



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Geodésica,
Cartográfica y Topográfica

Redes Neuronales Adversariales para la generación
sintética de imágenes satelitales de superresolución en
entornos costeros.

Trabajo Fin de Grado

Grado en Ingeniería Geomática y Topografía

AUTOR/A: Diéguez López, Alejandro

Tutor/a: Colomer Granero, Adrián

Director/a Experimental: CRESPO PEREMARCH, PABLO

CURSO ACADÉMICO: 2022/2023

Agradecimientos

En primer lugar, quiero agradecer a mi familia por su amor incondicional y su apoyo constante a lo largo de mi carrera académica. Vuestras palabras de apoyo y vuestros ánimos han sido mi fuente de inspiración y motivación en los momentos más duros. Gracias por estar siempre ahí.

A Adriana, mi compañera de vida y mi mayor apoyo, quiero agradecerte por tu paciencia, comprensión y por creer en mí, incluso cuando yo mismo dudaba, tus palabras y tu cariño fueron fundamentales.

También quiero agradecer a mis compañeros de estudio, sobre todo a Luis, Raúl y Andreea, que han estado a mi lado durante toda esta etapa universitaria. Vuestra amistad ha sido un pilar fundamental en mi vida y vuestro apoyo incondicional ha sido incalculable. Gracias por las largas noches de estudio juntos, las risas compartidas y por grandes recuerdos que nunca olvidaré.

Asimismo, deseo expresar mi agradecimiento a todo el personal docente que me ha ayudado en este camino, sobre todo a mis tutores Adrián y Pablo, cuya guía y conocimientos han sido esenciales para el desarrollo de este trabajo. Vuestra dedicación y orientación me han ayudado a crecer como estudiante y como profesional.

Por último, mi más sincero agradecimiento al equipo del CVB Lab, y en especial a Valery, por brindarme la increíble oportunidad de formar parte del equipo.

Compromiso

'El presente documento ha sido realizado completamente por el firmante; no ha sido entregado como otro trabajo académico previo y todo el material tomado de otras fuentes ha sido convenientemente entrecomillado y citado su origen en el texto, así como referenciado en la bibliografía'

Valencia, 06/07/2023

A handwritten signature in black ink, appearing to be 'Adrián', written over a horizontal line.

Resumen

En los últimos tiempos, debido a la afección del cambio climático y a otros procesos antrópicos, muchas zonas costeras del planeta están sufriendo cambios continuos que provocan variaciones constantes en las líneas de costa.

Para poder actuar ante estas variaciones costeras es necesario monitorizar las zonas que sufren estos sucesos para poder tomar las decisiones correctas. Una forma de realizar dicha monitorización es la implementación de diversas técnicas de teledetección. Estas técnicas permiten la extracción de información de la superficie terrestre y de cuerpos de agua sin estar en contacto con ellos; por ejemplo, mediante imágenes satelitales o aéreas. No obstante, existen una serie de desventajas como la resolución espacial de las imágenes satelitales disponibles que, en algunos casos, limita la precisión con la que se pueden obtener los resultados. Por esta razón, la premisa de este proyecto es generar dichas imágenes satelitales con una resolución espacial mayor, permitiendo así realizar un análisis de la línea de costa con una mayor precisión comparado al que se hubiese podido efectuar con las imágenes originales.

Para llevar a cabo esta tarea, se emplearán diferentes metodologías fundamentadas en aprendizaje profundo para aumentar la resolución espacial. En este sentido, las metodologías de deep learning, son capaces de sintetizar imágenes digitales de alta resolución a partir de imágenes con un menor número de píxeles.

Concretamente, en este TFG se explorarán técnicas basadas en arquitecturas de redes neuronales profundas que utilizan conexiones residuales para mejorar la calidad de las imágenes de baja resolución. El objetivo principal es desarrollar un modelo capaz de aumentar la resolución de imágenes satelitales de baja a media resolución, manteniendo la homogeneidad estructural de la imagen.

La consecución del presente trabajo pretende establecer una línea de investigación basada en imagen satelital de alta resolución, en la que se abordarán tareas de análisis posteriores como la detección de estructuras características o la segmentación de objetos de interés.

Resum

En els últims temps, a causa de l'afecció del canvi climàtic i a altres processos antròpics, moltes zones costaneres del planeta estan patint canvis continus que provoquen variacions constants en les línies de costa.

