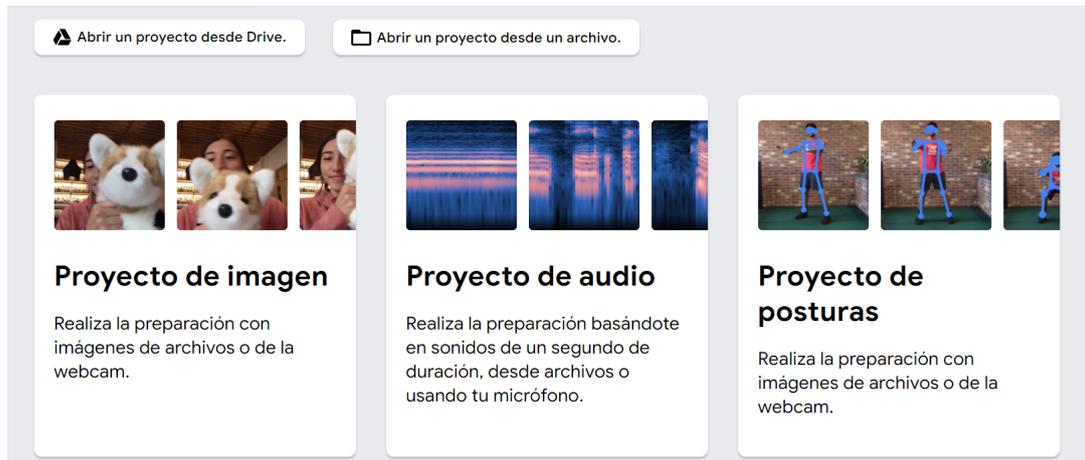


Figura 19: Elección de tipo de proyecto con "Teachable Machine 2.0".

"Teachable Machine 2.0" es una herramienta de aprendizaje automático de código abierto desarrollada por Google. Permite a los usuarios crear modelos de aprendizaje automático personalizados utilizando datos de entrada de imágenes, audio y texto. Los modelos creados pueden ser utilizados para clasificar y reconocer patrones en los datos de entrada, y se pueden exportar para su uso en aplicaciones y proyectos. Es fácil de usar y no requiere conocimientos previos en programación o aprendizaje automático. Es una interfaz totalmente gratuita y se puede acceder a ella a través del siguiente enlace: <https://teachablemachine.withgoogle.com/train>.

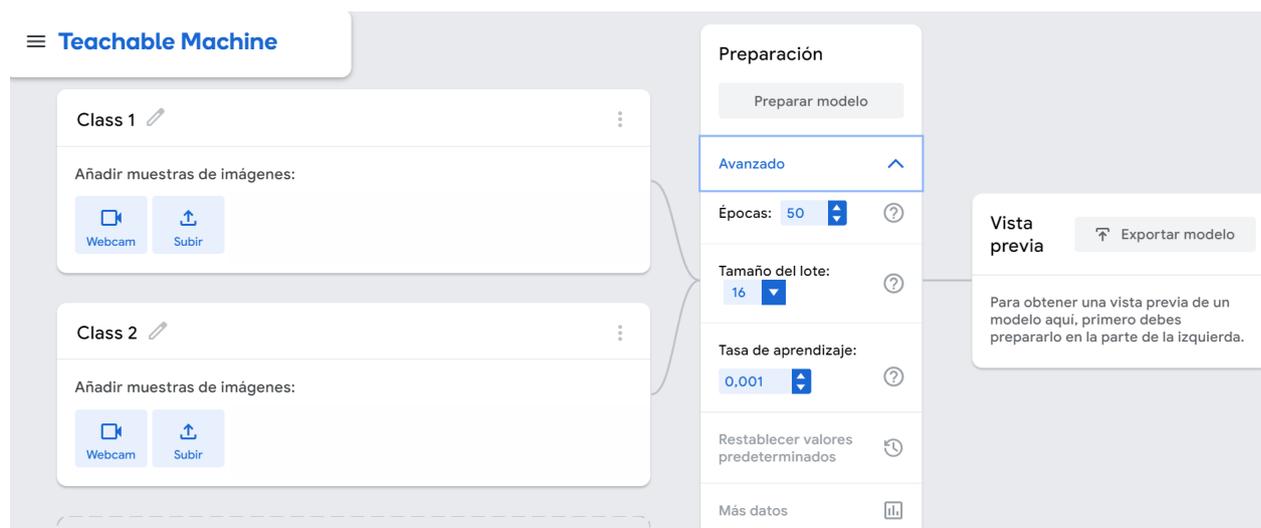


Figura 20: Selección de clases y parámetros en "Teachable Machine 2.0".

Una vez hayamos accedido al enlace anterior, seleccionaremos "Proyecto de imagen" (Figura 19). Tras esto, debemos cargar las imágenes en diferentes clases correspondientes a nuestras etiquetas

(de 0 a 8 para el número de octas y 0 ó 1 si tenemos sol obstruido o despejado)(véase Figura 20). A continuación se establecen los parámetros de entrenamiento. Estos son los siguientes:

- **Épocas:** una época se refiere a una iteración completa a través del conjunto de entrenamiento. Durante cada época, el modelo de aprendizaje automático procesa todos los ejemplos de entrenamiento y actualiza sus parámetros para adaptarse mejor a los datos. El número de épocas especificado en el entrenamiento del modelo determina cuántas veces el modelo procesará el conjunto de entrenamiento antes de detenerse. A medida que aumenta el número de épocas, el modelo puede mejorar su precisión al procesar los datos de entrenamiento más veces, pero también puede aumentar el riesgo de sobreajuste. Por lo tanto es fundamental una buena elección del mismo. Su valor por defecto es 50.
- **Tamaño del lote:** el tamaño del lote se refiere al número de ejemplos de entrenamiento que se utilizan en una sola iteración del algoritmo de optimización. El tamaño del lote se utiliza para calcular las actualizaciones de los parámetros del modelo en cada iteración, y se especifica en número de ejemplos. Por ejemplo, si el tamaño del lote es de 32, entonces el algoritmo de optimización utilizará 32 ejemplos para actualizar los parámetros del modelo en cada iteración. Puede ser ajustado según las necesidades del modelo y las características del conjunto de datos. Un tamaño de lote más grande puede aumentar la velocidad de entrenamiento, pero también puede requerir más memoria. Su valor por defecto es 16.
- **Tasa de aprendizaje:** la tasa de aprendizaje es un número entre 0 y 1 que controla la magnitud de las actualizaciones de los parámetros en cada iteración del algoritmo de optimización. Este parámetro afecta la velocidad a la que el modelo aprende y la precisión final del modelo. Una tasa de aprendizaje demasiado alta puede causar que el modelo oscile y no converja, mientras que una tasa de aprendizaje demasiado baja puede causar que el modelo tarde demasiado en converger. Es importante ajustar la tasa de aprendizaje adecuadamente para obtener una buena precisión del modelo y un tiempo de entrenamiento razonable. Su valor por defecto es 0.001.

