

### **(07-011) - Application of Deep Learning to satellite image processing**

Villanueva Balsera, Joaquín Manuel <sup>1</sup>; Fernández Gutiérrez, Adrián <sup>2</sup>; Cedillo Gancedo, Cristian <sup>1</sup>; Rodríguez Montequín, Vicente <sup>1</sup>; Alonso Fernández, Ángela <sup>2</sup>

<sup>1</sup> Universidad de Oviedo, <sup>2</sup> SERESCO SA

In the current context, the application of Deep Learning to satellite image processing is presented as a strategic opportunity. Numerous open source accessible sources, such as Sentinel, Landsat, and MODIS, which generate a constant flow of multispectral satellite data, creating a large and ever-growing amount of geospatial information. However, this data is almost unused.

The successful application of Deep Learning techniques requires large datasets, and fortunately, in the context of satellite imagery, we have a rich supply. By harnessing this wealth of information, we not only satisfy the essential requirement for Deep Learning, but also bring about a significant transformation. Automating the analysis of this imagery saves substantial time and economic resources, and certainly justifies its application.

This article explores the Deep Learning techniques with the greatest potential for use in project development, defining a classification of data sources, methods, and application. From enhancing the resolution of geospatial imagery, to terrain classification and terrain mapping, key areas that offer revolutionary potential in process optimisation are covered.

Keywords: Deep Learning; open data; multispectral satellite data; remote sensing; process automation; classification

#### **Aplicación de Deep Learning al procesamiento de imágenes satelitales**

En el contexto actual, la aplicación de Deep Learning al procesamiento de imágenes satelitales se presenta como una oportunidad estratégica. Existe una gran cantidad de fuentes de libre acceso, como Sentinel, Landsat y MODIS, que generan un flujo constante de datos satelitales multiespectrales, creando una gran cantidad de información geoespacial en constante crecimiento. Sin embargo, estos datos están infrutilizados.

La aplicación exitosa de técnicas de Deep Learning requiere extensos conjuntos de datos, y afortunadamente, en el contexto de las imágenes satelitales, contamos con un abundante suministro. Al aprovechar esta riqueza de información, no solo satisfacemos la necesidad esencial del Deep Learning, sino que también provocamos una transformación significativa. Automatizar el análisis de estas imágenes implica un ahorro sustancial en tiempo y recursos económicos, y sin duda justifica, de manera concluyente, su aplicación.

Este artículo explora las técnicas de Deep Learning con mayor potencial para ser usadas en el desarrollo de proyectos, definiendo una clasificación de fuentes de información, métodos y aplicación. Desde el aumento de la resolución de imágenes geoespaciales, hasta la clasificación y el mapeo del terreno, se abarcan áreas clave que ofrecen un potencial revolucionario en la optimización de procesos.

Palabras clave: Deep Learning; datos de libre acceso; datos satelitales multiespectrales; teledetección; automatización de procesos; clasificación



©2024 by the authors. Licensee AEIPRO, Spain. This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Correspondencia: Joaquín Manuel Villanueva Balsera, [jmvillanueva@uniovi.es](mailto:jmvillanueva@uniovi.es)

Agradecimientos: Este trabajo ha sido financiado gracias a la convocatoria de ayudas 2021 destinadas a proyectos de investigación y desarrollo en inteligencia artificial y otras tecnologías digitales y su integración en las cadenas de valor - C005/21-ED. IA4GEO: Visión Artificial y Deep Learning para la detección de cambios y usos del suelo en el territorio y su representación directa en las bases de datos cartográficas. 2021/C005/00147326

## 1. Introducción

El aprendizaje profundo, una rama de la inteligencia artificial, se ha convertido en una poderosa herramienta en la transformación de múltiples disciplinas, impulsando avances significativos en distintos campos. En el ámbito de la salud, ha prometido transformar la informática aplicada a la medicina al mejorar los diagnósticos y las terapias, aprovechando datos complejos para generar características e interpretaciones optimizadas (Ravi et al., 2017).

En el sector empresarial, el aprendizaje profundo es vital para el desarrollo de productos innovadores en reconocimiento de imágenes y voz, asistentes personales y vehículos autónomos. El acceso a grandes conjuntos de datos, un mayor poder computacional y los nuevos algoritmos de aprendizaje automático han permitido lograr resultados innovadores (Vieira & Vieira, 2017).

La ecología es otro campo que ha empezado a beneficiarse del aprendizaje profundo. La capacidad de estas técnicas para manejar el creciente tamaño y complejidad de los conjuntos de datos ecológicos ha abierto nuevos horizontes para la investigación y la conservación, permitiendo hacer predicciones y tomar decisiones más precisas (Christin et al., 2019).

La conducción autónoma ha experimentado una profunda transformación con la integración del aprendizaje profundo en la detección de objetos y el reconocimiento de escenas. Estas aplicaciones no solo prometen mejorar la seguridad vial sino también reducir la necesidad de intervención humana (Gupta et al., 2021).

Por otra parte, el programa Copernicus, liderado por la Unión Europea, representa un avance significativo en la observación terrestre mediante satélites, proporcionando datos globales en tiempo cuasirreal para diversas aplicaciones (Unión Europea, s. f.). Este programa incluye la familia de satélites Sentinel y es posiblemente el mayor proveedor de datos de observación de la Tierra, ofreciendo información precisa y actualizada sobre el clima y el medio ambiente diariamente. Estos datos son esenciales para la monitorización de la superficie terrestre, el entorno marino, la atmósfera y el cambio climático, entre otras (Jutz & Milagro-Pérez, 2020).

La accesibilidad y el manejo de estos grandes volúmenes de datos representan un desafío, abordado por proyectos como Copernicus App Lab, que facilita el acceso a la información de Copernicus, potenciando su utilización por parte de desarrolladores y fomentando la creación de nuevas aplicaciones basadas en estos datos (Bereta et al., 2019).

La abundancia de datos es precisamente lo que requieren los modelos de aprendizaje profundo para generar resultados significativos. La convergencia entre la necesidad de datos y la disponibilidad de imágenes satelitales representa una oportunidad sin precedentes para el desarrollo de proyectos. El acceso a grandes conjuntos de datos permitiría un entrenamiento más efectivo de los modelos de aprendizaje profundo, generando información muy valiosa. El objetivo de este artículo es explorar cómo la convergencia entre el aprendizaje profundo y la disponibilidad de datos satelitales de libre acceso puede contribuir a la relación de proyectos. La combinación de ambas tecnologías presenta un potencial significativo en proyectos relacionados con la necesidad de observación de la superficie terrestre y de la atmósfera.

El presente artículo pretende realizar un pequeño estado del arte con las principales técnicas de aprendizaje profundo que pueden aplicarse en proyectos relacionados con el procesamiento de imágenes satelitales, proporcionando una visión general de las técnicas más relevantes y especialmente aplicables a los datos.