Per a poder actuar davant aquestes variacions costaneres és necessari monitorar les zones que pateixen aquests successos per a poder prendre les decisions correctes. Una manera de realitzar aquest monitoratge és la implementació de diverses tècniques de teledetecció. Aquestes tècniques permeten l'extracció d'informació de la superfície terrestre i de cossos d'aigua sense estar en contacte amb ells; per exemple, mitjançant imatges satel·litàries o aèries. No obstant això, existeixen una sèrie de desavantatges com la resolució espacial de les imatges satel·litàries disponibles que, en alguns casos, limita la precisió amb la qual es poden obtenir els resultats. Per aquesta raó, la premissa d'aquest projecte és generar aquestes imatges satel·litàries amb una resolució espacial major, permetent així realitzar una anàlisi de la línia de costa amb una major precisió comparat al que s'haguera pogut efectuar amb les imatges originals.

Per a dur a terme aquesta tasca, s'empraran diferents metodologies fonamentades en aprenentatge profund per a augmentar la resolució espacial. En aquest sentit, les metodologies de deep learning, són capaces de sintetitzar imatges digitals d'alta resolució a partir d'imatges amb un menor nombre de píxels.

Concretament, en aquest TFG s'exploraran tècniques basades en arquitectures de xarxes neuronals profundes que utilitzen connexions residuals per a millorar la qualitat de les imatges de baixa resolució. L'objectiu principal és desenvolupar un model capaç d'augmentar la resolució d'imatges satel·litàries de baixa a mitja resolució, mantenint l'homogeneïtat estructural de la imatge.

La consecució del present treball pretén establir una línia d'investigació basada en imatge satel·litària d'alta resolució, en la qual s'abordaran tasques d'anàlisi posteriors com la detecció d'estructures característiques o la segmentació d'objectes d'interès.

Abstract

In recent times, due to the effects of climate change and other anthropogenic processes, many coastal areas of the planet are undergoing continuous changes that cause constant variations in coastlines.

In order to be able to act in the face of these coastal variations, it is necessary to monitor the areas that suffer these events to be able to make the right decisions. One way to carry out such monitoring is the implementation of various remote sensing techniques. These techniques allow the extraction of information from the land surface and water bodies without being in contact with them, e.g., through satellite or aerial imagery. However, there are a number of disadvantages, such as the spatial resolution of the available satellite images, which in some cases limits the accuracy with which results can be obtained. For this reason, the premise of this project is to generate such satellite images with a higher spatial resolution, allowing a more accurate analysis of the coastline than would have been possible with the original images.

To carry out this task, different methodologies based on deep learning will be used to increase the spatial resolution. In this sense, deep learning methodologies are capable of synthesising high-resolution digital images with images of a smaller number of pixels.

Specifically, this TFG will explore techniques based on deep neural network architectures that use residual connections to improve the quality of low-resolution images. The main objective is to develop a model capable of increasing the resolution of low to medium resolution satellite images while maintaining the structural homogeneity of the image.

The aim of this work is to establish a line of research based on high-resolution satellite images, in which subsequent analysis tasks such as the detection of characteristic structures or the segmentation of objects of interest will be tackled.

Índice de figuras

1.	Comparación de imagen de baja resolución y alta resolución	2
2.	Ejemplo de arquitectura CNN	8
3.	Arquitectura EDSR	9
4.	Arquitectura RDN	10
5.	Proceso de <i>Forward</i> y <i>backward</i>	11
6.	Remuestreo vecino más cercano	18
7.	Remuestreo Interpolación bilineal	18
8.	Partición del conjunto de datos OLI2MSI	20
9.	Huella del conjunto de datos OLI2MSI	21
10.	Ejemplo de algunos pares de imágenes del conjunto de datos OLI2MSI	21
11.	Ejemplo de algunos pares de imágenes del conjunto de datos propio .	22
12.	Huella del conjunto de datos propio	23
13.	Partición del conjunto de datos propio	24
14.	Diagrama de flujo metodológico	27
15.	Ejemplo de máscara binaria obtenida mediante el modelo de segmen- tación	34
16.	Matriz de confusión	35
17.	Proceso para obtener el área acotada en la zona de cambio	37
18.	Resultado del conjunto de datos OLI2MSI	40
19.	Prueba del modelo en una imagen de la Albufera	41
20.	Resultados obtenidos con el conjunto de datos propio	43
21.	Mascaras de segmentación de la línea de costa	44
22.	Segmentación de la línea de costa en la zona acotada	46