Se han realizado 18 modelos diferentes combinando los parámetros mencionados, con el objetivo de encontrar los mejores resultados. Los valores utilizados han sido los siguientes: **Épocas: 50, 100, 200, Tamaño del lote: 16, 32 y Tasa de aprendizaje: 0.01, 0.001, 0.0001.**

Así, tras el entrenamiento de los modelos, estos se exportan en formato “.h5”. En la plataforma “Google Colab” cargaremos los modelos utilizando un código Python. Este código carga dos modelos de Tensorflow (Keras) previamente entrenados y guardados en el disco, recorre una carpeta de imágenes llamada “validar” y utiliza cada imagen para hacer una predicción con ambos modelos. Una vez que se han procesado todas las imágenes, se guardan los resultados en un archivo “.csv”.

3. Capítulo 3: Resultados

Vamos a exponer los resultados obtenidos para cada clase de etiqueta. Tenemos un total de 18 modelos que compararemos objetivamente y así poder concluir si las CNN son competentes para la clasificación de imágenes de cielo. Así podremos concluir que modelo es el más apropiado para esta tarea.

Compararemos la etiqueta predicha por el modelo, con la etiqueta real, es decir, la etiquetada por nosotros mismos. En la Tabla 2 se exponen el número de imágenes del conjunto validación utilizadas para cada etiqueta en octas.

Tabla 2: Número de imágenes de validación para cada etiqueta según el número de octas.

Etiqueta número de octas	0	1	2	3	4	5	6	7	8
Imágenes de validación	199	88	63	57	47	43	57	50	68

Utilizaremos la siguiente notación de la Tabla 3 para evaluar cada modelo:

Tabla 3: Notación de los distintos modelos.

Modelo	Épocas	Tamaño de lote	Tasa de aprendizaje
1	50	16	0,01
2	50	16	0,001
3	50	16	0,0001
4	50	32	0,01
5	50	32	0,001
6	50	32	0,0001
7	100	16	0,01
8	100	16	0,001
9	100	16	0,0001
10	100	32	0,01
11	100	32	0,001
12	100	32	0,0001
13	200	16	0,01
14	200	16	0,001
15	200	16	0,0001
16	200	32	0,01
17	200	32	0,001
18	200	32	0,0001

3.1. Error absoluto medio

El error absoluto medio (MAE, por sus siglas en inglés) es una medida de la diferencia entre los valores reales y los valores previstos o estimados [19]. Se calcula como la suma de los errores absolutos dividida por el número de datos. Matemáticamente, se puede representar como:

$$\frac{1}{N} \cdot \sum_{i=1}^N |x_i - y_i| \quad (2)$$

donde N es el número de datos, x_i es el valor predicho para el i -ésimo dato y y_i es el valor real para el i -ésimo dato.

El error absoluto medio es una medida robusta que no se ve afectada por valores atípicos o extremos, y es comúnmente utilizada en la evaluación de modelos de machine learning. Sin embargo, también tiene algunas limitaciones, como no penalizar los errores grandes más que los errores pequeños. Por lo tanto, a veces se utiliza en conjunto con otras métricas como la raíz del error cuadrático medio o el error medio.

En la tabla 4 se presentan los valores calculados para la métrica en cuestión para cada modelo y cada etiqueta distinta. Además, se incluye el valor de la métrica para el conjunto completo de imágenes de validación procesadas por cada modelo analizado.

Tabla 4: Error absoluto medio en octas.

Error Absoluto Medio (octas)										
Modelo	0	1	2	3	4	5	6	7	8	Todas
1	0,04	0,31	0,67	0,67	1,40	1,33	1,68	1,22	0,18	0,60
2	0,09	0,52	0,71	0,92	0,68	1,12	0,79	1,14	0,28	0,54
3	0,05	0,43	0,25	0,76	1,06	1,09	0,88	0,66	0,24	0,47
4	0,06	0,44	0,79	1,14	0,92	0,67	0,68	0,50	0,37	0,48
5	0,25	0,55	0,86	0,92	0,75	0,79	1,04	0,52	0,24	0,51
6	0,08	0,44	0,54	1,00	0,83	0,93	0,98	0,70	0,19	0,49
7	0,09	0,39	0,49	0,72	1,21	1,49	1,93	1,26	0,47	0,67
8	0,06	0,40	0,59	0,76	0,85	0,88	0,84	0,80	0,25	0,46
9	0,06	0,31	0,48	0,82	0,85	1,00	0,86	0,70	0,21	0,44
10	0,23	0,28	0,35	0,84	0,83	0,81	1,09	0,78	0,37	0,51
11	0,09	0,30	0,73	0,90	0,66	0,98	1,30	0,76	0,31	0,52
12	0,07	0,44	0,46	0,80	0,96	0,93	0,88	0,54	0,25	0,46
13	0,09	0,27	0,38	0,88	0,77	1,07	1,24	0,76	0,18	0,48
14	0,10	0,43	0,54	0,86	0,81	0,91	1,11	0,62	0,28	0,49
15	0,04	0,47	0,57	0,90	0,94	0,91	1,07	0,88	0,22	0,50
16	0,02	0,71	0,60	1,18	1,11	0,88	0,97	0,52	0,21	0,53
17	0,02	0,47	0,75	0,86	0,83	0,95	0,98	0,72	0,29	0,49
18	0,07	0,36	0,60	0,78	0,79	0,77	0,93	0,70	0,22	0,45

Podemos observar, que el error absoluto medio crece de 0 a 4 octas y de 4 a 8 comienza a decrecer. Podemos concluir que las octas extremas, 0 y 8 octas, están muy bien definidas mientras que los cielos parcialmente cubiertos son más variables.

