



---

**Universidad de Valladolid**

**FACULTAD DE CIENCIAS**

**TRABAJO FIN DE GRADO**

**Grado en Física**

**Análisis de imágenes de cielo con una inteligencia artificial**

**Autor: Sergio Sanz Huidobro**

**Tutores: Roberto Román Díez y Juan Carlos Antuña Sánchez**

**Año 2023**



## **Resumen**

En este trabajo se trata la clasificación de las nubes, los equipos de medida para la cubierta de nubes y, como objetivo principal, el análisis de imágenes de cielo a través de un modelo de segmentación semántica con el fin de poder conocer la localización exacta de las nubes en nuestras imágenes. Para ello se ha utilizado U-Net, un tipo de red convolucional ampliamente utilizada para este tipo de tareas de segmentación. Tras el entrenamiento y la validación de este modelo se han obtenido resultados de hasta el 83% de efectividad.

## **Abstract**

This work addresses the classification of clouds, the measurement equipment for cloud cover, and, as the main objective, the analysis of sky images through a semantic segmentation model to accurately determine the location of clouds in our images. U-Net, a widely used convolutional neural network for segmentation tasks, has been employed for this purpose. After training and validating this model, results with up to 83% effectiveness have been obtained.



# Índice

<b>1. Introducción</b>	<b>1</b>
1.1. Definición de nube . . . . .	1
1.2. Clasificación de las nubes . . . . .	1
1.2.1. Clasificación por altura . . . . .	1
1.2.2. Clasificación general . . . . .	2
1.3. Equipos de medida para la cubierta de nubes . . . . .	8
1.3.1. Equipos terrestres . . . . .	9
1.3.2. Equipos satélite . . . . .	10
1.3.3. Cámaras de cielo . . . . .	11
1.4. Objetivos . . . . .	12
<b>2. Instrumentación y metodología</b>	<b>13</b>
2.1. Cámaras de cielo . . . . .	13
2.2. Red neuronal . . . . .	14
2.2.1. Arquitectura de las Convolutional Neural Networks (CNN)	15
2.2.2. Elementos Clave de una CNN . . . . .	16
2.2.3. Arquitectura de U-Net . . . . .	17
2.3. Etiquetado de las imágenes de cielo . . . . .	18
<b>3. Resultados</b>	<b>22</b>
3.1. Código para la validación del modelo . . . . .	22
3.2. Imágenes . . . . .	22
3.3. Matriz de confusión . . . . .	25
<b>4. Conclusiones</b>	<b>28</b>



## Lista de acrónimos

<b>GOA:</b>	Grupo de Óptica atmosférica
<b>OMM:</b>	Organización meteorológica mundial
<b>CCD:</b>	En inglés, dispositivo de carga acoplada
<b>CMOS:</b>	En inglés, semiconductor complementario de óxido metálico
<b>HDF5:</b>	En inglés, formato de datos jerárquicos
<b>CNN:</b>	En inglés, red neuronal convolucional



# 1. Introducción

Las nubes pueden tener un impacto notable en la vida de una persona. Pueden afectar directamente el balance radiativo de la Tierra y pueden generar problemas como tormentas o sequías. Por estas razones, se han estudiado desde los inicios de la humanidad.

En nuestro caso, este trabajo consistirá en el estudio mediante una red neuronal U-Net de imágenes de cielo obtenidas con las cámaras disponibles para el Grupo de Óptica Atmosférica de la Universidad de Valladolid, GOA. Pero en primer lugar definamos qué es una nube y cuáles son las maneras de clasificarlas.

## 1.1. Definición de nube

Según el atlas de nubes de la OMM [3] (Organización Meteorológica Mundial) la definición de nube viene dada como:

*Una nube es un hidrometeoro consistente en diminutas partículas de agua líquida o hielo, o de ambos, suspendidas en la atmósfera y que, por lo general, no tocan el suelo. También puede contener partículas de agua líquida o hielo de mayores dimensiones, así como partículas líquidas no acuosas o partículas sólidas procedentes, por ejemplo, de gases industriales, humo o polvo*

También en el ya mencionado atlas obtenemos que un meteoro es *Un fenómeno observado en la atmósfera o sobre la superficie de la Tierra que consiste en una precipitación, una suspensión o un depósito de partículas líquidas o sólidas, así como en una manifestación óptica o eléctrica*. Por otro lado la característica diferencial de los hidrometeoros respecto al resto de estos es que, en su mayoría, estos están formados por partículas de agua en cualquiera de sus estados.

## 1.2. Clasificación de las nubes

### 1.2.1. Clasificación por altura

Para esta sección, hay unos términos que deberemos tener en cuenta como lo son la altura y altitud de la base de la nube y la extensión vertical de dicha nube. La altura nos dirá la distancia vertical desde el punto de observación hasta la base de la nube, la altitud será la distancia desde el nivel del mar hasta la base de la nube y la extensión vertical será la distancia desde la base hasta el punto más alto de la nube.

Con esto en cuenta, ahora clasificaremos las nubes en tres niveles (alto, medio y bajo) cuyas divisiones no son estables debido a la fluctuación en las altitudes de los niveles atmosféricos según la localización geográfica, pero que siempre estarán entre el nivel del mar y el nivel más alto de la troposfera. Si las dividimos en estos niveles podremos identificar varios tipos de nube en cada nivel:

- Nivel alto: Este nivel es el situado entre los 5 y los 13 km de altitud si nos encontramos en latitudes medias. En estos niveles solemos encontrar cirrus (En la Figura 1 señalado como (Ci), cirrocumulus (Cc) y cirrostratus (Cs).

- Nivel medio: Delimitado por los 2 y los 7 km de altitud en este nivel podremos encontrar altocumulus (Ac), altostratus (As) y nimbostratus (Ns).
- Nivel bajo: Este nivel es aquel que se encuentra entre la superficie de la Tierra y los 2 km de altitud y en ella podemos encontrar stratus (St), stratocumulus (Sc), cumulus (Cu) y cumulonimbus (Cb).

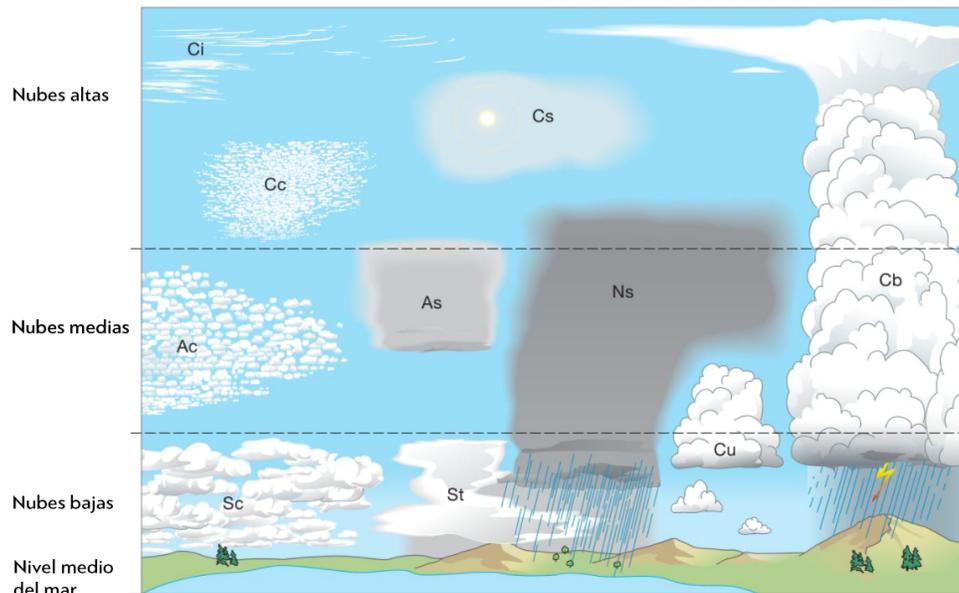


Figura 1: Esquema que representa las distintas nubes a distintos niveles. Fuente: Atlas de las Nubes [3].

En la Figura 1 podemos observar que algunas nubes no se ajustan solamente a la clasificación por altura y que esta presenta las siguientes excepciones: el altostratus es propio del nivel medio, pero con frecuencia se extiende a alturas superiores, el nimbostratus casi siempre se sitúa en el nivel medio, pero suele extenderse a los otros dos niveles; el cumulus y el cumulonimbus tienen habitualmente su base en el nivel bajo, pero a menudo presentan tal extensión vertical que su cima puede alcanzar los niveles medio y alto. Por estas razones, podemos diferenciar las nubes siguiendo otros criterios.

### 1.2.2. Clasificación general

La manera más común de clasificar nubes se basa en la utilización de tres criterios denominados géneros, especies y variedades [3]. Estos criterios están basados en la apariencia visual de las nubes. En primer lugar veamos los géneros, cuyos nombres deberían resultar familiares ya que hemos utilizado estos nombres en el apartado anterior. Se dividen en 10 géneros que podemos ver en la Figura 2:

- Cirrus (Ci): Se trata de una nube filamentada con un color blanco o casi blanco con filamentos finos.

- Cirrocumulus (Cc): Es una capa de nubes delgada y blanca compuesta por pequeñas ondas o granulaciones.
- Cirrostratus (Cs): Velo transparente y blanquecino que cubre parcial o completamente el cielo.
- Altostratus (As): Banda delgada grisácea o azulada de aspecto estriado con ranuras o canales dispuestos paralelamente a la corriente de aire, que cubre el cielo pero tiene partes suficientemente delgadas que permiten observar vagamente el Sol.
- Nimbostratus (Ns): Capa de nubes gris oscura, con un aspecto velado por la continua precipitación de lluvia o nieve, que en muchos de los casos llega al suelo. Esta capa es por todas partes suficiente para no permitir ver en ninguna zona el Sol. Por debajo de un Nimbostratus pueden aparecer nubes bajas que pueden o no ser parte de este.
- Stratocumulus (Sc): Banco, banda delgada o capa de nubes de color blanco o gris, o a la vez blanco y gris, que casi siempre presentan partes oscuras y están compuestos por losetas, masas redondeadas, rodillos, etc., que no son fibrosos (excepto la virga) y que pueden o no estar unidos; la mayoría de los pequeños elementos están distribuidos con regularidad y tienen, por lo general, una anchura aparente superior a 5 grados.
- Stratus (St): Capa de nubes generalmente gris, con una base relativamente uniforme de la que pueden caer llovizna o incluso nieve. Cuando el Sol es visible, se puede distinguir su contorno con claridad. Estas nubes no generan fenómenos de halo excepto a temperaturas muy frías.
- Cumulus (Cu): Nubes separadas, normalmente densas y con el contorno definido, que crecen en la dirección vertical a través de protuberancias, torres o cúpulas que en su parte superior pueden parecerse a una coliflor. Las partes que están iluminadas por el Sol son, en su mayoría, de un blanco brillante; la zona inferior es bastante oscura y prácticamente horizontal.
- Cumulonimbus (Cb): Nube densa, con un potente desarrollo vertical, en forma de montaña o de enorme torre. Al menos, una porción de su área superior exhibe una superficie suave, fibrosa o rayada, generalmente con una forma aplanada. Esta sección tiende a extenderse en una configuración similar a un yunque o un amplio penacho. Debajo de la base de esta nube, que a menudo tiene un aspecto muy oscuro, es común observar nubes bajas desgarradas, conectadas o no a la mencionada nube superior, acompañadas de precipitaciones que a veces se presentan en forma de virga.