Índice de tablas

1.	Resultados en los índices PSNR y SSIM para el conjunto de datos OLI2MSI	5
2.	Bandas del satélite Landsat-8	15
3.	Bandas del satélite Sentinel-2	16
4.	Características del dataset OLI2MSI	19
5.	Dimensiones de la red EDSR con factor de escala de <i>upsampling</i> de 3	28
6.	Dimensiones de la red EDSR con factor de escala de <i>upsampling</i> de 4	30
7.	Dimensiones de las capas del modelo RDN con factor de escala 3. . .	31
8.	Valores medios de SSIM y PSNR con el modelo EDSR.	38
9.	Valores medios de SSIM y PSNR con el modelo RDN.	38
10.	Valores medios de SSIM y PSNR dataset propio	42
11.	Comparación de los índices Dice y Kappa para las diferentes combinaciones de imágenes.	44
12.	Comparación de los índices Dice y Kappa para las diferentes combinaciones de imágenes en la zona acotada.	45
13.	Costes de personal	50
14.	Coste de hardware	51

Índice

1. Introducción	1
1.1. Motivación y explicación de la adversidad	1
1.2. Estado del arte	3
2. Objetivos	5
2.1. Objetivo principal	5
2.2. Objetivos secundarios	5
3. Marco teórico	6
3.1. Redes Neuronales Artificiales	6
3.2. Redes Neuronales Convolucionales	6
3.3. EDSR	8
3.4. RDN	9
3.5. Fase de entrenamiento	10
3.6. Parámetros de entrenamiento	12
4. Datos	14
4.1. Satélites y sensores	14
4.1.1. Landsat-8	14
4.1.2. Sentinel-2	15
4.1.3. Ortoimágenes ICV	16
4.2. Métodos en el preprocesado de las imágenes	16
4.2.1. Reproyección	17
4.2.2. Remuestreo	17
4.3. Bases de datos	19
4.3.1. OLI2MSI	19
4.3.2. Base de datos propia	21
4.3.3. Base de datos conjunta	24
4.4. Entorno de desarrollo	25
4.4.1. Librerías más empleadas	25
4.4.2. Otros software utilizados	25
5. Metodología	27
5.1. Diagrama de flujo	27
5.2. Implementación de los modelos	28
5.2.1. EDSR aplicado en conjunto de datos OLI2MSI	28
5.2.2. EDSR aplicado en <i>dataset</i> propio	29
5.2.3. RDN aplicado al <i>dataset</i> conjunto	30
5.3. Evaluación de la calidad de la imagen sintética de alta resolución ge- nerada	32
5.3.1. SSIM	32
5.3.2. PSNR	33
5.4. Evaluación mediante la segmentación de la línea costera	33

5.4.1. Matriz de confusión	34
5.4.2. DICE	36
5.4.3. KAPPA	36
5.4.4. Cálculo de los índices acotados al área de la línea de costa . .	36
6. Resultados	38
6.1. Resultados OLI2MSI	38
6.2. Resultados Conjunto de datos propio	41
6.2.1. Análisis de imágenes superresueltas	41
6.2.2. Análisis de línea de costa	43
6.2.3. Análisis en la zona acotada a la línea de costa	45
6.3. Análisis comparativo entre otros artículos científicos	47
7. Conclusiones	48
8. Presupuestos	50
8.1. Partida de personal	50
8.2. Partida de gastos informáticos	50
8.3. Total	51
ANEXO I	55
ANEXO II	57

Listado de siglas

HR: Imagen de alta resolución
LR: Imagen de baja resolución
RN: Redes Neuronales
EDSR: Enhanced Deep Super-Resolution
RDN: Residual Dense Network
CNN: Red Neuronal Convolutacional
VDSR: Very Deep Super-Resolution
SRCNN: Super-Resolution Convolutional Neural Network
SRGAN: Super-Resolution Generative Adversarial Network
OLI: Operational Land Imager
NIR: Infrarrojo cercano
RGB: Bandas del espectro visible (rojo, verde, azul)
MSE: Error cuadrático medio
SSIM: Índice de Similitud Estructural
PSNR: Relación Señal-Ruido Pico
ICV: Institut Cartogràfic Valencià

Sección 1

1. Introducción

1.1. Motivación y explicación de la adversidad

Durante los últimos años, debido a diversas causas como el cambio climático y procesos antrópicos, se han producido importantes variaciones en las zonas costeras, lo que ha alterado la línea de costa. Estas variaciones representan una gran problemática, ya que las ciudades costeras y las viviendas ubicadas en estas áreas se ven cada vez más afectadas, ocasionando daños significativos a las ciudades y poniendo en peligro la seguridad de los residentes y sus propiedades.