El valor obtenido para el total del número de imágenes no es la media de el resto de valores.

Respecto a la comparación entre modelos, vemos un comportamiento similar en todas las etiquetas, siendo el modelo número 9 el que mejores resultados presenta para todas las imágenes.

3.2. Error medio

El error medio (también conocido como la media aritmética de los errores) es una medida de la diferencia entre los valores predichos y los valores reales en un conjunto de datos. Se utiliza para estimar la *exactitud* [20].

El error medio se calcula sumando los errores absolutos para cada punto de datos y dividiéndolos por el número de puntos de datos. El error absoluto es simplemente la diferencia entre el valor previsto y el valor real.

Matemáticamente, el error medio se puede expresar como:

$$\frac{1}{N} \cdot \sum_{i=1}^N (x_i - y_i) \quad (3)$$

Es importante tener en cuenta que el error medio es una medida global de la exactitud de un modelo.

Mostramos en la Tabla 5 los resultados obtenidos. Tenga en cuenta que la calidad del modelo se corresponde con la información de esta métrica, ya que una métrica cercana a cero reflejará una gran exactitud.

Tabla 5: Error medio en octas.

Error Medio (octas)										
Modelo	0	1	2	3	4	5	6	7	8	Todas
1	0,04	-0,06	-0,38	-0,35	-0,85	-0,67	-1,09	-0,42	-0,18	-0,30
2	0,10	0,16	0,40	0,56	-0,13	0,05	-0,51	-0,54	-0,28	0,01
3	0,10	-0,05	0,27	0,04	0,26	0,23	-0,56	-0,38	-0,24	-0,07
4	0,06	-0,03	0,48	0,62	0,49	0,44	-0,30	-0,38	-0,37	0,07
5	0,25	0,25	0,60	0,56	0,40	-0,05	-0,54	-0,32	-0,24	0,09
6	0,08	0,06	0,06	0,12	0,15	0,23	-0,56	-0,30	-0,19	-0,03
7	0,09	0,11	-0,24	-0,36	-0,92	-1,21	-1,74	-0,90	-0,47	-0,42
8	-0,06	0,24	0,30	0,20	0,00	-0,09	-0,49	-0,48	-0,25	-0,02
9	0,06	0,03	0,13	0,22	-0,17	0,12	-0,44	-0,38	-0,21	-0,04
10	0,23	0,22	0,06	0,20	0,06	-0,30	-0,74	-0,46	-0,37	-0,04
11	0,09	0,09	0,16	0,42	-0,06	-0,32	-0,98	-0,48	-0,31	-0,09
12	0,07	0,06	0,14	0,12	0,06	0,00	-0,42	-0,30	-0,25	-0,03
13	0,09	0,00	-0,10	-0,12	-0,21	-0,56	-1,00	-0,28	-0,18	-0,17
14	0,09	0,09	0,29	0,42	-0,30	-0,16	-0,65	-0,34	-0,28	-0,05
15	0,04	0,06	0,13	0,34	0,00	-0,21	-0,54	-0,44	-0,22	-0,06
16	0,02	-0,18	0,25	0,58	0,43	0,42	-0,19	0,00	-0,21	0,06
17	0,02	0,03	0,14	0,39	0,06	-0,02	-0,42	-0,44	-0,29	-0,05
18	0,07	0,14	0,32	0,14	0,02	-0,07	-0,65	-0,42	-0,22	0,04

Observamos que para valores de octas altos, nuestros modelos subestiman el resultado (valores negativos), mientras que para los valores de octas bajos los modelos (a excepción del modelo 8) sobreestiman el resultado. Estos resultados eran esperables pues la etiqueta cero octas solo se puede sobreestimar y la de 8 octas infraestimar. Se concluye que los modelos 2 y 8 son los que mayor exactitud muestran, aunque todos presentan gran exactitud pues tienen errores medios muy próximos al cero.

3.3. Desviación estándar

La desviación estándar es una medida de la dispersión de un conjunto de datos alrededor de su media. Es una medida de la variabilidad de los datos y representa la cantidad de variación o “dispersión” de los datos en torno a la media.

La desviación estándar se calcula como la raíz cuadrada de la varianza, que es otra medida de la dispersión de los datos. La varianza se calcula como la media de los cuadrados de las diferencias entre cada dato y la media del conjunto de datos.

Matemáticamente, la desviación estándar se puede calcular como:

$$\sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x}_i)^2}{N}} \quad (4)$$

La desviación estándar es una medida para describir la precisión del modelo. Es útil en muchas aplicaciones, incluyendo estadísticas, investigación de mercado, física y mucho más [20].

Tabla 6: Desviación estándar en octas.

Desviación estándar (octas).										
Modelo	0	1	2	3	4	5	6	7	8	Todas
1	0,20	0,62	0,72	1,00	1,61	1,49	1,46	1,82	0,46	1,06
2	0,40	1,08	0,90	1,10	1,04	1,29	1,14	1,35	0,89	0,99
3	0,23	0,69	0,63	1,13	1,39	1,32	1,24	0,91	0,60	0,87
4	0,27	0,79	1,10	1,29	1,05	0,87	0,94	0,89	0,54	0,87
5	1,05	1,19	1,02	1,13	1,02	1,16	1,17	0,93	0,55	0,99
6	0,28	1,00	0,83	1,32	1,25	1,22	1,27	1,01	0,55	0,93
7	0,29	0,94	0,68	0,89	0,99	1,34	1,50	1,87	1,25	1,18
8	0,27	0,89	0,87	1,04	1,19	1,40	1,19	1,02	0,65	0,90
9	0,28	0,59	0,72	1,15	1,29	1,31	1,17	0,98	0,53	0,84
10	0,50	0,57	0,73	1,08	1,16	1,09	1,32	1,24	0,86	0,94
11	0,32	0,81	1,07	1,06	1,10	1,38	1,37	1,15	0,73	0,98
12	0,29	0,90	0,73	1,09	1,39	1,35	1,15	0,94	0,58	0,88
13	0,60	0,71	0,71	1,05	1,13	1,37	1,43	1,04	0,45	0,94
14	0,62	0,94	0,79	1,19	1,03	1,31	1,46	0,84	0,59	0,97
15	0,23	0,92	0,88	1,19	1,20	1,29	1,42	1,10	0,54	0,93
16	0,14	0,87	1,03	1,44	1,30	1,28	1,32	0,89	0,58	0,95
17	0,14	1,03	0,96	1,13	1,17	1,30	1,27	1,17	0,60	0,93
18	0,28	0,84	0,89	1,11	1,06	1,21	1,24	0,92	0,57	0,87