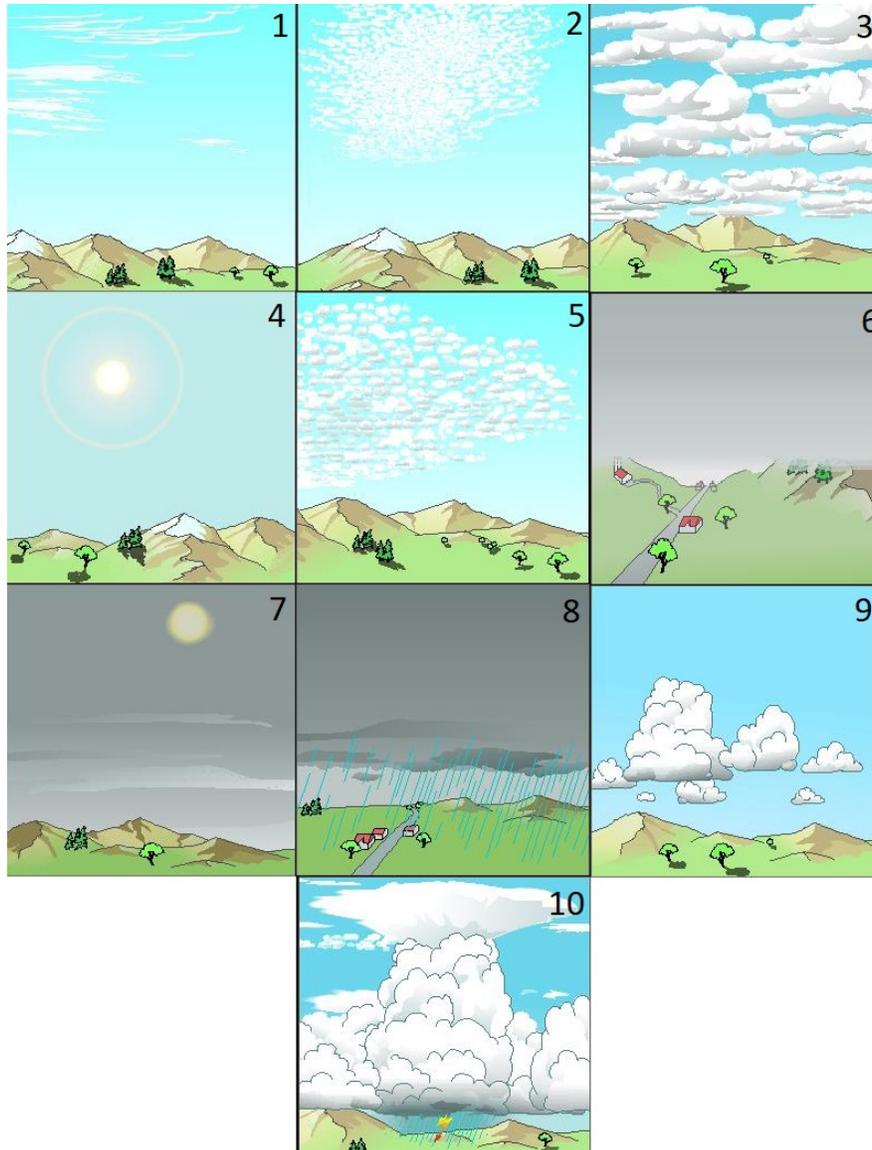


Figura 2: Representación de los distintos géneros de nube: (1) Cirrus, (2) Cirrocumulus, (3) Stratocumulus, (4) Cirrostratus, (5) Altostratus, (6) Stratus, (7) Altostratus, (8) Nimbostratus, (9) Cumulus, (10) Cumulonimbus. Fuente: Atlas de las nubes [3]

Aparte de los géneros, también podemos clasificar más concretamente las nubes ayudándonos de las especies, que subdividen cada género en función de sus formas y estructura interna. Estas especies son excluyentes, es decir, una nube solo puede pertenecer a una especie descartando así las demás. Dividiremos esta clasificación en los distintos géneros y dentro de ellos sus distintas especies:

- Cirrus:

- Uncinus: sin partes grises, a menudo en forma de coma, que acaba en un gancho o penacho que no termina en una protuberancia redondeada.

- Spissatus: su color es gris pueden y tapar el Sol o oscurecerlo. A menudo se forman en la parte superior de un Cumulonimbus.

- Cumulus

- Humilis: Cumulus caracterizado por una extensión vertical pequeña y un aspecto generalmente aplastado.
- Mediocris: Cumulus de tamaño moderado, con pequeños penachos y brotaciones en su cima.
- Congestus: Cumulus que presenta protuberancias muy claras, generalmente de formas muy marcadas y normalmente, de gran tamaño en la dirección vertical. A menudo la parte alta protuberante de un Cumulus congestus se parece a una coliflor.

- Cumulonimbus

- Calvus: Cumulonimbus en el que al menos algunas protuberancias de la parte alta empiezan a perder su forma de cumulo, pero en el que no se aprecian formas típicas de cirrus. Las protuberancias tienden a formar una masa blanquecina que presenta estrías verticales paralelas a la corriente de aire y que reflejan el flujo de aire.
- Capillatus: Cumulonimbus caracterizado por la presencia de zonas claramente cirriformes con una estructura claramente fibrosa o estriada, a veces con forma de yunque, penacho o enorme masa de cabellos. Suele ir acompañado de lluvia, tormenta y a veces granizo; frecuentemente una virga. Podemos observar un ejemplo en la Figura 3b

- Especies que aparecen en distintos géneros

- Floccus: cada elemento nuboso está formado por una pequeña protuberancia de aspecto cumuliforme rasgada en su parte inferior y a veces acompañada de virga. Cirrus, Cirrocumulus, Altocumulus y Stratocumulus
- Castellanus: presentan en su parte alta protuberancias cumuliformes con forma de torretas o torres, algunas conectadas por una base común y parecen estar alineadas. Esto es fácilmente observable si se mira la nube desde el lateral. Cirrus, Cirrocumulus, Altocumulus y Stratocumulus.
- Fibratus: velo de nubes delgado en forma de filamentos que no acaban en gancho. Puede aparecer en Cirrus (ver la Figura 3a) y Cirrostratus.
- Stratiformis: Nube extendida en una capa horizontal de grandes dimensiones. Este término se aplica a los Altocumulus, Stratocumulus y, en ocasiones, a los Cirrocumulus.
- Nebulosus: Nube con forma de velo nebuloso e indefinido, sin claridad en los detalles. Este término se aplica principalmente a los Cirrostratus y Stratus.

- Lenticularis: Nubes con forma de almendra, a menudo alargadas y con detalles nítidos; pueden presentar irisaciones. Estas nubes aparecen normalmente con origen orográfico aunque también aparecen en regiones en las que no hay un relieve marcado. Este término se aplica principalmente a los Cirrocumulus, Altocumulus y Stratocumulus
- Volutus: Masa de nubes largas, habitualmente situadas a baja altura, horizontal, aislada y con forma de tubo que con frecuencia parece girar lentamente sobre un eje horizontal. La nube giratoria es un cuerpo aparte separado de otras nubes. Esta especie aparece principalmente con el Stratocumulus y raramente con el Altocumulus.
- Fractus: Nubes en forma de jirones claramente rasgados. Este término se aplica a los Stratus y a los Cumulus.



Figura 3: (a) Cirrus fibratus radiatus. (b) Cumulonimbus Capillatus

Para concretar aún más en el análisis de nubes utilizamos las variedades, las cuáles describen cómo están dispuestos los elementos visibles a simple vista en una nube y cuánta luz pueden dejar pasar. Es esencial considerar que una nube específica puede mostrar diferentes disposiciones, lo que indica que estas variedades no se excluyen entre sí. Las variedades translucidus y opacus son excepciones a lo mencionado anteriormente, ya que sí son excluyentes. Algunas variedades pueden encontrarse en varios géneros de nubes. La existencia de una clasificación de variedades no implica que una nube en particular deba ser necesariamente etiquetada con una o más de esas variedades.

De manera similar a las categorías anteriores, enumeraremos y definiremos cada variedad.

- Intortus: Se describe como cirrus con filamentos que presentan curvaturas extremadamente irregulares, a menudo dando la apariencia de estar enredados de manera caótica e impredecible
- Vertebratus: Se utiliza para describir la disposición de elementos en ciertas nubes, específicamente en los cirrus, donde los elementos están dispuestos de manera que recuerdan a vértebras, costillas o la estructura ósea de un pez.

- **Undulatus:** Se hace referencia a nubes que se organizan en bancos, bandas estrechas o capas, y muestran patrones de ondulaciones. Estas ondulaciones pueden ser visibles tanto en capas nubosas relativamente uniformes como en nubes compuestas. En ocasiones, se observa un sistema doble de ondulaciones. Este concepto se aplica principalmente a los tipos de nubes Cirrocumulus, Cirrostratus, Altocumulus, Altostratus, Stratocumulus y Stratus.
  
- **Radiatus:** Se refiere a nubes que muestran bandas anchas y paralelas, creando la ilusión de converger hacia un punto en el horizonte debido al efecto de perspectiva. Cuando estas bandas atraviesan todo el cielo, parecen converger hacia dos puntos opuestos en el horizonte llamados “puntos de radiación”. Este término se aplica principalmente a los tipos de nubes Cirrus, Altocumulus, Altostratus, Stratocumulus y Cumulus.
  
- **Lacunus:** Se refiere a nubes dispuestas en bancos, bandas delgadas o capas, generalmente de grosor reducido y que se caracterizan por tener agujeros redondeados distribuidos de manera relativamente uniforme. Muchos de estos agujeros presentan bordes con flecos. La disposición de los elementos nubosos y las áreas claras recuerda a una red o a un panal. Este término se aplica principalmente a los tipos de nubes Cirrocumulus y Altocumulus, y ocasionalmente, aunque de forma muy poco común, a los Stratocumulus
  
- **Duplicatus:** Hace referencia a bancos, bandas delgadas o capas de nubes que se encuentran en niveles ligeramente diferentes en dos o más capas, y en ocasiones estas capas están parcialmente conectadas. Este término se aplica principalmente a los tipos de nubes Cirrus, Cirrostratus, Altocumulus, Altostratus y Stratocumulus.
  
- **Translucidus:** Se refiere a nubes organizadas en bancos, bandas delgadas o capas extensas, en las cuales la mayor parte es lo suficientemente traslúcida como para permitir ver la posición del Sol o la Luna. Este término se aplica a los tipos de nubes Altocumulus, Altostratus, Stratocumulus y Stratus. Se puede ver un ejemplo en la Figura 4a
  
- **Opacus:** Este término describe un banco, una banda delgada o una capa de nubes extensas, donde la mayor parte es tan opaca que oculta completamente al Sol o a la Luna. Se aplica principalmente a los tipos de nubes Altocumulus, Altostratus, Stratocumulus y Stratus. Podemos observarlo en la Figura 4b
  
- **Perlucidus:** Se refiere a un banco, una banda delgada o una capa de nubes extensas con espacios nubosos bien definidos entre sus elementos, aunque estos espacios a veces son muy pequeños. Estos espacios permiten ver claramente el Sol, la Luna, el cielo azul y otras nubes a altitudes superiores. También puede observarse en combinación con las variedades translucidus u opacus. Este término se aplica principalmente a los tipos de nubes Altocumulus y Stratocumulus.