Además, los espacios naturales también se ven comprometidos debido a los cambios en la línea de costa. Un ejemplo de ello es la Albufera, donde el aumento del nivel del mar y la pérdida de arena en las playas pueden afectar la estabilidad de este valioso hábitat, poniendo en riesgo la diversidad biológica y alterando el equilibrio natural del ecosistema.

Es igualmente importante mencionar que el turismo de playa constituye una fuente vital de ingresos para España. Las playas españolas son destinos populares tanto para turistas nacionales como internacionales, y la calidad y belleza de estas playas son fundamentales para mantener su atractivo turístico. Sin embargo, debido a los problemas mencionados anteriormente, las playas se ven afectadas negativamente, lo que requiere la implementación de medidas como los trasvases de arena para preservar su aspecto y funcionalidad. Estos trasvases implican la redistribución de arena desde áreas donde se acumula hacia áreas donde se ha perdido, con el objetivo de asegurar el disfrute de las playas por parte de los turistas.

Por todas estas razones, es cada vez más necesario llevar a cabo un mapeo exhaustivo de estas zonas para poder actuar de manera rápida y efectiva. Para realizar este mapeo existen diferentes formas, una de las más utilizadas según el estado del arte se basa en aplicar técnicas de teledetección sobre imágenes satelitales.

Dichas imágenes satelitales presentan una gran cantidad de características que aportan valor a dicha metodología. Una de las características destacadas es el registro periódico de las imágenes satelitales, lo que permite realizar un seguimiento continuo de los cambios en la línea de costa a lo largo del tiempo. Otra característica importante es la cobertura global de las imágenes satelitales, lo que significa que se pueden obtener imágenes de prácticamente cualquier parte del planeta. Además, es relevante mencionar que muchas imágenes satelitales están disponibles como datos de acceso libre.

Aunque dicha metodología presenta diversos inconvenientes, entre los cuales podemos destacar la resolución espacial que presentan las imágenes satelitales que pueden sesgar la precisión con la cual obtenemos los resultados. Por esta razón, se pretende mejorar la resolución espacial de las imágenes satelitales basándonos en Redes Neuronales (RN).

Las RN son modelos de procesamiento de información que consisten en un conjunto interconectado de unidades de procesamiento llamadas neuronas artificiales. Estas neuronas están organizadas en capas y se comunican entre sí mediante conexiones ponderadas. Utilizando algoritmos de aprendizaje, las redes neuronales pueden entrenarse para reconocer patrones complejos, realizar predicciones, clasificar datos y llevar a cabo diversas tareas de procesamiento de información. Su estructura y funcionamiento se inspiran en el cerebro humano, permitiendo que las redes neuronales sean capaces de aprender de manera autónoma y adaptarse a diferentes situaciones. Un ejemplo es la aplicación de la superresolución de imágenes satélites, donde la RN es capaz de aprender a generar imágenes de alta resolución (HR) introduciendo como *input* del modelo una de baja resolución (LR).

Como se puede observar en la Figura 1, se aprecian claramente las diferencias de resolución espacial entre la imagen LR (10 metros por píxel) y HR (3,33 metros por píxel). Es importante tener en cuenta que las imágenes HR no tienen la misma disponibilidad que las imágenes satelitales de resolución media. No obstante, la implementación de las RN para generar imágenes HR posibilita conseguir la cobertura global y la baja resolución temporal, junto con una resolución espacial mejorada. Esto brinda una herramienta valiosa para el mapeo detallado y el seguimiento de cambios en zonas costeras con mayor precisión.