En la Tabla 6 podemos observar que solo las etiquetas de 0, 1, 2 y 8 octas tienen valores menores que 1, contrastando y validando la idea de cometer un error de ± 1 octa. Como en las métricas anteriores, los valores centrales presentan mayor dispersión que los valores extremos. Si tenemos en cuenta todos los datos, todos los modelos, excepto el modelo 1 y 7, presentan buenos resultados sin embargo es el modelo 9 el que mejores resultados tiene.

3.4. Raíz del error cuadrático medio

La raíz del error cuadrático medio (también conocido como RMSE, por sus siglas en inglés) es una medida de la diferencia entre los valores reales y los valores previstos en un conjunto de datos. Se utiliza para evaluar el rendimiento de un modelo de predicción o para comparar diferentes modelos [19].

El error cuadrático medio se calcula sumando los errores cuadrados para cada punto de datos y dividiéndolos por el número de puntos de datos. El error cuadrático es la diferencia entre el valor previsto y el valor real elevado al cuadrado.

La raíz del error cuadrático medio se calcula como la raíz cuadrada del error cuadrático medio. Matemáticamente, se puede expresar como:

$$\sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (x_i - y_i)^2} \quad (5)$$

El RMSE es una medida más precisa que el error medio, ya que penaliza más los errores grandes y refleja mejor la magnitud de los errores en el modelo. Sin embargo, puede ser sensible a valores atípicos en los datos y no es necesariamente una medida óptima para todas las aplicaciones.

Tabla 7: Raíz del error cuadrático medio en octas.

Raíz del Error Cuadrático Medio (octas).										
Modelo	0	1	2	3	4	5	6	7	8	Todas
1	0,20	0,62	0,82	1,06	1,71	1,61	2,12	1,75	0,46	1,11
2	0,41	1,08	0,98	1,05	1,05	1,29	1,25	1,45	0,93	0,99
3	0,24	0,69	0,56	1,13	1,41	1,32	1,36	0,99	0,62	0,87
4	0,28	0,79	1,20	1,44	1,16	0,98	0,98	0,97	0,65	0,87
5	1,08	1,22	1,18	1,27	1,10	1,16	1,29	0,98	0,59	0,99
6	0,29	1,01	0,84	1,33	1,26	1,24	1,39	1,05	0,58	0,93
7	0,30	0,94	0,72	0,96	1,35	1,80	2,31	2,07	1,34	1,25
8	0,28	0,92	0,92	1,06	1,19	1,40	1,28	1,13	0,65	0,90
9	0,28	0,59	0,74	1,18	1,31	1,00	1,32	1,25	1,05	0,57
10	0,55	0,61	0,74	1,10	1,16	1,13	1,51	1,32	0,93	0,94
11	0,33	0,81	1,08	1,14	1,10	1,41	1,69	1,25	0,80	0,99
12	0,63	0,30	0,91	0,75	1,10	1,39	1,35	1,23	0,99	0,88
13	0,61	0,71	0,71	1,06	1,15	1,48	1,74	1,08	0,49	0,96
14	0,62	0,94	0,84	1,26	1,07	1,32	1,60	0,91	0,65	0,97
15	0,24	0,92	0,89	1,24	1,20	1,30	1,52	1,18	0,58	0,93
16	0,14	0,89	1,07	1,56	1,37	1,35	1,33	0,89	0,62	0,94
17	0,14	1,03	0,97	1,19	1,18	1,30	1,34	1,25	0,66	0,93
18	0,28	0,85	0,94	1,12	1,06	1,21	1,40	1,01	0,61	0,87

Para los valores de la raíz cuadrada del error cuadrático medio (véase Tabla 7), observamos de nuevo que los valores para las etiquetas intermedias presentan mayor error que los valores de las etiquetas 0 y 8 octas. El modelo 9 es el que mejores resultados globales presenta con un valor de 0,57 octas.

3.5. Distribución de las diferencias

Un histograma es una representación gráfica que muestra la frecuencia con la que aparecen los valores en un conjunto de datos. Es una herramienta importante en estadísticas y son útiles para identificar tendencias y patrones en los datos, así como para evaluar la forma de la distribución.

A continuación se presentan histogramas de la distribución de las diferencias entre el modelo y las medidas para cada modelo, los cuales fueron creados usando la librería “matplotlib” de Python. En el eje X se muestran las discrepancias entre el número de octas predicho y el real, mientras que en el eje Y se indica la cantidad de imágenes que tienen esa diferencia, es decir, su frecuencia absoluta. Estos histogramas nos permitirán evaluar la información dada por la desviación estándar sobre la distribución de los valores predichos. Sería ideal que hubiera una gran cantidad de resultados concentrados en torno a cero, ya que eso significaría un elevado número de discrepancias nulas y una baja distribución.