Figura 4: (a) *Altocumulus translucidus lacunosus*. (b) *Altostratus opacus undulatus*

### 1.3. Equipos de medida para la cubierta de nubes

La cubierta de nubes [11] se refiere a la fracción del cielo que está cubierto por nubes en un momento específico. En el estudio de la cubierta nubosa se divide el cielo en ocho octas que se utilizan como unidad de medida, de manera que si el cielo está cubierto por completo hablaríamos de una cubierta de 8 octas mientras que si no se aprecia ninguna nube se trataría de 0 octas.

La cubierta de nubes puede tener un impacto significativo en el clima y en los recursos naturales, como el agua. La cubierta de nubes se evalúa a menudo por cómo afecta la cantidad de luz solar que llega a la superficie de la Tierra, lo que influye en los recursos de energía solar. Los científicos estudian la cobertura nubosa y sus efectos en la radiación en la superficie utilizando diversos métodos y fuentes de datos, como imágenes de satélite y cámaras que observan el cielo.

Esta sección está basada en el artículo de Tapakis y Charlambides [16] en el que se mencionan distintas maneras de realizar estas medidas. En este capítulo se presentan los numerosos instrumentos de medición utilizados para obtener los datos necesarios para determinar las características de las nubes. Los equipos terrestres se utilizan principalmente para medir datos en una ubicación específica, mientras que los satélites se emplean para medir datos sobre continentes. Los equipos terrestres tienen un campo de visión reducido, lo que limita la capacidad de monitorear la formación de nubes y su movimiento sobre un área extensa del cielo. No obstante, dado que el estado del cielo se mide en el lugar, proporcionan una precisión suficiente para las variaciones locales de la irradiación solar debido a las nubes.

Por otro lado, los satélites proporcionan información a gran escala sobre las nubes y medidas multiespectrales de diferentes sensores, aunque los datos proporcionados tienen una baja resolución y pueden contener errores. A menudo, se pasan por alto las nubes pequeñas debido al análisis limitado y las nubes de baja altura o delgadas no se distinguen fácilmente de la superficie terrestre (Dybbroe et al., 2005a, 2005b [6]; Heinle et al., 2010 [8]). El costo del equipo es un factor importante en las mediciones de nubes, ya que el uso de satélites es mucho más caro que la mayoría de los instrumentos terrestres. Sin embargo, algunos sistemas de observación meteorológica terrestre especializados, como los radares (RADARs), pueden ser igualmente costosos que los satélites (Clothiaux et al., 1998 [5]; Wang and Sassen, 2004

[14]).

### 1.3.1. Equipos terrestres

En primer lugar hablaremos de los dispositivos terrestres, que podemos dividir en cuatro categorías. El primero incluye los instrumentos diseñados para medir la irradiancia, como los piranómetros, y también los radiómetros, cuyo propósito es medir la radiación incidente de una determinada longitud de onda. La idea detrás del diseño de estos instrumentos se fundamenta en que todo cuerpo con una temperatura superior al cero absoluto emite radiación electromagnética, y que el espectro de esta radiación se puede utilizar para distinguir diferentes objetos, como por ejemplo el Sol, las nubes, moléculas de aire, tierra, mar, hielo, etc.

El segundo grupo de los equipos de tierra está formado por las cámaras de cielo que fotografían el cielo en el espectro visible, infrarrojo o ultravioleta [12]. En la Figura 5 se puede ver una de las cámaras utilizadas.

El tercer grupo incluye radares que miden la retrodispersión de una señal electromagnética emitida por el propio aparato. Estos instrumentos miden la diferencia de tiempo entre la emisión y recepción de la señal, así como la pérdida de intensidad de la señal, para determinar la posición de las nubes (Feister et al., 2010 [7], Kato et al., 2001 [9]). Los ceilómetros forman parte de este tercer grupo, se utilizan para indicar la altura de la base de una o varias capas de nubes.

El último de estos cuatro grupos es la radiosonda, un dispositivo electrónico incorporado en globos meteorológicos que permite monitorizar las nubes. Un globo meteorológico se llena con helio para que alcance una altitud de más de 30 km. La radiosonda incluye sensores capaces de medir temperatura, presión del aire y humedad relativa. Durante su camino ascendente, la radiosonda detecta nubes aportando información sobre las mismas como la altura de su base, el espesor y la altura de la parte superior de la nube. La mayor ventaja de las radiosondas es que pueden detectar múltiples capas de nubes en cada nivel de la atmósfera. Además, el desplazamiento horizontal del globo indica la velocidad y dirección del viento que mueve las nubes en cada capa de la atmósfera.



Figura 5: Cámara de cielo SONA 202-NF utilizada por el grupo GOA en la Universidad de Valladolid. Fuente: Tesis doctoral de J. C. Antuña [2]

### 1.3.2. Equipos satélite

En el caso de los satélites algunos adoptan una órbita geoestacionaria, orbitando alrededor de la Tierra en sincronización con su rotación y escaneando constantemente una área específica. Por otro lado, algunos siguen una órbita polar, cruzando ambos polos durante cada órbita, o mantienen una órbita heliosíncrona, pasando por la misma ubicación a una hora consistente todos los días. Estas órbitas diversas satisfacen distintos objetivos meteorológicos.

Los satélites meteorológicos tienen como función principal la observación y análisis de las condiciones climáticas en la Tierra. Están equipados con un multirradiómetro llamado “imager”, que incorpora todos los equipos necesarios para medir radiación visible (VIS) pasiva o activa, infrarroja reflejada (IR), infrarroja térmica (IR térmica), ultravioleta (UV) y microondas (MW), entre otras mediciones. Los imágenes utilizan varios canales radiométricos, empleando sensores diferentes para cada canal con el fin de medir datos en un rango específico de longitudes de onda (Weng, 2011). Los avances en tecnología de teledetección han llevado al desarrollo de mejores imágenes, con más canales espectrales, mayor resolución y mayor sensibilidad. Como resultado, las nuevas versiones de satélites que incorporan imágenes de nueva generación pueden manejar una gama más amplia de datos, adquiridos de varios tipos de imágenes satelitales. Tsuchiya y Tokuno [17] (1992) y Li et al. (2007) presentaron una comparación de diferentes imágenes.

Los datos del canal visible (VIS) básicamente se asemejan a una fotografía de una área específica de la Tierra, capturando la luz solar reflejada desde la superficie superior de las nubes. Las nubes densas aparecen más brillantes en las fotografías porque tienen una reflectividad más alta que las nubes delgadas. Por otro lado, los canales de infrarrojo (IR) y ultravioleta (UV) se utilizan para medir la radiación IR o UV emitida y absorbida de la misma área de la Tierra. En la Figura 6 podemos observar una imagen obtenida a partir de un satélite en Baja California.

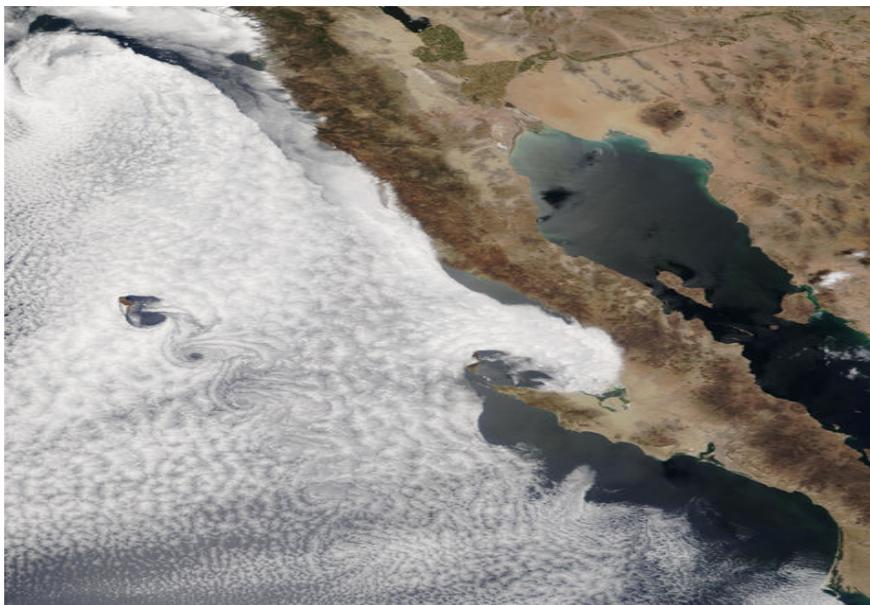


Figura 6: Imagen de satélite de Camino de vórtices cerca de Baja California.  
Fuente: Jacques Descloitres, MODIS Land Rapid Response Team, NASA/GSFC [19]

### 1.3.3. Cámaras de cielo

En nuestro caso nos centraremos en las cámaras de cielo, ya que son el instrumento de medida utilizado en este trabajo. Las cámaras espectrales en el rango visible clasifican las nubes según las diferencias en intensidad de color entre los diversos tipos de nubes y el cielo.

Estas cámaras capturan automáticamente imágenes en intervalos de tiempo establecidos, variando según el día y la noche, según el estudio en cuestión. Estas cámaras permiten estudiar la cobertura de nubes en una vista de  $180^\circ$ . Una cámara típica de cielo consta de dos componentes: una caja resistente a la intemperie con la óptica ensamblada y un conjunto de control de la computadora. Dentro del ensamblaje óptico se ubica un sensor CCD o CMOS apuntando hacia arriba. Una disposición común de una cámara de cielo incluye la cámara y una lente ojo de pez colocada en la parte superior para capturar imágenes amplias del cielo.

Estas cámaras se emplean para analizar y clasificar la cobertura de nubes en función de sus características de color e intensidad. La captura de imágenes se realiza de manera automática y se programa en intervalos regulares, pudiendo variar según el día o la noche, dependiendo del estudio en cuestión. Este tipo de cámaras proporcionan una visión completa de  $180^\circ$  del cielo y permiten el estudio detallado de la cobertura nubosa. En la Figura 7 podemos observar un ejemplo de estas imágenes.

En cuanto a la estructura de la cámara, consta de dos elementos principales. En primer lugar, está la carcasa resistente a las inclemencias del tiempo que alberga la óptica y, en segundo lugar, el conjunto de control de la computadora. El ensamblaje óptico incluye un sensor CCD o CMOS que apunta hacia arriba para capturar las imágenes del cielo.

Para obtener datos y análisis precisos sobre la nubosidad que cubre una amplia vista del cielo, estas cámaras son herramientas esenciales. Proporcionan información valiosa para comprender la atmósfera y su comportamiento en diferentes condiciones y momentos del día.