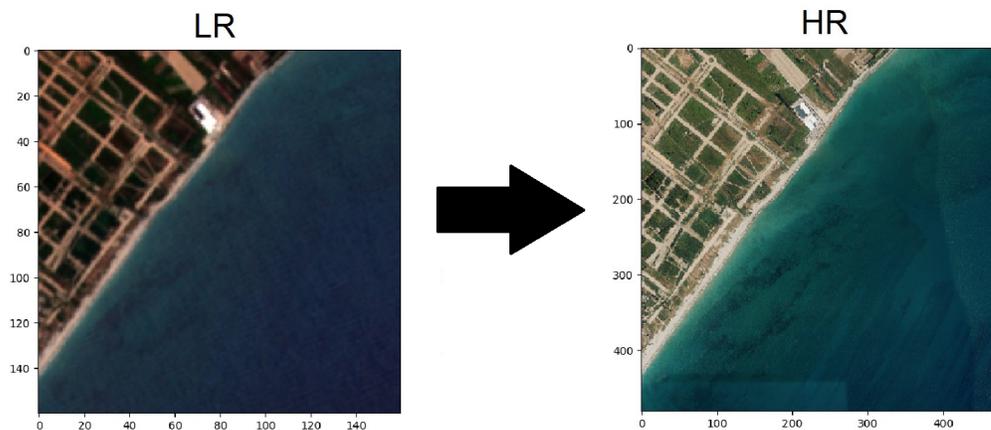


Figura 1: Comparación de imagen de baja resolución y alta resolución. Fuente: Elaboración propia

1.2. Estado del arte

La idea de superresolución de imágenes se remonta algunas décadas atrás cuando se empezó a realizar investigaciones de cómo mejorar la calidad de las imágenes mediante diferentes métodos basados en la reducción de ruido y fusión de imágenes, un ejemplo de esto es el artículo científico *Fast and robust multiframe super resolution* [1] donde se plantean implementaciones clásicas para abordar la tarea de superresolución.

Con el transcurso del tiempo se han ido proponiendo nuevas propuestas sobre cómo abordar la mejora de la calidad en todo tipo de imágenes. Hasta la actualidad, donde las técnicas que están más presentes en estado del arte se fundamentan en enfoques de aprendizaje profundo como las que se explican en el artículo [2].

En los últimos años, se han desarrollado varias arquitecturas de RN profundas para la tarea de superresolución aplicada a todo tipo de imágenes. Estas arquitecturas han demostrado ser capaces de producir resultados de alta calidad en la reconstrucción de imágenes de baja resolución y media resolución espacial. Algunos ejemplos de los modelos más empleados en este campo incluyen *Super-Resolution Convolutional Neural Network* (SRCNN) [3], *Very Deep Super-Resolution* (VDSR) [4], *Super-Resolution Generative Adversarial Network* (SRGAN) [5], *Enhanced Deep Super-Resolution* (EDSR) [6] o *Residual Dense Network* (RDN) [7].

La red SRCNN presentada en el artículo [3] emplea una red neuronal convolucional la cual consta de tres etapas principales que son la extracción de características donde se utilizan varias capas de convolución para obtener caracteres relevantes de la imagen LR. Después viene la etapa de mapeo no lineal donde se aplica una activación no lineal posterior al resultado de las capas convolucionales anteriores. Finalmente, se aplica la etapa de reconstrucción en la imagen HR partiendo de las características extraídas en las etapas pasadas. Dicha red fue implementada con imágenes de un *dataset* público llamado *ImageNet* el cual está compuesto por todo tipo de imágenes.

Por otro lado, la VDSR utiliza una red neuronal profunda con múltiples capas convolucionales, a diferencia de otras RN que están pensadas para la tarea de superresolución, la VDSR se centra en encontrar diferencias entre las imágenes HR y LR, como se explica en [4]. También está compuesta por tres etapas que son la extracción de características muy similar a la comentada anteriormente en [3]. Después pasamos a una etapa residual enfocada en detectar y aprender diferencias residuales entre las imágenes HR y LR, por último, la etapa de reconstrucción también muy similar a la red comentada anteriormente. Como en el caso del artículo [3] también implementaron la red con el *dataset* público *ImageNet*.