A su vez, se muestra un ejemplo de aplicación real que demuestra la utilidad y efectividad de la combinación entre el aprendizaje profundo y la teledetección. Con ello se pretende ofrecer un caso práctico que ilustre el potencial de estas técnicas.

El presente artículo está estructurado para mostrar el estado del arte de varias técnicas de aprendizaje profundo aplicadas a datos satelitales. Se iniciará con la clasificación, segmentación y detección de objetos, las cuales constituyen la base para entender las imágenes satelitales. Posteriormente, se profundizará en técnicas más especializadas y relevantes como son la detección y borrado de nubes, la super resolución y la detección de cambios. Se concluirá con un caso de estudio práctico, ilustrando la aplicación real de la detección de cambios en un contexto específico.

## 2. Técnicas

En esta sección se examinan las técnicas de aprendizaje profundo que presentan un mayor potencial de aplicación en el análisis de datos satelitales. Estas técnicas han sido seleccionadas tras una revisión de la documentación existente sobre ellas.

### 2.1 Clasificación

La clasificación es una técnica del aprendizaje profundo que se basa en asignar una etiqueta o categoría predefinida a unos datos. Este proceso implica la construcción de modelos o clasificadores para asignar etiquetas a las imágenes basándose en sus características (Kesavaraj & Sukumaran, 2013).

Los avances recientes en el campo han demostrado que el empleo de Redes Neuronales Convolucionales Profundas (DCNN) pre entrenadas son eficaces realizando tareas de clasificación de la superficie terrestre. El estudio Qian et al. (2022) utiliza imágenes del satélite ZiYuan-3 para obtener imágenes espaciales multibanda. Utilizando estos datos, el estudio explora varias combinaciones de bandas y modelos para averiguar cuales ofrecen los mejores resultados en la clasificación del suelo.

Existen múltiples conjuntos de datos preparados para su uso en tareas de clasificación, es el caso del conjunto de imágenes satelitales EuroSat (Helber et al., 2019), un recurso extremadamente útil para la clasificación del uso y la cobertura del suelo, utilizando imágenes de satélite Sentinel-2 accesibles a través del programa Copernicus (Unión Europea, s. f.) ya mencionado. Este conjunto de datos comprende 27000 imágenes georreferenciadas y etiquetadas en 10 clases distintas y 13 bandas espectrales. Las imágenes, de 64x64 píxeles han sido revisadas a mano para asegurarse de que son válidas para su integración en los datos. Las redes neuronales convolucionales profundas (CNN o DCNN) empleadas en este modelo han alcanzado una precisión del 98.57% en EuroSat, demostrando así su eficacia.

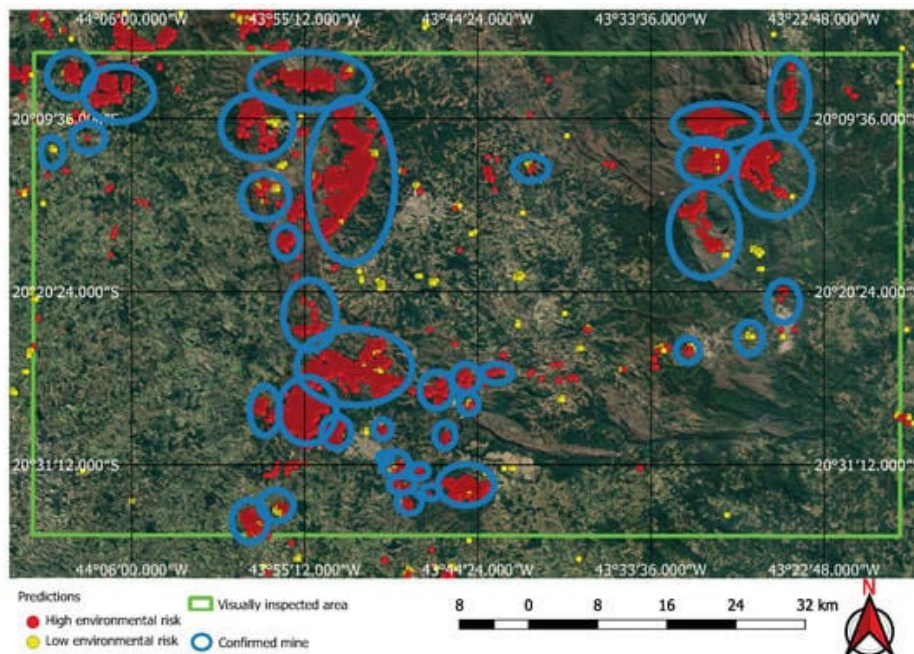
**Figura 1: Muestra de imágenes del conjunto de datos EuroSat (Helber et al., 2019) junto a la clase a la que pertenecen.**



La técnica de clasificación no solo sirve para aplicarse directamente sobre la superficie terrestre, se puede utilizar con fines más concretos como la identificación y categorización de

minas. Utilizando datos gubernamentales de libre acceso, imágenes satelitales Sentinel-2 procesadas en la plataforma Google Earth Engine, y redes neuronales convolucionales implementadas en TensorFlow 2 y Google Colab, el estudio Balaniuk et al. (2020) logra descubrir 263 minas no registradas en Brasil. Este trabajo destaca como conseguir un gran impacto con tecnologías de libre acceso y bajo costo, al mismo tiempo que aborda problema de la minería ilegal y le aparición de presas por los residuos generados. En la figura 2 se puede observar un ejemplo de resultado de la red. En amarillo y rojo aparecen las predicciones del modelo, mientras que en azul se señalan aquellos lugares en los que la predicción se ha confirmado de manera visual. Todos aquellos puntos fuera de las elipses son falsos positivos.

**Figura 2: Aciertos y fallos del modelo. Imagen procedente del artículo Balaniuk et al. (2020)**



Otro artículo relevante relacionado con las aplicaciones particulares de los problemas de clasificación es el realizado por Chen et al. (2021), donde se propone una metodología basada en el aprendizaje profundo para la clasificación automatizada de redes de carreteras en cuatro categorías distintas según sus características visuales. Este enfoque implica la generación de una representación visual de la red de calles denominada "Diagrama Jerárquico de Carreteras Coloreado" y su clasificación mediante una red neuronal convolucional, específicamente una ResNet-34. Los resultados obtenidos muestran que el modelo logra una precisión del 87.5%.

El estudio (Chen et al., 2021) se lleva a cabo utilizando datos de redes de carreteras de nueve ciudades en todo el mundo, obtenidos de OpenStreetMap. Además de la clasificación de las redes de carreteras, el artículo examina la utilidad de este enfoque para predecir la vitalidad urbana. Se utiliza un modelo de regresión basado en árboles (LightGBM) para analizar la relación entre las características derivadas de la clasificación de la red de carreteras y los indicadores de vitalidad urbana, revelando que, si bien el efecto sobre ella es limitado, existe una asociación positiva entre ambos.