Figura 7: Imagen capturada por una cámara de cielo. Fuente: Grupo GOA Universidad de Valladolid

## 1.4. Objetivos

Como ya hemos mencionado en las secciones anteriores, el estudio de las nubes ha sido uno de los objetivos perseguidos por los humanos desde la Antigüedad. En nuestro caso, seguiremos con esta labor centrándonos en el estudio de la cubierta de nubes a través de una inteligencia artificial. Si bien este tema ya ha sido tratado en otros trabajos ([1], [10]), en nuestro caso usaremos una peculiaridad respecto a estos. Mientras el enfoque de estos trabajos era utilizar redes neuronales para calcular cuántas octas del cielo están cubiertas, nuestro trabajo tendrá como objetivo principal poder localizar las nubes dentro de cada imagen de cielo y no solamente calcular qué proporción de la imagen se encuentra cubierta por nubes.

Otros objetivos de este trabajo han sido:

- Profundizar en el conocimiento de las nubes y su clasificación.
- Dar a conocer los distintos equipos de medida para la cubierta nubosa y una explicación física de su funcionamiento.
- Trabajar con CNN en general y U-Net en particular como herramienta para analizar la cubierta de nubes y aprender más sobre su funcionamiento y características.
- Incidir en el uso de este tipo de redes neuronales para la segmentación semántica. Esta herramienta se puede utilizar también para otros fines.

## 2. Instrumentación y metodología

### 2.1. Cámaras de cielo

Esta sección describe varias estaciones de medida utilizadas en este trabajo, destacando su ubicación y el equipo instrumental instalado en ellas. A continuación, se detallan las principales estaciones utilizadas:

- Estación de Valladolid (41.66°N, 4.70°O, 705m s.n.m.), España: Equipamiento: Se mencionan cámaras SONA, modelos SONA202-NF y OMEA-3C. Información adicional: Esta estación se utiliza como sitio instrumental de referencia para el Grupo de Óptica Atmosférica de la Universidad de Valladolid (GOA-UVa).
- Estación de Fuencaliente en La Palma (28.48°N, 17.84°O, 630m s.n.m.), España: Equipamiento: Cámara OMEA-3C. Información adicional: Fue instalada para monitorizar aerosoles generados por la erupción del volcán Tajogaite en La Palma.
- Estación de Izaña (28.30°N, 16.49°O, 2400m s.n.m.), España: Equipamiento: Cámara OMEA-3C. Información adicional: Esta estación pertenece al Centro de Investigación Atmosférica de Izaña y se encuentra por encima de una capa de inversión de temperatura.

Estas estaciones están equipadas con diferentes modelos de cámaras de cielo y han sido fundamentales para la recolección de datos necesarios para el desarrollo de la investigación. Además, se menciona el esfuerzo logístico y personal realizado para instalar y mantener estos instrumentos en ubicaciones a menudo desafiantes. En los siguientes párrafos se proporcionan más detalles sobre los instrumentos utilizados en la estación de Valladolid.

En julio de 2018, se instaló el modelo de cámara de cielo SONA202-NF fabricado por Sieltec Canarias S.L [12]. Este dispositivo está equipado con un sensor CMOS SONY IMX249, que ofrece una resolución de imagen de 1172 x 1158 píxeles, equivalente a 2.35 megapíxeles, y una profundidad de 10 bits. El sensor emplea un mosaico de filtros Bayer, junto con un filtro tri-banda que reduce el solapamiento entre los canales de color, aunque no lo elimina por completo.

Esta cámara de cielo está equipada con una lente de ojo de pez que proporciona una visión de 185°, lo que le permite capturar imágenes de todo el cielo. Todo el sistema está encapsulado en una carcasa resistente a las condiciones meteorológicas, con una cúpula de cristal transparente para su protección.

También se trabaja con una cámara OMEA-3C que está equipada con un sensor IMX178 de SONY, que tiene una resolución de imagen de 3096 x 2080 píxeles, una resolución de 6.44 megapíxeles y ofrece una profundidad de color de 14 bits. Este sensor utiliza un mosaico Bayer con un patrón RGGB y se le ha añadido un filtro que bloquea la radiación infrarroja. Las respuestas espectrales de estos filtros se pueden observar en la Figura 2.7.

Esta cámara utiliza una lente ojo de pez y también cuenta con un sensor externo para medir la temperatura y la humedad. Además, posee un sistema de calefacción interna diseñado para prevenir la formación de condensación en el interior de la cámara y eliminar rápidamente las gotas de agua en caso de lluvia o rocío.

Para su funcionamiento, estas cámaras necesitan estar conectadas a un ordenador. Se ha desarrollado una aplicación específica para capturar imágenes y comprimir los datos cuyo funcionamiento explicaremos a continuación.:

- El modelo SONA202-NF de cámara de cielo viene con un ordenador integrado que permite su gestión y control a través de una interfaz web. A través de esta interfaz, los usuarios tienen la capacidad de configurar varias opciones, incluyendo la captura de imágenes en formato RAW y la definición de una secuencia de tiempos de exposición específica.
- El modelo OMEA-3C utiliza la aplicación GOA OMEA Capture para capturar y comprimir imágenes, así como para enviar datos. Las imágenes capturadas se almacenan en un archivo H5 junto con toda la información relevante proporcionada por la cámara. Esta aplicación también ofrece la capacidad de definir tiempos de exposición en función de diferentes intervalos de ángulo cenital solar (SZA), lo que ayuda a evitar la infraexposición durante el amanecer y el atardecer. Para la noche, es posible configurar diferentes tiempos de exposición en dos escenarios: con o sin luna, lo que genera distintas secuencias en cada caso. La aplicación permite seleccionar los intervalos de tiempo entre las secuencias de múltiples exposiciones.

En general, todas las cámaras están configuradas para capturar secuencias cada 5 minutos durante el día y cada 2 minutos durante la noche.

Este control flexible facilita la adaptación de la cámara según las necesidades de la investigación y las condiciones atmosféricas en un momento dado.

Las imágenes se registran en formato RAW para obtener una sensibilidad mayor. Sin embargo, este formato genera un volumen considerable de datos para organizar y optimizar la gestión de estos datos, se emplea el formato de archivos HDF5 abreviado como H5. Este enfoque es fundamental para asegurar que el volumen masivo de datos se mantenga organizado y que se pueda acceder a ellos de manera eficiente.

## 2.2. Red neuronal

Una vez tenemos todas las imágenes almacenadas, trabajaremos con ellas a través de una red neuronal que sigue el modelo de una red U-Net. Esta red será aplicada en nuestro caso para realizar una tarea de segmentación semántica, que es una técnica esencial de dividir una imagen digital en segmentos o regiones, asignando a cada píxel una etiqueta o categoría específica. Este proceso posibilita la identificación y diferenciación precisa de varios objetos y características en la imagen.

Este enfoque, basado en algoritmos de aprendizaje profundo, tiene aplicaciones variadas que abarcan desde la conducción autónoma hasta la generación de imágenes médicas e inspección industrial. Por ejemplo, en conducción autónoma, la segmentación semántica es fundamental para distinguir el camino transitable de los obstáculos, como peatones, aceras, postes y otros vehículos, facilitando así la toma de decisiones del vehículo autónomo. (Ver Figura 8).



Figura 8: Ejemplo de segmentación semántica. [15]

Es relevante destacar que la segmentación semántica no está limitada a solo dos categorías. Es posible ajustar el número de categorías para clasificar el contenido de la imagen, permitiendo una clasificación más detallada. Por ejemplo, una imagen podría dividirse en clases como nube, cielo despejado, y Sol. Además, la segmentación semántica representa una alternativa valiosa a la detección de objetos, ya que permite que el objeto de interés abarque diversas áreas de la imagen a nivel de píxel. Esta técnica es especialmente eficaz para detectar objetos con formas irregulares, en contraste con la detección de objetos, que se basa en la delimitación mediante un cuadro.

La implementación de la segmentación semántica se apoya en redes neuronales convolucionales (CNN), reconocidas por su versatilidad y capacidad para obtener resultados destacados en aplicaciones de visión por ordenador. Estas redes se entrenan para procesar imágenes y, según la cantidad de fotogramas a procesar por segundo, pueden aplicarse también a ciertos tipos de videos en tiempo real. En conclusión, la segmentación semántica representa una técnica poderosa y flexible en el ámbito de la visión por computadora, permitiendo la identificación detallada y la clasificación de elementos en una imagen a nivel de píxel. Su aplicación en diversos sectores y campos de investigación sigue expandiéndose al evolucionar y perfeccionarse las técnicas de aprendizaje profundo.

### 2.2.1. Arquitectura de las Convolutional Neural Networks (CNN)

Las CNN vienen definidas por la palabra convolución, una operación matemática representada de forma general de esta manera:

$$F(t) = (f * g)(t) = \int f(\tau)g(t - \tau)d\tau$$

En la terminología de las redes neuronales convolucionales, el primer argumento de la convolución ( $f$  en la ecuación) sería la entrada o input. Su segundo componente sería el núcleo o kernel ( $g$ ) y la función de salida ( $F$ ), que genera el mapa de características o feature map. En la mayoría de los casos de machine learning<sup>1</sup>, la entrada es un vector unidimensional de datos y el núcleo es un conjunto de parámetros.

<sup>1</sup>El machine learning es el término inglés para aprendizaje automático una rama de la Inteligencia Artificial que permite a los ordenadores aprender a tomar decisiones basadas en datos

Antes de adentrarnos en la estructura de las CNN, debemos llamar la atención sobre dos ventajas cruciales de estas sobre el resto de redes neuronales:

- **Conexiones dispersas:** A diferencia de las redes neuronales tradicionales, donde cada neurona está conectada con todas las neuronas de las capas previas y posteriores, las CNN emplean un núcleo considerablemente más pequeño que la entrada. Esto posibilita que sus neuronas trabajen con conexiones dispersas, es decir, están conectadas solo a un subconjunto de la capa posterior y también solo a un subconjunto de la capa anterior. Así, el resto de píxeles de la imagen no influye en la salida. La capa que se conecta con una neurona de la capa  $n + 1$ , se denomina campo receptivo local.
- **Compartición de parámetros:** En una misma capa, todas las neuronas ocultas comparten los mismos pesos y sesgos. Por lo tanto, todas las neuronas de la capa oculta detectan la misma característica, aunque lo hacen en diferentes partes o ubicaciones de la imagen de entrada.

Estas ventajas hacen de las CNN una elección ideal para la identificación y clasificación de imágenes. La compartición de parámetros permite que la red sea invariante ante movimientos de traslación en la entrada. Por otro lado, gracias a las conexiones dispersas, las neuronas pueden concentrarse y especializarse en un pequeño rectángulo de la imagen de entrada, lo que les permite detectar de manera efectiva las características de bajo nivel en las primeras capas ocultas, antes de fusionarlas y enfocarse progresivamente en niveles más altos de abstracción.

Esta estructura de las CNN, centrada en características locales y con la capacidad de aprender jerarquías complejas de características, las ha posicionado como una herramienta poderosa en el procesamiento de imágenes y vídeos. Su capacidad para reconocer patrones y características relevantes en una imagen, aprendiendo automáticamente a través del entrenamiento, las convierte en una elección popular para aplicaciones como detección de objetos, segmentación de imágenes, reconocimiento facial, entre otras.