En cuanto a la red SRGAN, introduce la idea de implementar una red neuronal generativa adversarial (GAN) para producir imágenes HR. Dicha red está compuesta

por dos componentes que son el Generador donde se generan las imágenes de HR partiendo de una imagen de LR. Dicho componente está compuesto por varias capas convolucionales, donde cada capa realiza una transformación no lineal para aumentar gradualmente la resolución y los detalles de la imagen. Además, se emplean bloques residuales para capturar las diferencias entre la imagen HR y LR, permitiendo que el generador aprenda las características y detalles finos que faltan en la imagen LR. Por otra parte, el Discriminador es una red neuronal convolucional que tiene como objetivo distinguir entre imágenes HR reales y las imágenes SR creadas por el modelo generador. Se entrena para clasificar correctamente las imágenes HR reales y las generadas por el modelo. A medida que el generador mejora, el discriminador se vuelve más efectivo según el artículo [5]. Por otra parte, esta red fue implementada con los *datasets* públicos *div2k* y *Flickr2K* formados por todo tipo de imágenes.

La arquitectura EDSR [6] es una variante de la arquitectura comentada anteriormente VDSR que emplea un enfoque de red residual profunda, pero tienen diversas diferencias en su enfoque y diseño, entre estas se destacan las diferencias en la profundidad de la red. Ya que el modelo VDSR lleva este enfoque a un nivel más extremo al ser una red más profunda. De igual modo, la arquitectura de bloques residuales también es una diferencia notable a causa de que la red EDSR usa conexiones residuales para aprender la diferencia residual entre las imágenes HR y LR. Estos bloques residuales permiten que la red se enfoque en los detalles y patrones finos que faltan en la imagen LR. En la implementación de este modelo se emplearon diversas bases de datos, algunas son: *Set 5*, *Set14*, *Urban100* y *BSDS100*.

Finalmente, la RDN [7] es una arquitectura de red neuronal profunda utilizada para tareas de superresolución de imágenes mediante la entrada de dos pares de imágenes (HR y LR). La red se compone de bloques densos residuales que permiten la extracción y fusión de características de diferentes escalas en la imagen. Estos bloques densos utilizan conexiones residuales para mejorar el flujo de información a través de la red y reducir la degradación causada por el entrenamiento profundo. Para la implementación de este modelo se emplearon los mismos conjuntos de datos que en el caso anterior, es decir, *Set 5*, *Set14*, *Urban100* y *BSDS100*.

Todas las arquitecturas anteriores también fueron implementadas en el artículo [8] donde se utilizó un *dataset* creado en el mismo artículo llamado OLI2MSI el cual explicaremos con mayor profundidad en la siguiente sección del Trabajo Fin de Grado, pero está compuesto por pares de imágenes satelitales siendo las imágenes LR las captadas por el satélite Landsat-8 y las imágenes HR las captadas por el Sentinel-2. En dicho artículo se muestran los resultados que se obtuvieron al aplicar los diferentes modelos expuestos anteriormente, tomando como datos de entrada el *dataset* OLI2MSI. Dichos resultados se muestran en la Tabla 1.

Algoritmo	PSNR (OLI2MSI)	SSIM (OLI2MSI)
SRCNN	35.26	0.907
VDSR	35.87	0.919
SRGAN	33.63	0.879
EDSR	37.68	0.946

Tabla 1: Resultados en los índices PSNR y SSIM para el conjunto de datos OLI2MSI

En general, la aplicación de RN para la superresolución de imágenes de satélite ha demostrado ser una técnica efectiva y prometedora para mejorar la calidad de las imágenes de baja o media resolución.

2. Objetivos

2.1. Objetivo principal

El objetivo fundamental del Trabajo Fin de Grado es el diseño y desarrollo de un *framework* basado en técnicas de aprendizaje profundo que permita la generación sintética de imágenes satelitales de superresolución. Cabe contextualizar que dichas imágenes de superresolución generadas se emplearán posteriormente para realizar diversos análisis en zonas costeras, como puede ser la delimitación de línea de costa.

2.2. Objetivos secundarios

Del objetivo principal definido anteriormente, surgen una serie de objetivos específicos que permitirán acometer el presente Trabajo Final de Grado con éxito:

- Desarrollar, entrenar, optimizar y evaluar un modelo de superresolución basado en técnicas de aprendizaje profundo para mejorar la resolución espacial de imágenes Landsat-8 (30 metros por píxel) tomando como imágenes de referencia tomas del satélite Sentinel-2 (10 metros por píxel).
- Desarrollar, entrenar, optimizar y evaluar un modelo de superresolución basado en técnicas de aprendizaje profundo para mejorar la resolución espacial de imágenes Sentinel-2 (10 metros por píxel) en zonas costeras, tomando como imágenes de referencia tomas de alta resolución (3,33 metros por píxel).
- Evaluar los tres tipos de imágenes HR, LR y SR en entornos costeros para cuantificar las variaciones de la línea de costa entre ellas.