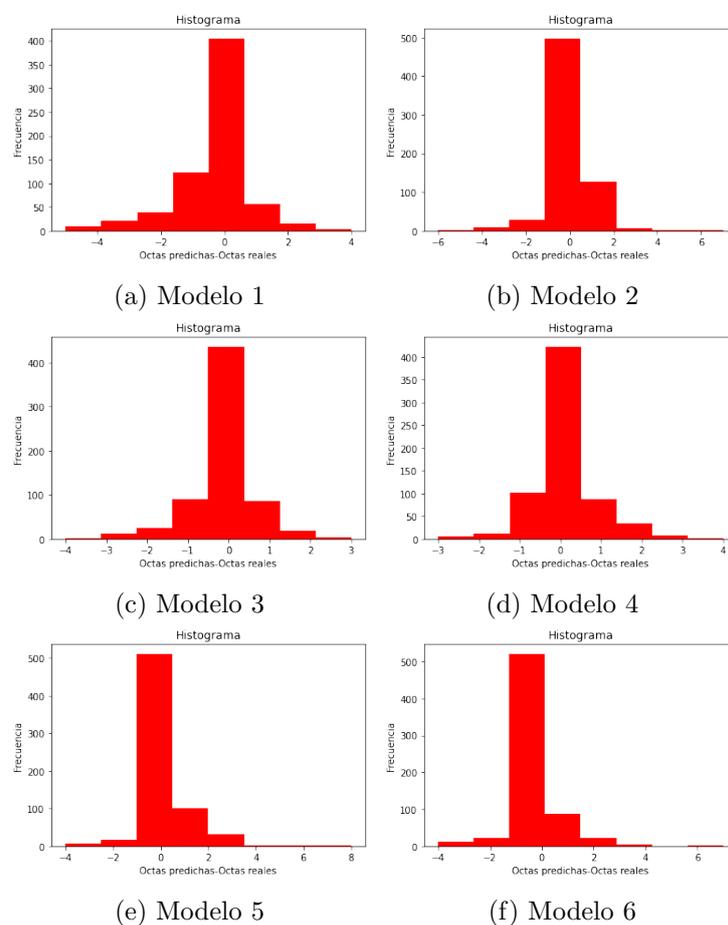


Figura 21: Histogramas de los modelos 1-6, representando la frecuencia de las diferencias entre los valores predichos y reales.

En la Figura 21, Figura 22 y Figura 23, observamos que los valores se concentran entorno al cero, decayendo la frecuencia de forma progresiva hacia valores más extremos del eje x. Se puede decir, sin duda alguna, que todos los modelos son eficaces para la clasificación, y su rendimiento es muy bueno. No hay diferencias notables entre ellos, aunque es cierto que en los modelos 2, 5, 6, 8, 11, 12 y 14 hay una distinción significativa alrededor de las seis octas.

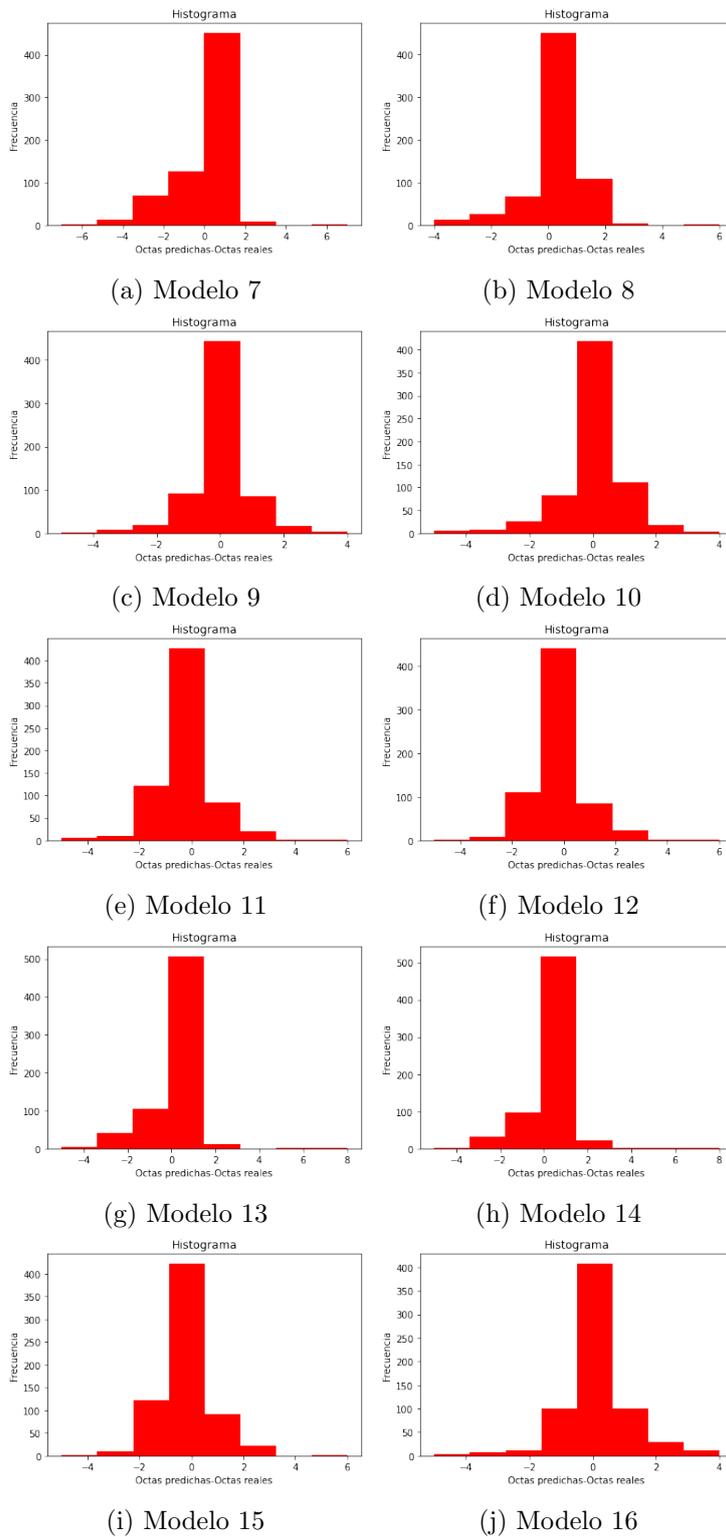


Figura 22: Histogramas de los modelos 7-16, representando la frecuencia de las diferencias entre los valores predichos y reales.