## 2.2 Segmentación

La segmentación es, desde hace tiempo, otra de las grandes beneficiadas del aumento de las imágenes satelitales de alta resolución, ya que esta técnica permite extraer gran cantidad de información. La segmentación se basa en dividir una imagen satelital en regiones más pequeñas y significativas, llamadas segmentos. A diferencia de la clasificación, que asignaba



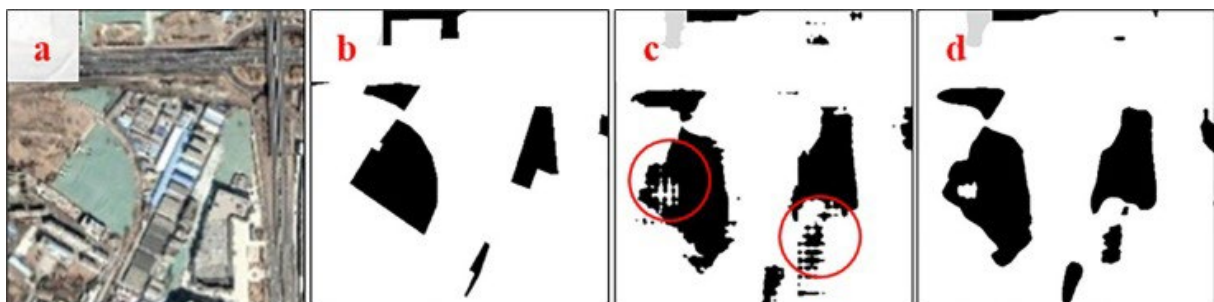
una etiqueta a toda la imagen, la segmentación permite identificar áreas específicas de interés dentro de ella, centrándose en la contigüidad espacial de los segmentos en los que ha dividido la imagen (Lucieer & Stein, 2002).

Estudios más recientes (Riabko, 2023) reconocen que la segmentación aplicada a imágenes satelitales, que pueden ser muy complejas, es una técnica esencial para la extracción de información en teledetección. El objetivo es agrupar los segmentos con atributos similares, como color, textura o intensidad en un mismo conjunto o característica.

La segmentación tiene aplicaciones muy diversas y particulares en algunos casos, como puede ser la detección de parcelas vacantes en ciudades super pobladas. La investigación de Mao et al. (2022) propone utilizar la segmentación semántica de imágenes satelitales de alta resolución para detectar terrenos urbanos vacíos a gran escala. Uno de los aspectos más distintivos de la metodología empleada es la estratificación de las ciudades antes del entrenamiento. Este paso implica dividir las ciudades en diferentes categorías y seleccionar una ciudad representativa en cada una para establecer un modelo común al resto de ciudades en la categoría. Esta estratificación se basa en las características de los terrenos vacantes.

El modelo entrenado utilizando esta metodología logra una precisión superior al 90%, lo que representa una mejora significativa en comparación con los métodos manuales tradicionales. Además de mejorar la precisión, aumenta la eficiencia del proceso de detección, reduciendo el tiempo y el costo necesarios para realizar la tarea.

**Figura 3: Resultados del modelo. (a) es la imagen satelital cortada. (b) es la máscara. (c) el resultado del modelo sin optimizar. (d) el resultado con los bordes tratados. Imagen procedente del artículo Mao et al. (2022).**



Otro caso de aplicación particular, relacionado con Chen et al. (2021), sobre la clasificación de carreteras, tiene que ver con la aplicación de la segmentación a encontrar conexiones o cortes en estas. Este problema lo aborda Batra et al. (2019) presentando una metodología novedosa para la extracción de redes viales a partir de imágenes satelitales. El estudio aborda el problema de la obtención de mapas de carreteras fragmentados, derivado de las técnicas de segmentación tradicionales.

Para superar estas limitaciones, Batra et al. (2019) propone una tarea de conectividad denominada Aprendizaje de Orientación, inspirada en el método de trazar carreteras manteniendo una orientación específica, y un modelo convolucional multi-rama que aprovecha tanto la segmentación como la orientación. Esto permite al modelo predecir máscaras de carreteras conectadas, contribuyendo así a la creación de mapas viales más precisos a partir de imágenes satelitales.

**Figura 4: Imagen de ejemplo del uso de la orientación junto a la segmentación para extraer un mapa de carreteras más acertado. (a) es la imagen de máscara, creada a mano. (b) es representa el resultado de una segmentación tradicional. (c) carreteras trazadas con orientación. (d) resultado obtenido de juntar ambas técnicas. Imagen procedente del artículo Batra et al. (2019).**



Pese al enorme potencial que presenta la segmentación, también tiene debilidades. Uno de los problemas habituales de la segmentación es la especialización de los modelos a entornos específicos, limitando su transferibilidad y escalabilidad a otros emplazamientos. Para abordar esta problemática se dispone del conjunto LoveDA (Land-cOVER Domain Adaptive semantic segmentation) (J. Wang et al., 2021). Este conjunto de datos cuenta con 5987 imágenes de tres ciudades distintas y abarca dos dominios: una zona urbana y otra rural. Mientras que el entorno urbano se caracteriza por su densidad de construcciones y estructuras, el entorno rural cuenta con una distribución de elementos más dispersa y mayor vegetación. Esta variedad permite la evaluación de algoritmos de segmentación adaptados a diferentes condiciones geográficas mediante el uso de la Adaptación de Dominio No Supervisada (UDA). Estos algoritmos consisten en aplicar modelos entrenados en un dominio, a otros distintos, mostrando su potencial de aplicación en lugares distintos al original.

### 2.3 Detección de objetos

La detección de objetos es una técnica de visión por computador que permite identificar y localizar objetos en imágenes o vídeos. Utiliza algoritmos de aprendizaje profundo para procesar y entender visualmente el contenido de estas. Su aplicación a las imágenes satelitales de libre acceso es más limitada, debido a la poca resolución que ofrecen y a buscar objetos pequeños en una imagen más bien grande (Tahir et al., 2022).

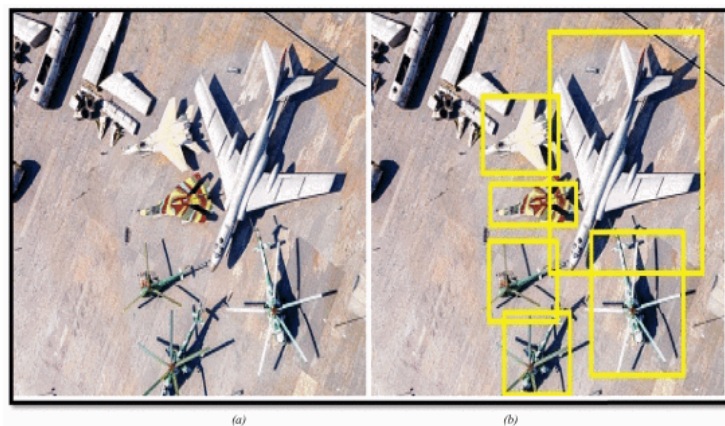
Estudios como el de Singh & Nene (2022), abordan los desafíos inherentes a la detección de objetos en imágenes satelitales, como la variabilidad en la escala y orientación de los objetos debido a la perspectiva de las imágenes satelitales. Los autores proponen un modelo que

utiliza mecanismos auto supervisados para identificar detalles en las imágenes a pesar del desafío que presenta la complejidad y el alto costo computacional inducido por la resolución de las imágenes.