### 2.2.2. Elementos Clave de una CNN

- **Neurona:** Es la unidad fundamental en una CNN. Aplica una función a los valores de sus entradas provenientes de las conexiones con otras neuronas, generando un nuevo valor que será procesado por lo que se denomina función de activación, generando así el valor de salida.
- **Filtro o kernel:** Este elemento corresponde a la matriz de pesos de una capa de convolución integrada en la CNN. El núcleo siempre tiene los mismos canales que la entrada, pero con un tamaño mucho menor. Normalmente, la entrada es una imagen en escala de grises o en color.

Normalmente, se emplea un número elevado de núcleos en las CNN destinadas a la clasificación de imágenes.

Estos elementos conforman la base de las CNN, permitiendo un procesamiento efectivo de imágenes y la extracción de características relevantes para su clasificación.

Las CNN constan de múltiples capas de filtros convolucionales, que se utilizan para extraer características de los datos de entrada. Hay tres tipos de capas:

- Capas convolucionales: Estas capas aplican filtros convolucionales a los datos de entrada para extraer características.
- Capas de agrupación (pooling): Estas capas reducen las dimensiones espaciales de los mapas de características para introducir invariancia translacional.
- Capas completamente conectadas: Estas capas conectan cada neurona en una capa con cada neurona en la siguiente capa y se utilizan para clasificar los datos de entrada en función de las características extraídas

### 2.2.3. Arquitectura de U-Net

El deep learning o aprendizaje profundo se enfoca en el entrenamiento de redes neuronales artificiales con múltiples capas, lo que se conoce como redes neuronales profundas. Estas redes están diseñadas para aprender y representar patrones y características complejas a partir de datos, de manera similar a cómo funciona el cerebro humano, pero en un entorno digital.

U-Net es una arquitectura de deep learning ampliamente utilizada que fue diseñada para aprovechar de manera efectiva una cantidad menor de datos, manteniendo al mismo tiempo la velocidad y precisión. Estas redes tienen este nombre debido a la forma de U que tiene su representación esquemática (ver Figura 9).

Se trata de una red neuronal del tipo CNN, o Red Neuronal Convolucional, que son un tipo de fundamental de red neuronal artificial que se utiliza principalmente para aplicaciones de visión por ordenador, como el reconocimiento de imágenes o vídeos.

En el ámbito del Aprendizaje Profundo, la obtención de conjuntos extensos de datos es fundamental para entrenar modelos eficaces. Sin embargo, recopilar volúmenes considerables de datos para abordar un problema de clasificación de imágenes puede presentar desafíos significativos en términos de tiempo, costos y recursos de hardware.

Además, el proceso de etiquetado de estos datos implica un conocimiento experto y la participación de varios desarrolladores, aquí es donde U-NET emerge como una solución para estos obstáculos. Incluso con un conjunto de datos limitado, U-NET funciona de manera eficaz y supera en precisión a los modelos convencionales.

En comparación con la arquitectura estándar de los autocodificadores, que comprime la información de entrada y luego la expande, U-NET adopta un enfoque más sofisticado. Evita la pérdida de características clave al realizar la expansión en la etapa de decodificación, en contraste con la compresión lineal utilizada por los autocodificadores.

La arquitectura de U-Net es única en el sentido de que consta de una ruta de contracción (contraction path) y una ruta de expansión (expansion path). La ruta de contracción contiene capas de codificación (encoder) que capturan información contextual y reducen la resolución espacial de la entrada, mientras que la ruta de expansión contiene capas de decodificación (decoder) que decodifican los datos codificados y utilizan la información de la ruta de contracción a través de conexiones de salto (skip connections) para generar un mapa de segmentación.

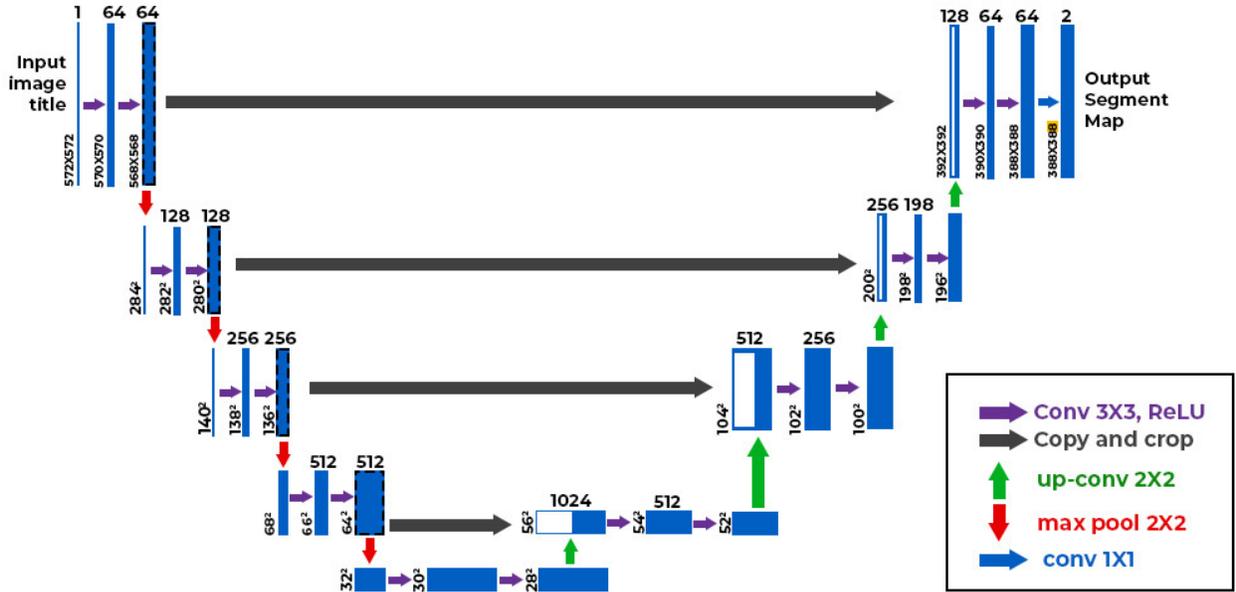


Figura 9: Esquema de una red U-Net que convierte una imagen de dimensiones 572x572x1 en escala de grises en una imagen de output en binario y segmentada de tamaño 388x388x2. Fuente: GeeksforGeeks

La ruta de contracción en U-Net es responsable de identificar las características relevantes en la imagen de entrada. Las capas de codificación realizan operaciones convolucionales que reducen la resolución espacial de los mapas de características al tiempo que aumentan su profundidad, capturando representaciones cada vez más abstractas de la entrada. Esta ruta de contracción es similar a las capas de avance en otras redes neuronales convolucionales.

Luego, la ruta de expansión toma el mapa de características del cuello de botella y lo convierte nuevamente en una imagen del mismo tamaño que la entrada original. Esto se logra utilizando capas de aumento de tamaño, que aumentan la resolución espacial del mapa de características mientras reducen el número de canales. Las conexiones de salto desde la ruta de contracción se utilizan para ayudar a las capas de decodificación a ubicar y refinar las características en la imagen. Finalmente, cada píxel en la imagen de salida representa una etiqueta que corresponde a un objeto o clase particular en la imagen de entrada. En resumen, las capas de decodificación en la ruta de expansión aumentan el tamaño de los mapas de características, al tiempo que realizan operaciones convolucionales. Las conexiones de salto desde la ruta de contracción ayudan a preservar la información espacial perdida en la ruta de contracción, lo que ayuda a las capas de decodificación a ubicar las características de manera más precisa.

Esta arquitectura distintiva en forma de U de U-NET marca la diferencia, ya que conserva características esenciales y ofrece resultados sobresalientes en la clasificación de imágenes.

### 2.3. Etiquetado de las imágenes de cielo

Para el etiquetado de las imágenes nos hemos conectado al servidor remoto del Grupo de Óptica Atmosférica para poder trabajar en un servidor conjunto y no tener que lidiar con

el inconveniente de compartir la información a través de otro software. La conexión a este servidor se ha realizado a través de Mobaxterm.

El etiquetado ha sido manual, a través de una aplicación desarrollada por el propio GOA-UVa para este fin. Se han etiquetado 226 imágenes, de manera que después de seleccionar todos los píxeles cubiertos por nubes, se genera una máscara que, idealmente, cubrirá todas las partes del cielo que estén cubiertas por estas y dejará sin seleccionar los claros en la cubierta nubosa. Esto se hará de la siguiente forma:

- En primer lugar (ver Figura 10) obtendremos la imagen y una máscara simplemente dividida en cielo claro (azul) y no-cielo (negro) como se ve en la siguiente imagen. Podemos ver que la parte denominada no-cielo es un marco predeterminado que ya viene dado por defecto, luego solo deberemos dividir nuestra imagen entre nubes y cielo claro.

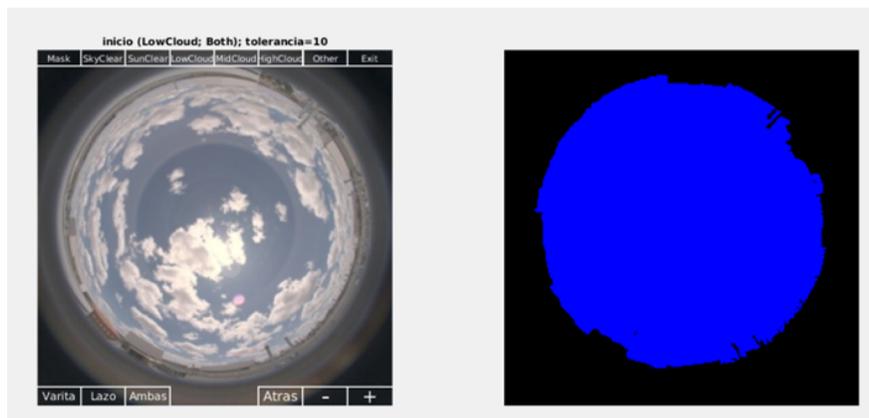


Figura 10: Captura del proceso de etiquetado manual 1.

- A continuación seleccionaremos una sección de nubes con nuestro cursor. Si seleccionamos la opción “varita” el programa nos seleccionará todos los píxeles con un color similar a los seleccionados como se aprecia en la Figura 11. Si en cambio elegimos la opción “lazo” el programa nos permitirá seleccionar regiones enteras teniendo en cuenta nuestra selección y uniendo con una línea recta nuestros primer y último puntos seleccionados. En nuestro caso utilizaremos la opción “ambas” que nos permite utilizar el lazo como si fuera una varita, es decir, seleccionaremos una región y todos los píxeles con colores de esa región serán seleccionados por el programa.

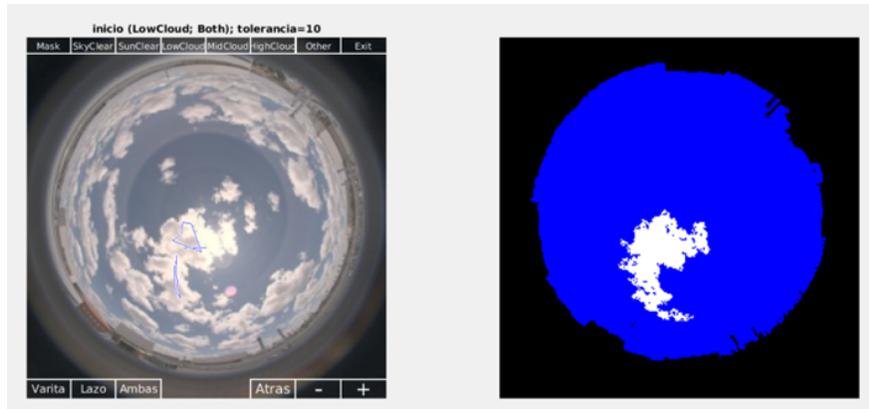


Figura 11: Captura del proceso de etiquetado manual 2.