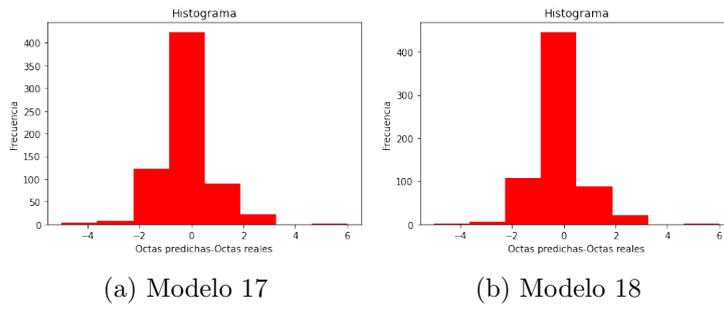


Figura 23: Histogramas de los modelos 17 y 18, representando la frecuencia de las diferencias entre los valores predichos y reales.

Además los modelos 3 y 9, en concordancia a lo obtenido en los resultados de la desviación estándar, presentan una mayor dispersión pues si nos fijamos en sus valores extremos tienen los mayores valores de frecuencia absoluta.

3.6. Matrices de confusión

Una matriz de confusión es una tabla que se utiliza para evaluar el rendimiento de un modelo de clasificación. La matriz de confusión muestra la cantidad de veces que un modelo predice correctamente o incorrectamente cada clase en un conjunto de datos.

La matriz de confusión tiene dos ejes: en el eje horizontal se listan las clases reales de los datos, mientras que en el eje vertical se listan las clases predichas por el modelo. Cada celda en la matriz representa una combinación de una clase real y una clase predicha. La diagonal principal representa las predicciones correctas, mientras que las otras celdas representan predicciones incorrectas. Por lo tanto, los mejores modelos tendrán valores más elevados en la diagonal, pues el valor predicho y el real coincidirán.

Mostramos las matrices de confusión de cada modelo en las Figuras 24 y 25. Hemos añadido un mapa de calor en cada una de ellas para obtener conclusiones de forma rápida y visual.

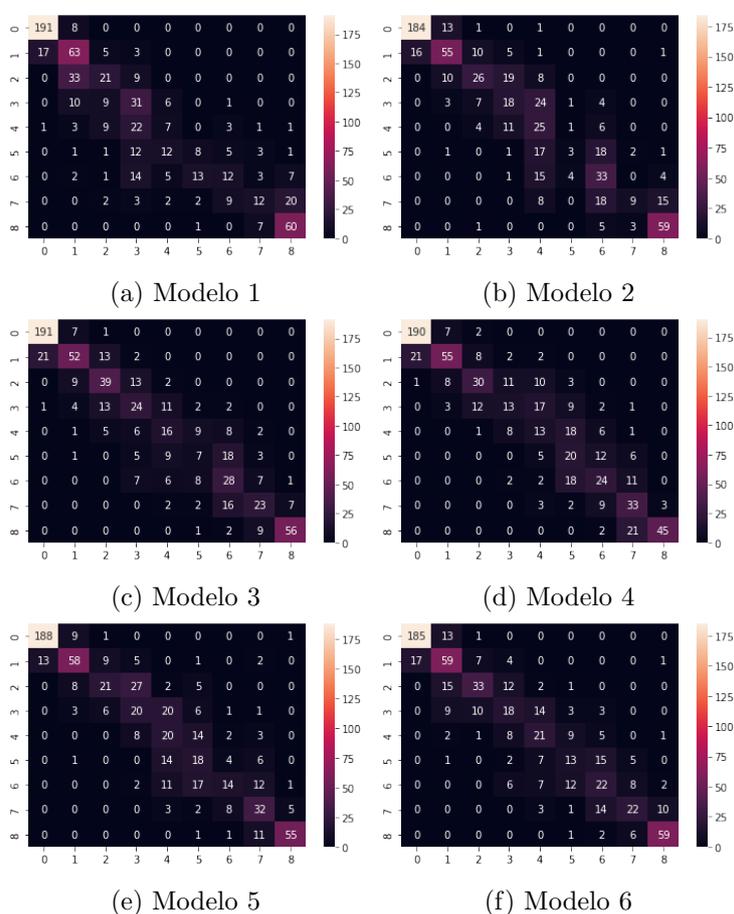


Figura 24: Matrices de confusión de los modelos 1-6.