La metodología empleada en la citada investigación incluye el uso de técnicas de data augmentation, como el recorte aleatorio y la creación de mosaicos de imágenes, para mejorar la robustez y capacidad de generalización del modelo. Se ha utilizado una porción del dataset iSAID, que contiene imágenes de alta resolución con 655451 objetos de 15 categorías, para entrenar y evaluar el modelo propuesto.

El resultado de este estudio, ejemplificado en la figura 5, demuestra mejoras significativas en la detección de objetos utilizando imágenes satelitales de muy alta resolución. El modelo es capaz de detectar objetos de diferentes clases con una gran precisión, pero es importante reconocer que el éxito obtenido se basa en el uso de imágenes de satélites cuyo acceso no parece ser completamente libre, como el GF-2 y JL-1. Esto sugiere que, para replicar o superar estos resultados, se requieren imágenes satelitales de alta calidad que no siempre están disponibles de manera gratuita o fácil para la comunidad investigadora en general. Por tanto, aunque el artículo demuestra la viabilidad de la detección de objetos en imágenes satelitales, queda por investigar cómo estos métodos se traducirían al trabajar con imágenes satelitales de libre acceso, como las proporcionadas por el programa Copernicus, que, si bien son accesibles, no ofrecen la misma resolución. Este aspecto subraya la necesidad de continuar explorando y adaptando técnicas de detección de objetos que puedan operar con efectividad dentro de las limitaciones que ofrecen las imágenes satelitales de libre acceso.

**Figura 5: Imagen de ejemplo de la inferencia del modelo. (a) es la imagen de entrada. (b) representa la salida del modelo. Imagen procedente del artículo Singh & Nene (2022).**



Uno de los estudios que sí consigue utilizar imágenes de libre acceso, de Sentinel-2, para la detección de objetos, es el Kolokoussis & Karathanassi (2018). Este se realiza en dos áreas: la zona A, un vertido natural al sur de la isla Zakynthos (Grecia) y la zona B, un vertido artificial procedente del petrolero “Agia Zoni II”. Si bien no se proporcionan resultados numéricos específicos, sí se describe la metodología aplicada: la segmentación multi-resolución. Este proceso se basa en fusionar los píxeles en la imagen en segmentos u objetos con cierto criterio de homogeneidad. Se utilizan dos niveles de segmentación: el primero para crear objetos pequeños basados principalmente en comportamiento espectral y el segundo para objetos más grandes. Ambos objetos no pueden superponerse. El ajuste de los parámetros se cambia según se trate de elementos grandes o pequeños.

Finalmente, basado en las propiedades espectrales y geométricas de los objetos de la imagen, se calculan diversas capas, como el índice de diferencia normalizada de agua (NDWI) para la diferenciación entre mar y tierra, la relación B2/B11, y una característica particular del



estudio que multiplica la desviación estándar de la banda 2 del objeto de imagen con el valor de la relación anterior.

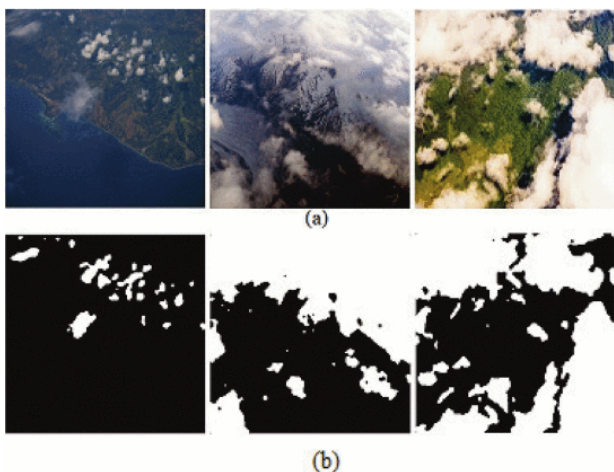
## 2.4 Detección y borrado de nubes

La detección de nubes y sus sombras es crucial en el preprocesamiento de imágenes satelitales, como se destaca en el artículo de Hollstein et al. (2016), y también es necesaria para poder realizar una corrección atmosférica. La presencia de nubes puede obstruir significativamente la observación de la superficie terrestre, haciendo imposible la posterior utilización de las imágenes. Además, identificar nubes y cirros (nubes altas) es fundamental para filtrar y preparar los datos para análisis más detallados. Esto es especialmente relevante dado el enorme volumen de datos generados por Sentinel-2.

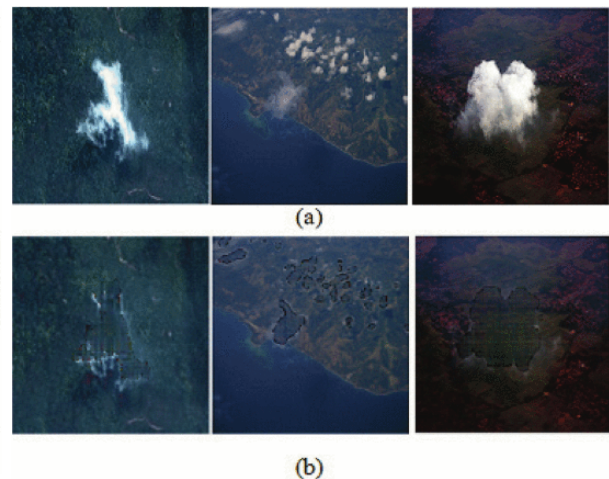
El ISCCP-FD (International Satellite Cloud Climatology Project-Flux Data) estima que la tasa de nubes mundial anual es del 66%. Por tanto, la presencia de nubes y sus sombras es un gran problema para el uso de las imágenes satelitales. El reflejo de las nubes y la oscuridad que provocan sus sombras complica su uso, convirtiéndose en una de las principales fuentes de ruido de los datos satelitales. Zhu & Woodcock (2012) ya proponían hace años un algoritmo para detectar este problema en las imágenes Landsat utilizando como entradas: la reflectancia de las nubes en lo alto de la atmósfera y el brillo.