- Para seguir con nuestra segmentación, aumentaremos nuestra tolerancia, ya que al hacerlo nuestro programa aceptará los colores similares al seleccionado aceptando puntos con un color ligeramente distinto, es decir, podremos seleccionar con mayor facilidad las secciones cubiertas con el riesgo de seleccionar secciones de cielo sin nube que tengan un color más claro (ver Figura 12).

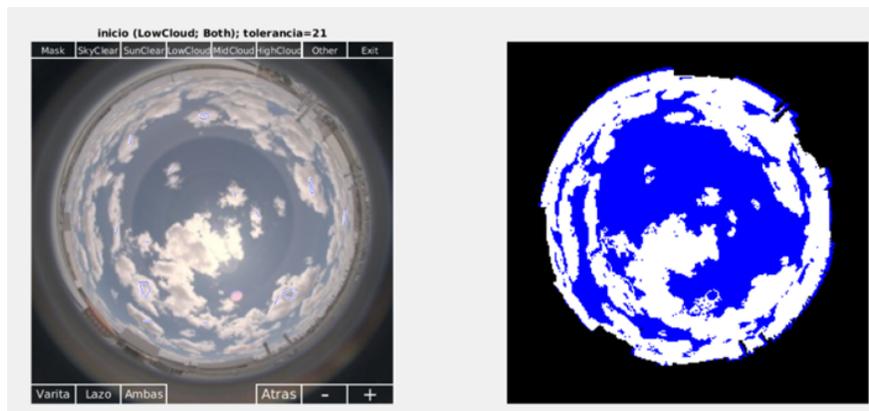


Figura 12: Captura del proceso de etiquetado manual 3.

- Finalmente, haciendo clic en el botón EXIT nos dará la opción de guardar la máscara al elegir entre YES o NO como se ve en la Figura 13. Si la máscara obtenida es de nuestro agrado, la guardaremos, ya que podremos utilizarla en nuestra investigación. Una vez seleccionado esto, aparecerá nuestra siguiente imagen y habrá que repetir el proceso. En caso de que la máscara no sea de nuestro agrado, podríamos comenzar de nuevo seleccionando la opción “SkyClear” en toda la imagen y realizar de nuevo el proceso anterior o, en su defecto, no guardar nuestra máscara y continuar con la siguiente imagen.

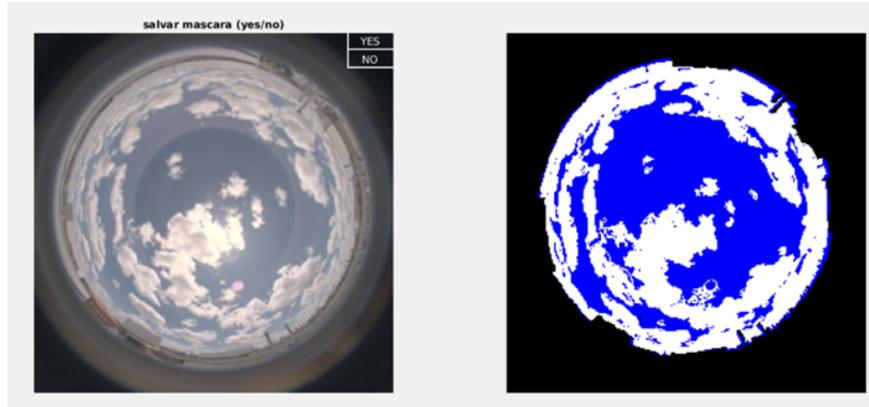


Figura 13: Captura del proceso de etiquetado manual 4.

En este caso podemos observar que la selección de nubes y claros es sencilla, ya que aunque la nube es alta, la iluminación es suficientemente buena y los límites entre nube y cielo son claros. Sin embargo, debido al volumen de medidas realizadas y a los distintos tipos de nubes, aparecerá en nuestras máscaras un error humano no despreciable. Estos errores humanos afectarán a la medida de precisión de nuestra red, ya que no estaremos comparando los resultados finales con imágenes perfectas, sino con unas que presentarán error.

Una vez obtenidas las máscaras manualmente, se ha utilizado el 75% de ellas como imágenes de entrenamiento del modelo U-Net mientras que el porcentaje restante se ha utilizado como máscaras de validación. Esto es, el modelo U-Net ha recibido la información de lo que es y no es una nube a partir de las máscaras realizadas manualmente para la carpeta de entrenamiento y posteriormente ha aplicado este conocimiento para realizar el mismo tipo de diferenciación en las imágenes de la carpeta de validación.

El principal objetivo de esta división es la comparación posterior de las máscaras generadas manualmente para las imágenes situadas en la carpeta de validación y las generadas por el modelo para las mismas imágenes.

## 3. Resultados

### 3.1. Código para la validación del modelo

Para desarrollar el código necesario en este trabajo se ha utilizado el programa Anaconda Spyder utilizando el lenguaje de programación Python. Anaconda es una empresa formada en 2012 con la intención de simplificar el uso del lenguaje Python en el creciente mundo del análisis de datos y negocios [18]. La parte de Spyder nos ofrece una interfaz intuitiva (ver Figura 14) en la que poder programar y ejecutar programas sin necesidad de utilizar una máquina virtual o dejar de trabajar en Windows.

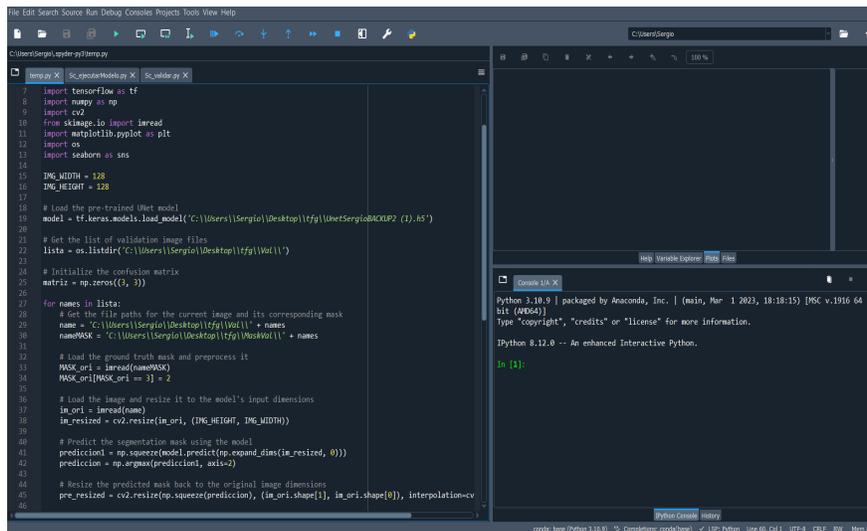


Figura 14: Interfaz de Anaconda Spyder. En la izquierda observamos el código, en la parte superior de la interfaz aparecen las múltiples tareas a realizar y la parte de la derecha es la que representa los outputs resultantes.

El código de nuestro programa busca en primer lugar redimensionar la imagen para que sea compatible con el modelo creado. A continuación obtendrá la máscara de segmentación realizada de manera manual en primer lugar, después hará lo propio con la máscara generada por el modelo. Esta nueva máscara será modificada de una escala de grises (ya que es lo que obtenemos del modelo identificando cada valor con una de las condiciones de cielo) a una en color, para poder apreciar las diferentes zonas con mayor claridad.

Finalmente nuestro programa comparará ambas máscaras y nos dará unos resultados de coincidencia concretos. Para evaluar cómo de válidos son nuestros resultados con este modelo tratando de eliminar el ya mencionado error humano. Esta comparación se realizará con todas las máscaras de la carpeta validación mediante un bucle para así tener una mayor muestra de datos y resultados y por lo tanto, reducir así el error.

### 3.2. Imágenes

Una vez ejecutamos el código, obtendremos tres imágenes. En primer lugar aparecerá en la interfaz de “output” la imagen de cielo original. A continuación se obtiene la máscara

automática generada por nuestro modelo y por último representará la segmentación realizada manualmente sobre esa imagen de cielo. A continuación se pueden ver varios ejemplos:

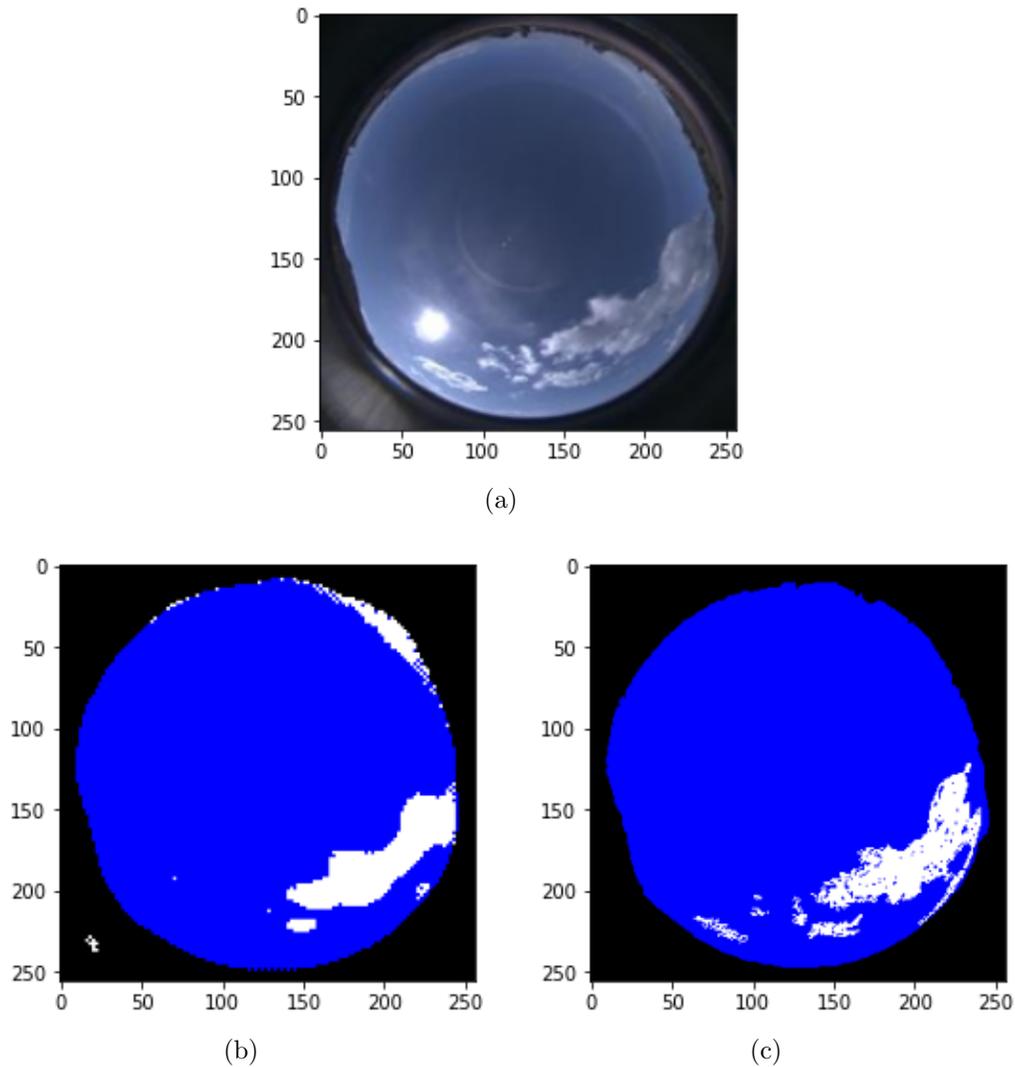


Figura 15: Output del programa número 1: a) Imagen de cielo original, b) Máscara generada por el modelo, c) Máscara manual.