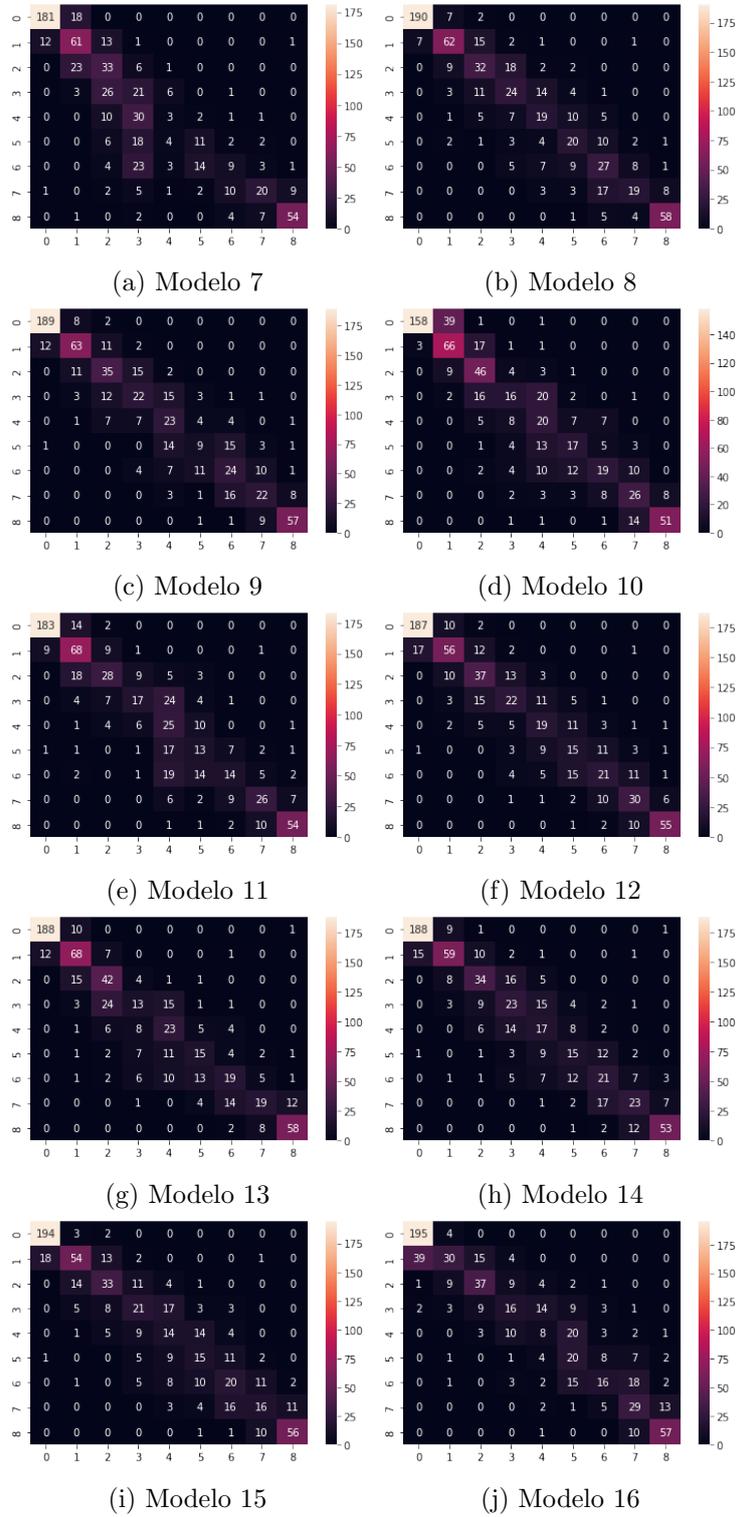


Figura 25: Matrices de confusión de los modelos 7-16.

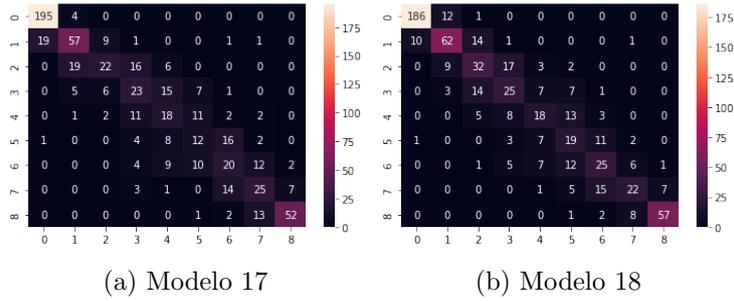


Figura 26: Matrices de confusión de los modelos 17 y 18.

De nuevo, tal y como muestran las Figuras 24 y 25, vemos esa diferencia entre los resultados extremos y los intermedios. Las tonalidades blancas muestran valores altos mientras que las moradas oscuras en las zonas intermedias reflejan los valores bajos.

Vemos como la mayoría de los datos en nuestras matrices se concentran en la diagonal, lo cual demuestra un buen comportamiento de nuestros modelos. Los valores de las octas intermedias son menores que los de las octas de los extremos pues como hemos explicado en las métricas anteriores, existe más subjetividad a la hora de seleccionar el número de octas para imágenes de cielo parcialmente cubiertas de nubes.

En la Tabla 8 mostramos la tasa de acierto total de cada modelo en tanto por uno, es decir, el cociente entre el número de imágenes predichas correctamente (la suma de la diagonal) y el total de imágenes. A este dato lo denominamos como Tasa de Acierto 1. También mostramos la tasa de acierto de cada modelo, pero en este caso el numerador serán las imágenes predichas más aquellas que contengan un error de ± 1 octa. A esto lo denominamos en la Tabla 8 como Tasa de Acierto 2. Esta segunda tasa contempla el error cometido por el investigador en el etiquetado.

Tabla 8: Tasas de acierto de cada modelo.

Modelo	Tasa de Acierto 1	Tasa de Acierto 2
1	0,60	0,87
2	0,61	0,89
3	0,65	0,91
4	0,63	0,91
5	0,63	0,91
6	0,64	0,91
7	0,59	0,86
8	0,67	0,91
9	0,66	0,93
10	0,62	0,91
11	0,64	0,90
12	0,66	0,92
13	0,66	0,91
14	0,64	0,91
15	0,63	0,91
16	0,61	0,91
17	0,63	0,91
18	0,66	0,92

Podemos concluir que el modelo 9 es el que presenta mejores resultados en la clasificación de imágenes de cielo pues tiene las mayores tasas de acierto. Para la Tasa de Acierto 2, vemos que obtenemos unos porcentajes alrededor del 93%.

4. Capítulo 4: Conclusiones

A lo largo de este trabajo se ha mostrado un método para reconocer la cubierta de nubes mediante inteligencia artificial utilizando imágenes obtenidas por una cámara de cielo. Estas imágenes fueron entrenadas con diferentes modelos para estimar el número de octas presente en las mismas. Como novedad, respecto a trabajos anteriores, este trabajo ha sido entrenado para predecir imágenes de cielo tanto de día como de noche. Tras la evaluación de los modelos hemos obtenido las siguientes conclusiones:

- Los modelos empleados se han analizado con distintas métricas (error absoluto medio, error medio, desviación estándar y error cuadrático medio) mostrando, en general, buenos resultados. Aunque no tenemos grandes diferencias entre ellos, podemos afirmar que el modelo 9 (Épocas: 100, Tamaño del Lote: 16, Tasa de aprendizaje: 0,0001) es el que mejores resultados presenta basándonos en que es el de menor error cuadrático medio. Podemos deducir que, escogiendo la menor tasa de aprendizaje y un número de épocas intermedio (100), encontraremos mejores resultados para nuestro estudio en concreto.
- Se comprueba que los modelos estimarán el número de octas que hay en la imagen con un margen de error de ± 1 octa; tal y como se había supuesto inicialmente.
- Para cielos parcialmente despejados observamos que todos los modelos presentan una mayor dispersión, pues el número de octas es más subjetivo de etiquetar. Al entrenar la red, esta se encontrará con imágenes similares etiquetadas de manera distinta, lo que provocará que el modelo obtenido presente peores resultados en este rango de octas.
- Las tasas de acierto de los modelos son aceptables. De echo, cuando consideramos como correctas aquellas imágenes con un error de una octa, obtenemos un 93 % de acierto.