La investigación realizada por Prinsa & Saritha (2018) presenta una metodología en la detección y eliminación de nubes en imágenes satelitales, mediante la combinación de múltiples características como color, textura y reflectancia. Utilizando la descomposición en valores singulares (SVD), no solo se consiguen identificar las áreas nubladas, sino que también se reconstruyen, mejorando así las imágenes para análisis posteriores. Este enfoque resulta ser una herramienta valiosa para mejorar la calidad de las imágenes en aplicaciones de teledetección, donde las nubes son sin duda un gran obstáculo.

**Figura 6: Ejemplo de detección de nubes. (a) son las imágenes utilizadas para test. (b) el resultado de la detección. Imagen procedente del artículo Prinsa & Saritha (2018)**



**Figura 7: Ejemplo de reconstrucción de la imagen tras las nubes. (a) son las imágenes utilizadas para test. (b) el resultado de la reconstrucción. Imagen procedente del artículo Prinsa & Saritha (2018)**



Encontrar y borrar nubes es una tarea complicada, puede requerir disponer de ciertos canales de una imagen o de una secuencia de estas. Estudios recientes como T. Wang et al. (2019) proponen metodologías novedosas como utilizar un esquema basado en la radiación de las nubes de la parte alta de la atmósfera. Esta propuesta pretende ser más accesible para los

usuarios que las anteriores. Su metodología consiste en utilizar como referencia una imagen de la radiancia emitida por la superficie de la tierra creada sin nubes. La presencia de nubes se determina comparando esta imagen de referencia con la imagen real satelital.

Se puede concluir, sin ninguna duda, que la detección de las nubes es un campo muy importante y tratado en la teledetección. No solo permite saber cuando las imágenes usadas están cubiertas de nubes, si no que además abren la puerta a reconstruir estas imágenes y poder usarlas de forma efectiva en el entrenamiento de modelos de aprendizaje profundo con cualquier otro fin.

## 2.5 Super resolución

La superresolución (SR) se presenta como una técnica de procesamiento de señales diseñada para obtener imágenes de alta resolución a partir de múltiples imágenes de baja resolución (Park et al., 2003) . La superresolución es una de las áreas más importantes de investigación en el procesamiento de imágenes. Se ha comprobado que su uso es útil en múltiples áreas, como las imágenes médicas, satelitales y vídeos.

Estudios recientes como Bull et al. (2021) presentan técnicas innovadoras para mejorar la resolución en imágenes de teledetección. Mediante la combinación de dos métodos de superresolución basados en aprendizaje profundo, DeepSUM y ESRGAN se consiguen imágenes de alta resolución a partir de varias imágenes de baja calidad. Posteriormente se mejora la definición de bordes y texturas en la imagen utilizando un tipo modificado de función de pérdida.

Si tomamos de punto de partida un vídeo corto de la zona en cuestión, Anger et al. (2020) proponen un modelo de interpolación capaz de combinar múltiples frames de baja resolución para crear una imagen de alta calidad. Este modelo puede implementarse fácilmente en gpu, lo que permite mejorar la eficiencia en el procesado de las imágenes. Esta eficiencia es fundamental para sus posibles futuras aplicaciones, ya que la super resolución puede ser un paso del preprocesamiento de una imagen antes de su uso en otro modelo.

Esta técnica es, sin duda alguna, una de las más importantes por su capacidad de mejorar el preprocesamiento de las imágenes de teledetección. Su correcta aplicación a las imágenes de entrada de un modelo permitirá obtener mejores resultados.

## 2.6 Detección de cambios

La detección de cambios en imágenes satelitales es vital para la observación de la superficie terrestre, la planificación urbana, la agricultura, evaluación de desastres y revisión de mapas. En el estudio Shi et al. (2020) se revisan varios métodos de detección de cambios basados en inteligencia artificial, explorando sus distintas aplicaciones. Se discuten en detalle los enfoques actuales de IA para la detección de cambios, incluyendo los datos de diferentes sensores como la teledetección e imágenes callejeras. También se presentan los desafíos y perspectivas futuras de la IA en este campo, como el procesamiento de grandes conjuntos de datos, la IA no supervisada y su fiabilidad.

Si se consideran como datos imágenes satelitales de alta resolución aparecen múltiples estudios que las tratan. Es el caso de Li et al. (2022). Este trabajo propone un modelo (DARNet) para este tipo de imágenes. Este modelo se basa en una arquitectura de codificador-decodificador en forma de U y emplea una red siamesa como extractor de características. Con esta propuesta consiguen resultados prometedores en cuanto al rendimiento respecto a otros modelos.

Sin embargo, si se pretenden usar imágenes multiespectrales de libre y fácil acceso, como las de Sentinel-2, Daudt et al. (2018) proponen el uso del conjunto de datos Onera Satellite Change Detection (OSCD). Este conjunto de imágenes está compuesto por un par de

imágenes de varias localizaciones en distintos momentos y una máscara de cambios entre ellas. Los buenos resultados, de hasta el 90% de precisión, obtenidos por la red siamesa en este estudio, evidencian que es posible detectar cambios en imágenes de 10 metros por píxel y utilizando únicamente cuatro de las trece bandas que ofrece Sentinel-2: RGB e infrarrojo.

### **3. Caso práctico**

En la siguiente sección abordaremos un caso de estudio concreto, desarrollado en colaboración con la empresa Seresco, que ilustra la aplicación de varias técnicas mencionadas en los apartados anteriores. El reto planteado posee dos objetivos fundamentales. En primer lugar, se propone estudiar la viabilidad de detectar cambios en el paisaje utilizando inteligencia artificial aplicada a imágenes satelitales. Este objetivo implica en primer lugar encontrar un conjunto de datos adecuado para poder evaluar distintos modelos de aprendizaje profundo, tales como las redes siamesas, que sean capaces de detectar variaciones en dos imágenes de un mismo lugar tomadas en dos instantes de tiempo diferentes. En segundo lugar, se busca desarrollar una herramienta funcional que permita la inferencia del modelo de detección de cambios. Esta herramienta estaría diseñada para procesar pares de imágenes satelitales, aplicar el modelo de aprendizaje profundo entrenado y generar como salida una máscara que identifique las áreas donde han ocurrido cambios. La consecución de estos objetivos resultaría en una herramienta diseñada para automatizar una actividad puramente humana, como puede ser el registro de fincas en el catastro.

A continuación, se presenta la metodología empleada para la realización del trabajo siguiendo una distribución habitual: adquirir imágenes satelitales válidas para el estudio, preprocesarlas de acuerdo con las necesidades planteadas y las de la red, elegir una arquitectura de red y entrenar el modelo. Finalmente se muestra el resultado obtenido.

#### **3.1 Adquisición de datos**

En esta fase se realizaron evaluaciones exhaustivas de diversas fuentes de imágenes satelitales con el objetivo de seleccionar la más adecuada. Entre las opciones consideradas se encontraba Google Earth y los satélites Landsat y Sentinel.