En este caso (Figura 15) podemos observar que las nubes están bien diferenciadas del cielo claro y que si bien la máscara automática puede sobreestimar el tamaño de las nubes ligeramente, la máscara generada manualmente tiene una mayor precisión en nubes pequeñas pero no marca algunas nubes que nuestro modelo sí.

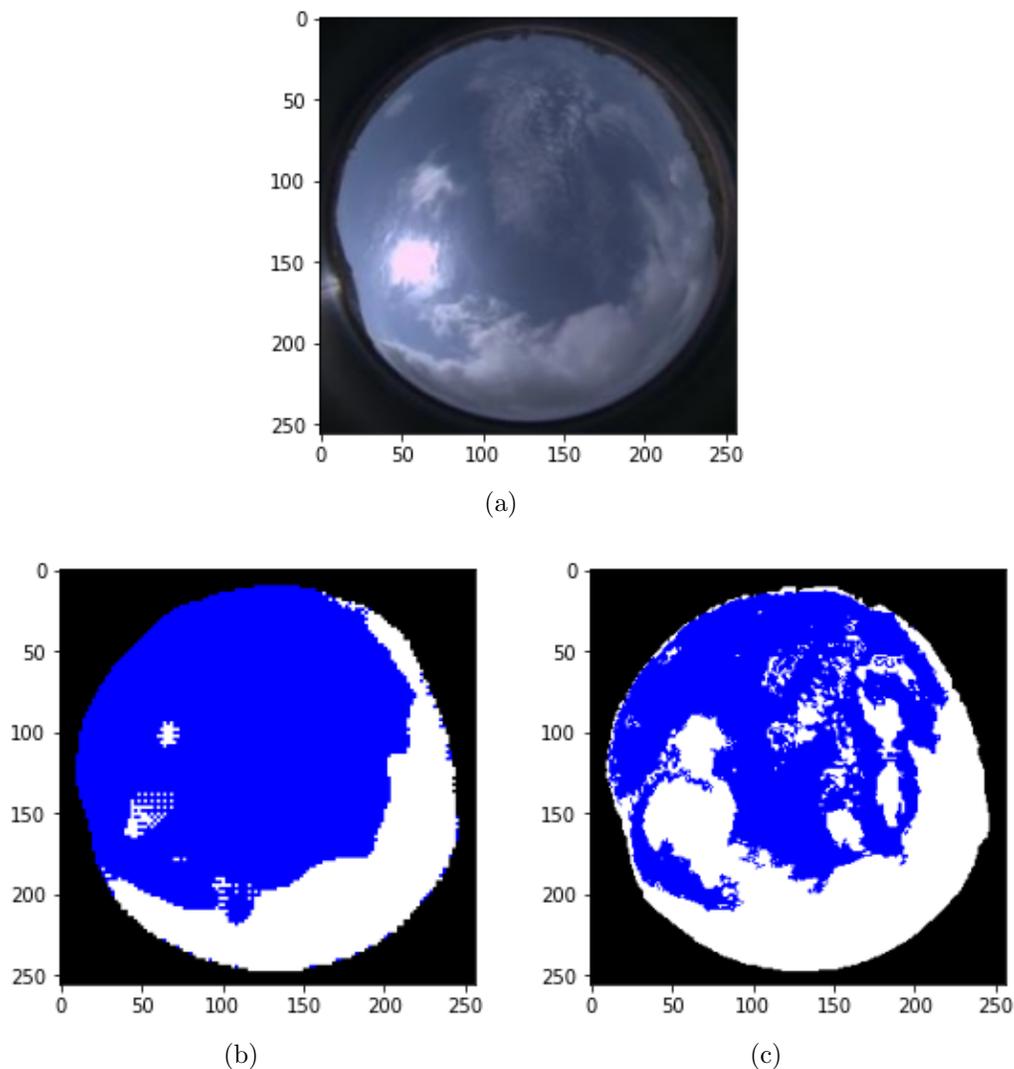


Figura 16: Output del programa número 2. a) Imagen de cielo original, b) Máscara generada por el modelo, c) Máscara manual.

En la Figura 16 se puede observar un caso en el que la máscara manual está sobreestimada mientras que la automática se ajusta de mejor manera a las nubes de la imagen original. Esto se debe a que en la realización de las máscaras manuales se buscó hacer una selección en la que todas las nubes estuvieran seleccionadas completamente para un mejor entrenamiento de la red neuronal en lugar de buscar una precisión exacta. Esto se verá reflejado en el resultado final, ya que la precisión de nuestro modelo no será perfecta.

Finalmente en la Figura 17 observamos uno de los últimos factores a tener en cuenta. A pesar de que ambas máscaras sean similares, se puede observar que en el caso de la generada por la red neuronal los bordes están menos delimitados. Esto se debe a que nuestra red neuronal identificará las zonas cubiertas por una nube con cualquier característica mientras que el ojo humano es capaz de determinar mejor los límites de las nubes y los detalles de la imagen.

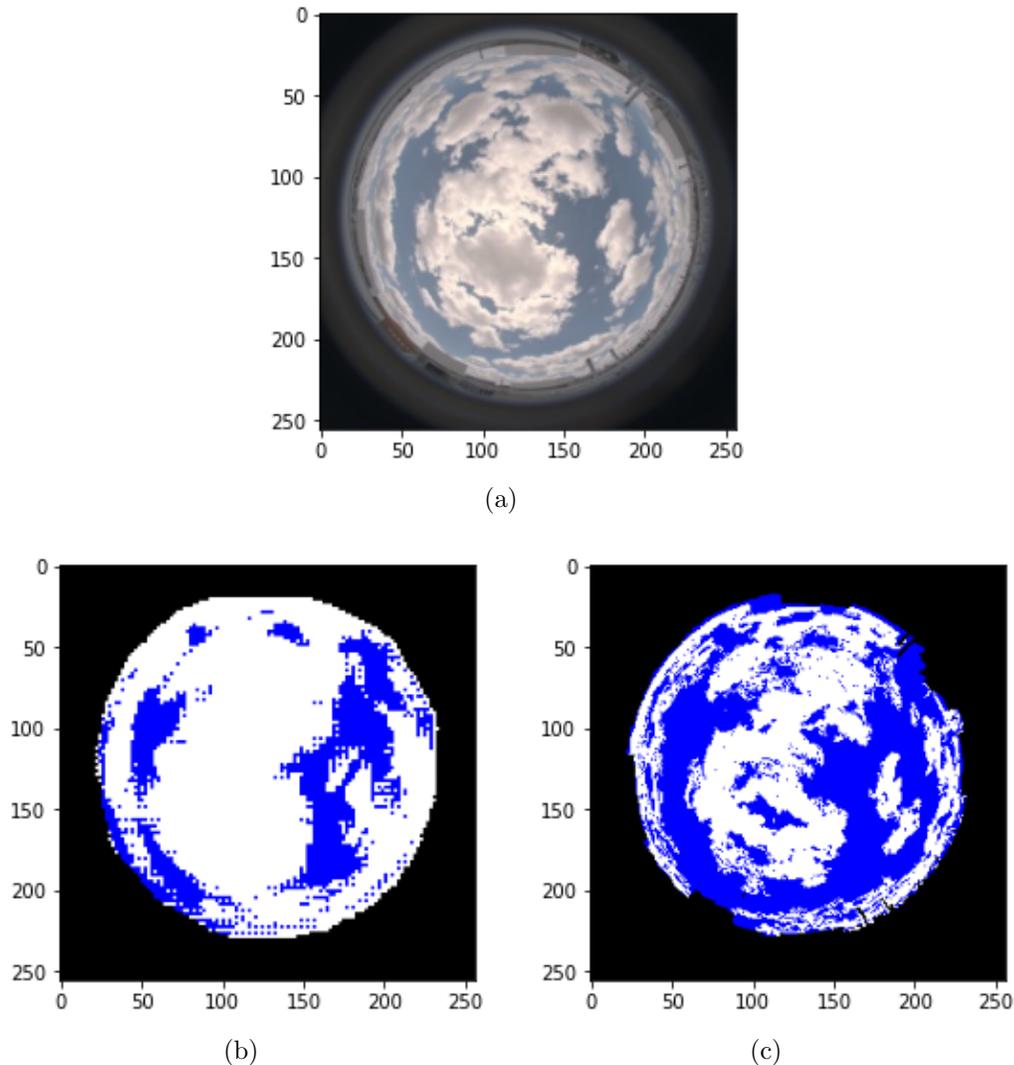


Figura 17: Output del programa número 3. a) Imagen de cielo original, b) Máscara generada por el modelo, c) Máscara manual.

### 3.3. Matriz de confusión

La matriz de confusión es una herramienta fundamental en la evaluación de modelos de clasificación en inteligencia artificial y machine learning. Con el fin de obtener una representación gráfica de nuestros resultados, hemos utilizado esta herramienta. Permite visualizar el desempeño de un modelo al comparar las predicciones hechas por el modelo con las etiquetas reales de los datos.

Se representa en forma de tabla con filas y columnas, donde cada entrada en la posición  $(i, j)$  de la matriz representa la cantidad de ejemplos que pertenecen a la clase  $i$  y fueron etiquetados por el modelo como clase  $j$ .

Los elementos en la diagonal principal de la matriz representan las predicciones correctas del modelo, mientras que los elementos fuera de la diagonal principal representan las predicciones incorrectas.

Esta matriz representa los datos obtenidos al comparar píxel a píxel todas las máscaras realizadas por nuestra U-Net a las máscaras generadas manualmente para todas las imágenes de valoración, es decir, aproximadamente el 25 % de las 226 imágenes utilizadas en nuestro trabajo.