Tras la realización del presente proyecto se concluye que el presente método es un método económico y de fácil instalación, con el que se puede generar una red de cámaras de cielo para así obtener imágenes de cielo en diferentes regiones a tiempo real. Así, podemos construir una base de datos útil para futuros estudios. El único punto más tedioso del método, es el tiempo que lleva etiquetar cada imagen manualmente, una a una, para después entrenar y evaluar los modelos.

Por último, como futuras líneas de investigación se proponen dos estudios. El primero es el desarrollo de modelos de IA más rápidos y eficientes. Los modelos de IA actuales pueden tardar mucho tiempo en clasificar grandes conjuntos de datos de imágenes de cielo, lo que limita su utilidad para ciertas aplicaciones. La investigación en el desarrollo de modelos de IA más rápidos y eficientes podría permitir la clasificación de grandes cantidades de datos en un período de tiempo más corto. Por otro lado, se propone el estudio para la identificación de objetos desconocidos. La IA también puede ser utilizada para identificar objetos desconocidos en el cielo, como los llamados “objetos transitorios”, que son objetos que aparecen y desaparecen en el cielo, como las supernovas, en un período de tiempo relativamente corto. La identificación de estos objetos podría tener implicaciones importantes para la investigación astronómica y la comprensión del universo.

Capítulo 5: Bibliografía

- [1] Reyes Rodriguez, I. “Cámaras de cielo ¿más allá de las nubes?”. Trabajo Fin de Grado.
- [2] A. M. Turing, “I-COMPUTING MACHINERY AND INTELLIGENCE”, *Mind*, vol. LIX, pp 443-460, 10 1950
- [3] B. Raphael, “Sir: A computer program for semantic information retrieval”
- [4] Alegre Fernández, S. “Clasificación automática de imágenes del cielo mediante inteligencia artificial”
- [5] Información obtenida de: <https://youtu.be/MRIv2IwFTPg>
- [6] Bosch Rué An., Casas Roma J. y Lozano Bagén T. “Deep Learning, principios y fundamentos”, primera edición, voc, 2019
- [7] “Redes neuronales” <https://ml4a.github.io/mla4a/es/neuralnetworks/>
- [8] “Commonly used activation functions”. <https://cs231n.github.io/neuralnetworks-1/>
- [9] Bagnato J. I. “Convolutional Neural Networks: La Teoría explicada en Español” [Internet]. *Aprendemachinelearning.com*, 2018
- [10] National Geographic. <https://www.nationalgeographic.es/medio-ambiente/que-son-las-nubes-y-que-tipos-existen?gallery=12876image=1123.600x450>
- [11] Quirantes Calvo J. A. y Gallego Poveda J. A., “Atlas de nubes y meteoros”, primera edición, Torrelavega, España, Cantabria Tradicional S.L. 2011
- [12] OMM, O. M. M. (2021). “Clouds definitions — international cloud atlas”. <https://cloudatlas.wmo.int/es/clouds-definitions.html>.
- [13] Martinez Celda, B. “Clasificación automática de imágenes de cielo mediante inteligencia artificial”.
- [14] Tapakis, R. y Charalambides, A. G. (2013). “Equipment and methodologies for cloud detection and classification: A review”.
- [15] Alonso-Montesinos, J., Batlles, F. J., y Portillo, C. (2015). “Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images. *Energy Conversion and Management*”, 105.
- [16] Román, R., Antón, M., Cazorla, A., Miguel, A. D., Olmo, F. J., Bilbao, J., y Alados-Arboledas, L. (2012). “Calibration of an all-sky camera for obtaining sky radiance at three wavelengths. *Atmospheric Measurement Techniques*”, 5.
- [17] R. Román, B. Torres, D. Fuertes, V.E. Cachorro, O. Dubovik, C. Toledano, A. Cazorla, A. Barreto, J.L. Bosch, T. Lapyonok, R. González, P. Goloub, M.R. Perrone, F.J. Olmo, A. de Frutos, L. Alados-Arboledas, “Remote sensing of lunar aureole with a sky camera: Adding information in the nocturnal retrieval of aerosol properties with GRASP code, *Remote Sensing of Environment*”, Volume 196, 2017, Pages 238-252.
- [18] Hastie, T., Tibshirani, R., y Friedman, J. (2009). “The elements of statistical learning: data mining, inference, and prediction”. Springer Science Business Media.
- [19] Bosch Rue A. ,Casas Roma J. y Lozano Bagen T. “Deep Learning, principios y fundamentos, primera edicion”, UOC, 2019.

- [20] Taylor John R. “An introduction to error analysis, the study of uncertainties in physical measurements”, 1997.
- [21] García de Pedraza L. y García Vega J., “Las nubes como indicadores del tiempo”, Núm. 20/84 HD, Ministerio de Agricultura, Pesca y Alimentación, 1984.
- [22] Román R., Cazorla A., Toledano C., Olmo F. J., Cachorro V.E., De Frutos A., y Alados-Aarboledas L. (2017a), “Cloud cover detection combining high dynamic range sky images and ceilometer measurements”, Atmospheric Research, 196.
- [23] Información obtenida de: <https://cloudatlas.wmo.int/es/descriptions-of-clouds.html>
- [24] Zahnle K, Schaefer L, Fegley B. “Earth’s earliest atmospheres. Cold Spring Harb Perspect Biol”. 2010 Oct.
- [25] Hansen J. et al, “Atmospheric CO2 Concentrations and the Earth’s Climate”, 2007
- [26] Margulis L. y Sagan D., “The Atmosphere and Life on Earth”, Scientific American, 1986
- [27] Sanchez Antuña, J.C., “Configuración y metodología para el uso de cámaras de todo cielo en la obtención de parámetros atmosféricos” Universidad de Valladolid