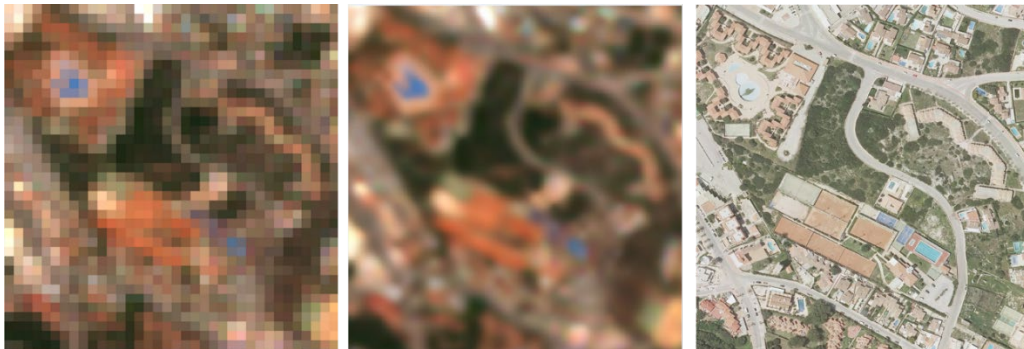
Inicialmente, Landsat fue descartado debido a que su resolución máxima de 15 metros por píxel aportaba desventajas significativas para los objetivos de nuestro proyecto. Por otro lado, aunque Google Earth proporciona imágenes de alta resolución con una precisión muy elevada, capaces de distinguir detalles como vehículos, piscinas y farolas, se planteó la posibilidad de que algunas de estas imágenes fueran de origen aéreo y no estrictamente satelital. Además, la descarga de imágenes desde Google Earth presenta complicaciones, ya que se obtienen en forma de capturas sin geolocalización precisa, y el proceso de añadir esta información resulta engorroso y poco práctico para nuestras necesidades.

Dado este contexto, se optó por utilizar Sentinel-2 como principal fuente de datos satelitales. Sentinel-2 destaca por ofrecer una resolución de 10 metros por píxel, la mejor encontrada en imágenes satelitales de libre acceso y por ofrecer facilidades de descarga. Además, proporciona tanto imágenes RGB como multiespectrales, incluyendo hasta 13 bandas de la superficie terrestre. En particular, se selecciona el conjunto de datos Sentinel-2 L2A, que incluye 12 bandas, pues utiliza previamente la banda 10 para mitigar el impacto de los cirros o nubes altas, una de las complicaciones habituales en la interpretación de datos satelitales. Estas imágenes también vienen normalizadas, facilitando su integración y análisis directo.

### 3.2 Preprocesado de los datos

Las imágenes satelitales adquiridas de Sentinel-2 se someten a un proceso de ampliación para ajustar la resolución espacial a una escala que permita obtener una máscara poco distorsionada, y por tanto más realista de los cambios entre las imágenes. Mediante la aplicación de un factor de ampliación de 8, la resolución original de 10 metros por píxel se reduce a 1.25 metros por píxel. Este ajuste en la resolución, que se puede observar en la figura 8, permite una mayor precisión y detalle en la visualización y análisis de las características de la imagen.

**Figura 8: La imagen de la izquierda procede directamente de Sentinel-2. La imagen central es resultado del proceso de ampliación x8. La imagen de la derecha es la comparativa con una imagen aérea de la misma localización.**



Durante el proceso de ampliación, es crucial mantener el geoposicionamiento exacto de cada imagen para asegurar que la localización geográfica de los píxeles se conserve. Esto es importante para la correcta obtención de las máscaras objetivo. La creación de la máscara es un paso fundamental en este proceso y se realiza superponiendo un archivo geoposicionado proporcionado por Seresco sobre la imagen ampliada. Al tener una resolución más alta (menos metros por píxel), la máscara puede adaptarse con mayor precisión a los píxeles de la imagen.

Como último paso antes del entrenamiento: se cortan las imágenes y sus máscaras en trozos de 256 x 256 píxeles con superposición entre ellos. Es decir, cada corte de la imagen comparte buena parte de sus píxeles en cada lado con los cortes adyacentes en cada una de las direcciones y sentidos. Este método se utiliza únicamente en el entrenamiento para aumentar el conjunto de datos disponible.

Una vez completado el preprocesamiento, las imágenes y máscaras preparadas están listas para ser utilizadas en la fase de entrenamiento del modelo de detección de cambios.

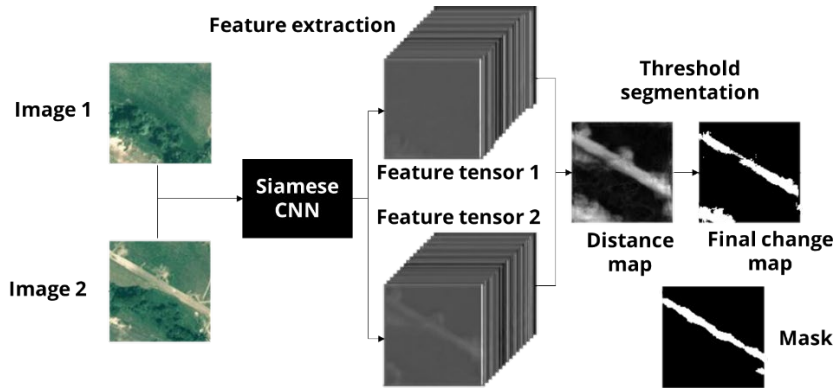
### 3.3 Arquitectura de red y entrenamiento

Para la arquitectura del modelo utilizamos una red convolucional siamesa. Esta arquitectura está diseñada para comparar dos entradas y determinar cuán similares o diferentes son entre sí. Estas redes se basan en la idea de compartir su configuración y pesos obtenidos durante el entrenamiento.

Una red convolucional siamesa se estructura en torno a tres componentes principales. Primero, cuenta con subredes gemelas que procesan las entradas de manera paralela. Estas subredes son idénticas, permitiendo la extracción de características de ambas imágenes. Segundo, una capa de fusión combina las características extraídas de cada subred, utilizando técnicas como la diferencia o la suma, preparando el terreno para la comparación. Finalmente, una capa de decisión evalúa las características combinadas para ofrecer una salida que refleja la probabilidad de similitud entre las imágenes. Para este estudio se utiliza un entrenamiento

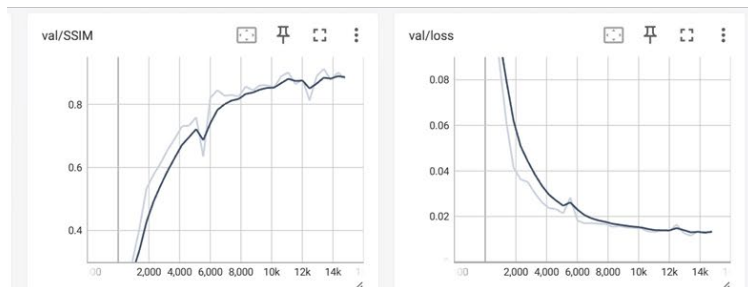
con 100 epochs, con una condición de parada de no mejorar la función de pérdida en validación durante 10 epochs.