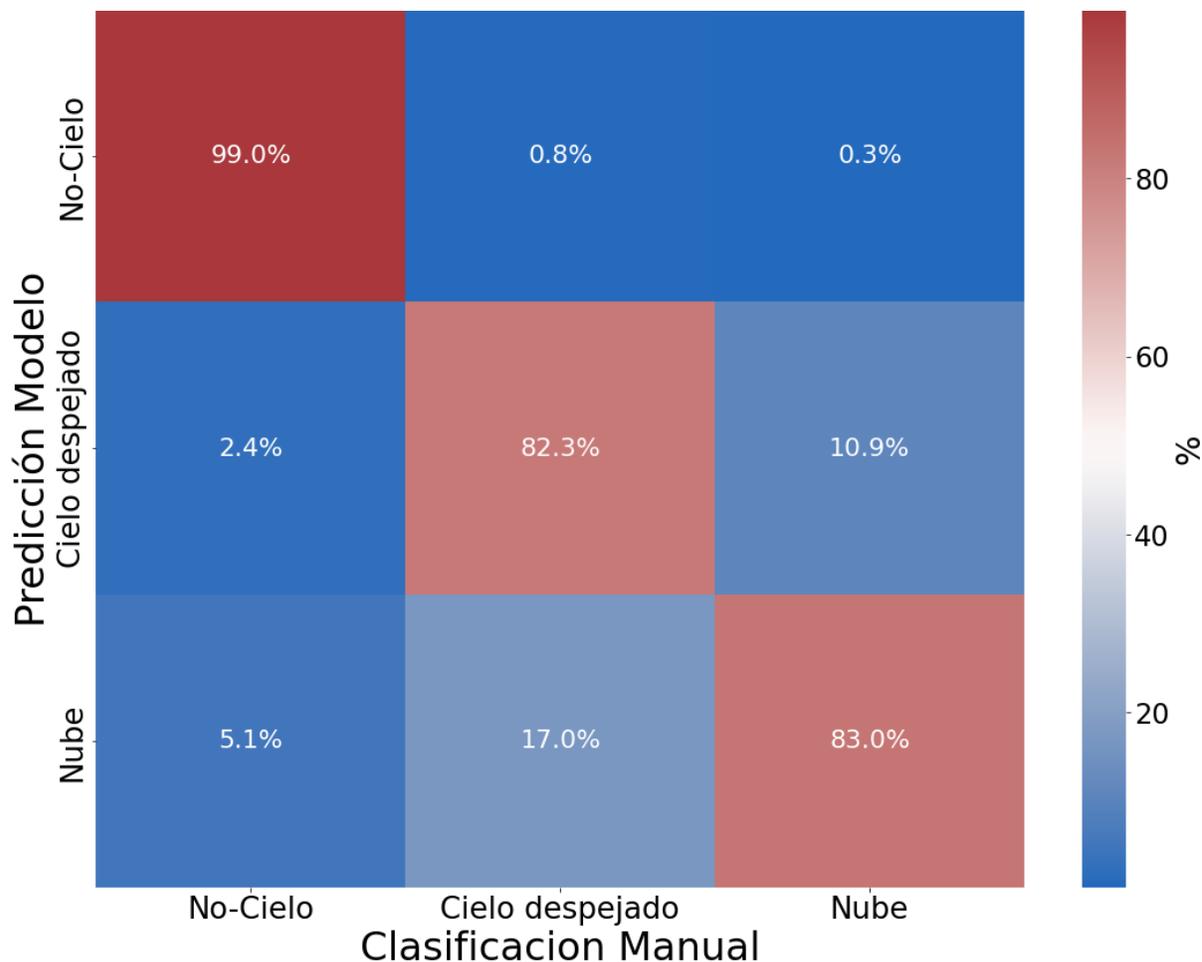


Figura 18: Matriz de confusión obtenida al comparar las máscaras manuales con las generadas automáticamente.

A través de esta matriz podemos observar que en el caso del no-cielo, la coincidencia es prácticamente exacta. Esto se debe a que, como ya hemos dicho antes, en nuestras imágenes analizadas a mano esta parte viene predeterminada. En cambio, el modelo reconoce algunos puntos de no-cielo como nube o cielo despejado. Esto se debe a que de manera ocasional el modelo reconocerá zonas brillantes causadas por algún reflejo solar en la zona de no-cielo como una zona de cielo.

En el caso de las dos opciones dentro de la propia imagen (sin tener en cuenta el “marco”) podemos observar que las coincidencias son del 82,3 % en el caso del cielo claro y del 83 % en el de nube. Contando con el error humano y la deliberada sobreestimación, estos son buenos resultados, ya que demuestran una efectividad de más del 80 %.

En los términos no diagonales se puede ver que nuestro modelo es más propenso a con-

fundir una parte de cielo claro, según la clasificación manual, con una nube que viceversa. Esto se observa en que el porcentaje de la casilla nube-cielo despejado es 6 puntos mayor que el de la casilla cielo despejado-nube.

## 4. Conclusiones

Las cámaras de cielo se pueden emplear para proporcionar imágenes que permitan analizar y clasificar la cobertura de nubes en función de sus características de color e intensidad. El objetivo principal de nuestro trabajo era poder localizar las nubes dentro de cada una de estas imágenes, a diferencia de otros estudios anteriores centrados en determinar qué proporción de la imagen se encuentra cubierta por nubes.

Como herramienta se ha utilizado una red neuronal convolucional denominada U-Net, que realiza una tarea de segmentación semántica. Este enfoque, basado en algoritmos de aprendizaje profundo, aplica una técnica esencial para dividir una imagen digital en segmentos o regiones, asignando a cada píxel una etiqueta o categoría específica y posibilitando la identificación y diferenciación precisa de objetos y características en la imagen.

El etiquetado de las imágenes ha sido manual, con más de 200 imágenes procesadas, en las que se seleccionan las nubes generando así una máscara que, idealmente, cubrirá todas las partes del cielo que estén cubiertas por nubes y dejará sin seleccionar los claros en la cubierta nubosa.

La matriz de confusión proporciona una representación gráfica de nuestros resultados, permitiendo visualizar el desempeño de nuestro modelo al comparar las predicciones hechas por el modelo con las etiquetas reales de los datos. Con una proporción de coincidencia entre el modelo y la clasificación manual mayor al 80 %, podemos determinar que este modelo funciona de una manera aceptable para nuestro trabajo.

A pesar de esto, se pueden diseñar algunas estrategias para mejorar este resultado. Una primera consistiría en corregir los errores humanos en el etiquetado manual, siendo más precisos en la delimitación de las zonas cubiertas en cada imagen procesada. Una segunda opción podría ser realizar el mismo trabajo con un mayor número de imágenes, lo que permitiría un mayor entrenamiento de la inteligencia artificial en uso. Por último, una mejora de la resolución de las imágenes con las que se trabaja posibilitaría establecer más claramente la diferencia entre nube y cielo despejado en cada imagen.

## Referencias

- [1] S. Alegre, *Clasificación automática de imágenes de cielo mediante inteligencia artificial*. Trabajo de Fin de Grado en Física, Universidad de Valladolid, 2021, disponible online en Uvadoc, <https://uvadoc.uva.es/handle/10324/58262> (último acceso noviembre de 2023).
- [2] J. C. Antuña Sánchez *Configuración y metodología para el uso de cámaras de todo cielo en la obtención de parámetros atmosféricos*. Tesis Doctoral, Universidad de Valladolid, 2021, disponible en Uvadoc, <https://uvadoc.uva.es/handle/10324/60022> (último acceso noviembre de 2023).
- [3] *Atlas Internacional de Nubes*, Organización Meteorológica Mundial, 2017, <https://cloudatlas.wmo.int/es/home.html> (último acceso noviembre de 2023).
- [4] H. I. Aysel, X. Cai, A. Prügel-Bennett, *Semantic Segmentation by Semantic Proportions*. Disponible en arxiv, <https://arxiv.org/abs/2305.15608> (último acceso noviembre de 2023).
- [5] E. E. Clothiaux, T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner, 2000: Objective Determination of Cloud Heights and Radar Reflectivities Using a Combination of Active Remote Sensors at the ARM CART Sites. *J. Appl. Meteor. Climatol.*, 39, 645–665, <https://doi.org/10.1175E>.
- [6] A. Dybbroe, K. Karlsson, A. Toss, *Cloud Detection and Analysis Using Dynamic Thresholds and Radiative Transfer Modeling. Part I: Algorithm Description*, *J. Appl. Meteor. Climatol.* 44, 39–54, <https://doi.org/10.1175/JAM-2188.1>, 2005.
- [7] U. Feister, G. Meyer, and U. Kirst, *Solar UV Radiation Exposure of Seamen – Measurements, Calibration and Model Calculations of Erythral Irradiance along Ship Routes*, AIP Conference Proceedings, Volume 1531, Issue 1, p.860–863, 2013, DOI:10.1063/1.4804906.
- [8] A. Heinle, A. Macke, A. Srivastav, *Automatic cloud classification of whole sky images*, *Atmos. Meas. Tech.*, 3, 557–567, <https://doi.org/10.5194/amt-3-557-2010>, 2010.
- [9] S. Kato, G. G. Mace, E. E. Clothiaux, J. C. Liljegren, R. T. Austin, *Doppler cloud radar derived drop size distributions in liquid water stratus clouds*, *Journal of Atmospheric Sciences*, vol. 58, Issue 19, 2895–2911, 2001.
- [10] B. Martínez Celda, *Clasificación automática de imágenes de cielo mediante inteligencia artificial*. Trabajo de Fin de Grado en Física, Universidad de Valladolid, 2020, disponible online en Uvadoc, <https://uvadoc.uva.es/handle/10324/50628> (último acceso noviembre de 2023).
- [11] M. Martínez-Chico, F. J. Batlles, *Clasificación de las nubes a partir de datos de radiación e imágenes de una cámara de cielo*. Acta de las Jornadas Científicas de la Asociación Meteorológica Española, ISSN-e 2605-2199, N<sup>o</sup>. 31, 2010.

- 
- [12] R. Román, M. Antón, A. Cazorla, A. de Miguel, F. J. Olmo, J. Bilbao, and L. Alados-Arboledas, *Calibration of an all-sky camera for obtaining sky radiance at three wavelengths*, Atmos. Meas. Tech., 5, 2013–2024, <https://doi.org/10.5194/amt-5-2013-2012>, 2012.
- [13] O. Ronneberger, P. Fischer and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Universidad de Friburgo, 2015.
- [14] K. Sassen, L. Wang, *Lidar applied to Cirrus cloud research*, 85th AMS Annual Meeting, American Meteorological Society - Combined Preprints, 2005.
- [15] O. A. Soto-Orozco, A. D. Corral-Sáenz, C. E. Rojo-González, J. A. Ramírez-Quintana, *Análisis del desempeño de redes neuronales profundas para segmentación semántica en hardware limitado*. ReCIBE, Revista electrónica De Computación, Informática, Biomédica Y Electrónica, 8(2), C–6. <https://doi.org/10.32870/recibe.v8i2.142>, 2019.
- [16] R. Tapakis, A. G. Charalambides, *Equipment and methodologies for cloud detection and classification: A review*. Solar Energy, Volume 95, 2013, 392–430, <https://doi.org/10.1016/j.solener.2012.11.015>.
- [17] K. Tsuchiya, M. Tokuno, *Comparison of different sensor data and characteristics of cirrus clouds*, Advances in Space Research, Volume 12, Issue 7, 1992, 343–352, ISSN 0273-1177, [https://doi.org/10.1016/0273-1177\(92\)90238-S](https://doi.org/10.1016/0273-1177(92)90238-S).
- [18] <https://www.anaconda.com> (último acceso noviembre de 2023).
- [19] <https://public.wmo.int/es/files/jacques-desclotres-modis-rapid-response-team-nasagsfc> (último acceso noviembre de 2023).