**Figura 9: Arquitectura de la red siamesa utilizada.**



Para el entrenamiento de la red se ha utilizado como función de pérdida FocalLoss, una función de pérdida muy utilizada en problemas de aprendizaje profundo. Este método es una modificación de la pérdida en entropía cruzada que se utiliza para problemas de clases desbalanceadas como es el caso en el que en una imagen la clase de presencia de cambio es mucho menos frecuente que la sin cambio. Para ello aplica un factor de corrección alpha para tratar el desequilibrio centrando el aprendizaje en los elementos difíciles de clasificar (Lin et al., 2018).

**Figura 10: Función de pérdida y métrica de resultados en validación.**

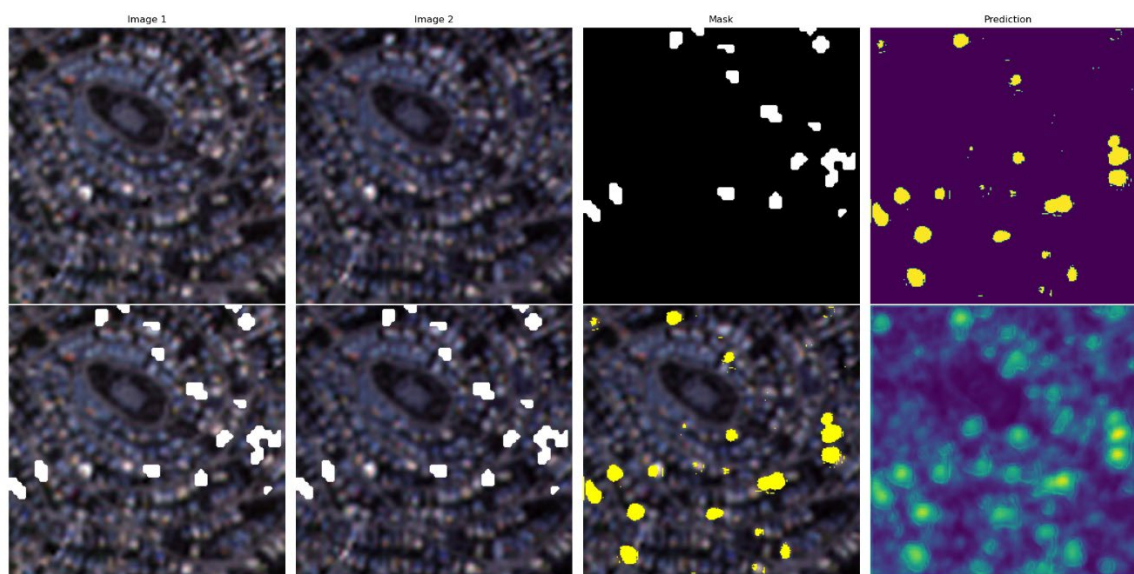


Para evaluar la bondad de los resultados del modelo se ha utiliza como el índice SSIM (Structural similarity index measure). Este método permite evaluar la calidad percibida de una imagen, principalmente es utilizado para medir la similitud de dos imágenes, en este caso, la máscara original y la predicha por el modelo. De esta manera se dispone de un índice que es capaz de cuantificar la calidad de la imagen frente a la generada por el modelo. Se ha utilizado este indicador dado que a diferencia del error cuadrático medio (MSE) considera cambios estructurales en la imagen (Z. Wang et al., 2003).

La figura 11 muestra un buen ejemplo del resultado obtenido con la red. Los cuatro cuadros de la izquierda representan las imágenes en un primer instante (primera columna) y en un segundo instante (segunda columna), ambas con la máscara objetivo superpuesta. La imagen inferior más a la derecha representa el resultado en bruto de la red con una nube de puntos representando la posibilidad de cambio. La imagen justo encima ha sido tratada para solo admitir como cambio aquellos puntos con una probabilidad igual o superior al 40%. En negro se muestra la máscara objetivo por separado y justo debajo la imagen resultante superpuesta sobre una de las entradas.



**Figura 11: Resultados del modelo.**



#### 4. Conclusiones

Existe mucha literatura que avala la utilización de las técnicas de entrenamiento profundo para el procesamiento de imágenes satelitales, como es el caso, entre otros, de realización de mapas viales y detección de vertidos de petróleo.

En este caso particular, la identificación de cambios en el paisaje construido entre dos imágenes de un mismo lugar tomadas en dos instantes de tiempo diferentes permite obtener resultados confiables susceptibles de ser aplicados en el desarrollo de proyectos.

La utilización de estas técnicas permitiría reducir el tiempo y el costo empleados en la búsqueda manual de diferencias entre imágenes. Dejando los recursos humanos destinados a tareas que aportan valor, como la validación o identificación de falsos positivos que genere el modelo de aprendizaje profundo.

El proyecto se consigue realizar únicamente utilizando imágenes multispectrales procedentes de fuentes de libre acceso, como es el caso de Sentinel-2. Estas fuentes abiertas tienen unas limitaciones de la resolución de la imagen a 10 metros por píxel. Lo que dificulta el entrenamiento de los modelos, pero permite una total disponibilidad de la información.

Los buenos resultados obtenidos abren la posibilidad de utilizar esta técnica para la identificación de cambios en la estructura del terreno, como pueden ser plantaciones, los provocados por desastres naturales, movimientos de tierras y demás cambios en la orografía.

Para las pruebas del modelo se han usado imágenes de zonas insulares con una gran cantidad de zonas marinas. En estas pruebas se ha observado que el modelo genera falsos positivos en las superficies del mar, dado que en las imágenes de entrenamiento no se presentaba esta circunstancia. En futuros trabajos si se pretende incorporar zonas marítimas es aconsejable incorporarlas también en el proceso de entrenamiento.

#### 5. Fuentes bibliográficas

Anger, J., Ehret, T., de Franchis, C., & Facciolo, G. (2020). FAST AND ACCURATE MULTI-FRAME SUPER-RESOLUTION OF SATELLITE IMAGES. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-1-2020, 57-64. <https://doi.org/10.5194/isprs-annals-V-1-2020-57-2020>

- Balaniuk, R., Isupova, O., & Reece, S. (2020). *Mining and Tailings Dam Detection In Satellite Imagery Using Deep Learning* (arXiv:2007.01076). arXiv. <https://doi.org/10.48550/arXiv.2007.01076>
- Batra, A., Singh, S., Pang, G., Basu, S., Jawahar, C. V., & Paluri, M. (2019). Improved Road Connectivity by Joint Learning of Orientation and Segmentation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10377-10385. <https://doi.org/10.1109/CVPR.2019.01063>
- Bereta, K., Caumont, H., Daniels, U., Goor, E., Koubarakis, M., Pantazi, D.-A., Stamoulis, G., Ubels, S., Venus, V., & Wahyudi, F. (2019). *The Copernicus App Lab project: Easy Access to Copernicus Data* (Versión 1) [dataset]. [object Object]. <https://doi.org/10.5441/002/EDBT.2019.46>
- Bull, D., Lim, N., & Frank, E. (2021). Perceptual improvements for Super-Resolution of Satellite Imagery. *2021 36th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, 1-6. <https://doi.org/10.1109/IVCNZ54163.2021.9653355>
- Chen, W., Wu, A. N., & Biljecki, F. (2021). Classification of urban morphology with deep learning: Application on urban vitality. *Computers, Environment and Urban Systems*, 90, 101706. <https://doi.org/10.1016/j.compenvurbsys.2021.101706>
- Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, 10(10), 1632-1644. <https://doi.org/10.1111/2041-210X.13256>
- Daudt, R. C., Le Saux, B., Boulch, A., & Gousseau, Y. (2018). Urban Change Detection for Multispectral Earth Observation Using Convolutional Neural Networks. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2115-2118. <https://doi.org/10.1109/IGARSS.2018.8518015>
- Gupta, A., Anpalagan, A., Guan, L., & Khwaja, A. S. (2021). Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. *Array*, 10, 100057. <https://doi.org/10.1016/j.array.2021.100057>
- Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). *EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification* (arXiv:1709.00029; Versión 2). arXiv. <https://doi.org/10.48550/arXiv.1709.00029>
- Hollstein, A., Segl, K., Guanter, L., Brell, M., & Enesco, M. (2016). Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images. *Remote Sensing*, 8(8), Article 8. <https://doi.org/10.3390/rs8080666>
- Jutz, S., & Milagro-Pérez, M. P. (2020). Copernicus: The European Earth Observation programme. *Revista de Teledetección*, 56, Article 56. <https://doi.org/10.4995/raet.2020.14346>
- Kesavaraj, G., & Sukumaran, S. (2013). A study on classification techniques in data mining. *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, 1-7. <https://doi.org/10.1109/ICCCNT.2013.6726842>
- Kolokoussis, P., & Karathanassi, V. (2018). Oil Spill Detection and Mapping Using Sentinel 2 Imagery. *Journal of Marine Science and Engineering*, 6(1), Article 1. <https://doi.org/10.3390/jmse6010004>
- Li, Z., Yan, C., Sun, Y., & Xin, Q. (2022). A Densely Attentive Refinement Network for Change Detection Based on Very-High-Resolution Bitemporal Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-18. <https://doi.org/10.1109/TGRS.2022.3159544>

- Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2018). *Focal Loss for Dense Object Detection* (arXiv:1708.02002). arXiv. <https://doi.org/10.48550/arXiv.1708.02002>
- Lucieer, A., & Stein, A. (2002). Existential uncertainty of spatial objects segmented from satellite sensor imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11), 2518-2521. <https://doi.org/10.1109/TGRS.2002.805072>
- Mao, L., Zheng, Z., Meng, X., Zhou, Y., Zhao, P., Yang, Z., & Long, Y. (2022). Large-scale automatic identification of urban vacant land using semantic segmentation of high-resolution remote sensing images. *Landscape and Urban Planning*, 222, 104384. <https://doi.org/10.1016/j.landurbplan.2022.104384>
- Park, S. C., Park, M. K., & Kang, M. G. (2003). Super-resolution image reconstruction: A technical overview. *IEEE Signal Processing Magazine*, 20(3), 21-36. <https://doi.org/10.1109/MSP.2003.1203207>
- Prinsa, K., & Saritha, E. (2018). Detection and Removal of Clouds on Remote Sensing Images. *2018 4th International Conference for Convergence in Technology (I2CT)*, 1-5. <https://doi.org/10.1109/I2CT42659.2018.9057879>
- Qian, M., Li, Y., Zhao, Y., & Yu, X. (2022). Prior Knowledge-Based Deep Convolutional Neural Networks for Fine Classification of Land Covers in Surface Mining Landscapes. *Sustainability*, 14(19), Article 19. <https://doi.org/10.3390/su141912563>
- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G.-Z. (2017). Deep Learning for Health Informatics. *IEEE Journal of Biomedical and Health Informatics*, 21(1), 4-21. <https://doi.org/10.1109/JBHI.2016.2636665>
- Riabko, A. (2023). Methods of Satellite Images Segmentation Analysis. *2023 IEEE 7th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC)*, 163-167. <https://doi.org/10.1109/MSNMC61017.2023.10329167>
- Shi, W., Zhang, M., Zhang, R., Chen, S., & Zhan, Z. (2020). Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges. *Remote Sensing*, 12(10), Article 10. <https://doi.org/10.3390/rs12101688>
- Singh, A., & Nene, M. J. (2022). Detection of Multiclass Objects in Satellite Images Using an Improved Algorithmic Approach. *2022 IEEE Bombay Section Signature Conference (IBSSC)*, 1-6. <https://doi.org/10.1109/IBSSC56953.2022.10037435>
- Tahir, A., Munawar, H. S., Akram, J., Adil, M., Ali, S., Kouzani, A. Z., & Mahmud, M. A. P. (2022). Automatic Target Detection from Satellite Imagery Using Machine Learning. *Sensors*, 22(3), Article 3. <https://doi.org/10.3390/s22031147>
- Unión Europea. (s. f.). *Sobre Copernicus*. Recuperado 4 de abril de 2024, de <https://www.copernicus.eu/es/sobre-copernicus>
- Vieira, A., & Vieira, A. (2017). *Business Applications of Deep Learning* (business-applications-of-deep-learning) [Chapter]. <https://Services.Igi-Global.Com/Resolvedoi/Resolve.Aspx?Doi=10.4018/978-1-5225-2545-5.Ch003>; IGI Global. <https://www.igi-global.com/gateway/chapter/www.igi-global.com/gateway/chapter/179088>
- Wang, J., Zheng, Z., Ma, A., Lu, X., & Zhong, Y. (2021). *LoveDA: A Remote Sensing Land-Cover Dataset for Domain Adaptive Semantic Segmentation*.
- Wang, T., Shi, J., Letu, H., Ma, Y., Li, X., & Zheng, Y. (2019). Detection and Removal of Clouds and Associated Shadows in Satellite Imagery Based on Simulated Radiance Fields. *Journal of Geophysical Research: Atmospheres*, 124(13), 7207-7225. <https://doi.org/10.1029/2018JD029960>

Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003). Multiscale structural similarity for image quality assessment. *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, 1398-1402. <https://doi.org/10.1109/ACSSC.2003.1292216>

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment, 118*, 83-94. <https://doi.org/10.1016/j.rse.2011.10.028>

**Comunicación alineada con los  
Objetivos de Desarrollo Sostenible**