3. Marco teórico

El Aprendizaje Profundo es una rama de la inteligencia artificial que se enfoca en el procesamiento, tratamiento y entrenamiento de conjuntos de datos utilizando RN. Cabe destacar, que las RN como se ha descrito en la introducción, tienen como premisa obtener modelos que sean capaces de aprender a interpretar situaciones y saber cómo actuar ante éstas. Un ejemplo de esto en el campo de imagen es la utilización de dichas RN para obtener imágenes de superresolución.

A continuación se explican en más detalle las redes que se emplearon y cómo funcionan los entrenamientos en la RN y los parámetros de los mismos a la hora de entrenar las redes para desarrollar este Trabajo Fin de grado.

3.1. Redes Neuronales Artificiales

Las redes neuronales artificiales son un tipo de modelo de aprendizaje automático inspirado en el funcionamiento del cerebro humano. Estas redes están compuestas por unidades básicas llamadas neuronas artificiales, que están interconectadas entre sí para procesar información y realizar tareas específicas. Cada neurona artificial recibe entradas, las procesa mediante una función de activación y produce una salida. Estas salidas pueden ser utilizadas como entradas para otras neuronas, permitiendo la construcción de redes complejas.

La estructura de una red neuronal artificial se organiza en capas, que son conjuntos de neuronas interconectadas. La primera capa se denomina capa de entrada y recibe los datos iniciales, mientras que la última capa se llama capa de salida y produce los resultados deseados. Entre la capa de entrada y la capa de salida, pueden existir una o más capas ocultas, que son responsables de procesar y transformar los datos a medida que se propagan a través de la red.

El entrenamiento de una red neuronal artificial implica ajustar los pesos y los sesgos de las conexiones entre las neuronas para que la red pueda realizar predicciones precisas o clasificar datos de entrada correctamente. Esto se logra mediante un proceso llamado retropropagación del error, en el cual se comparan las salidas producidas por la red con los valores deseados y se ajustan los parámetros de la red para reducir la diferencia entre ellos.

3.2. Redes Neuronales Convolucionales

En primer lugar, se explican las redes neuronales convoluciones, las cuales no se han empleado como tal directamente, pero son la base de las dos arquitecturas empleadas.

Una red neuronal convolucional (CNN) [9] es un tipo de arquitectura de aprendizaje profundo diseñada específicamente para el procesamiento de datos estructurados, como imágenes. Estas RN son capaces de aprender patrones y características relevantes directamente de los datos de imagen, sin necesidad de una extracción manual de características.

Las CNN están compuestas por múltiples capas, incluyendo capas convolucionales, capas de agrupación (pooling) y capas totalmente conectadas. Las capas convolucionales son el componente principal de las CNN y consisten en neuronas que están conectadas solo a una pequeña región de la entrada. En lugar de estar conectadas a todas las entradas, como en una capa totalmente conectada, estas neuronas realizan operaciones de convolución.

La convolución implica deslizar un filtro o kernel sobre la imagen de entrada para extraer características locales. Cada filtro aprende a detectar características específicas, como bordes, texturas o formas. Estos filtros se entrenan mediante el proceso de aprendizaje de la red neuronal, y a medida que avanza la información a través de las capas convolucionales, se van detectando características más complejas y de mayor nivel de abstracción.

Después de la operación de convolución, se aplica una función de activación no lineal, como la función ReLU, para introducir no linealidad en la red. Esto permite que la red aprenda representaciones más complejas y capture relaciones no lineales entre las características extraídas.

Las capas de agrupación *pooling* se utilizan para reducir la dimensionalidad de las características y proporcionar una invariancia a pequeñas traslaciones y deformaciones en la imagen. La capa de pooling toma una región de las características y reduce su tamaño tomando el valor máximo *max pooling* o promedio *average pooling* de los elementos en esa región.

Finalmente, las capas totalmente conectadas se utilizan para tomar las características extraídas y realizar la clasificación o la predicción final. Estas capas están completamente conectadas entre sí y utilizan métodos de aprendizaje supervisado para mapear las características a las clases o categorías deseadas.

Todas las características comentadas anteriormente se pueden observar en la Figura 2 que muestra un ejemplo de implementación de una red CNN.

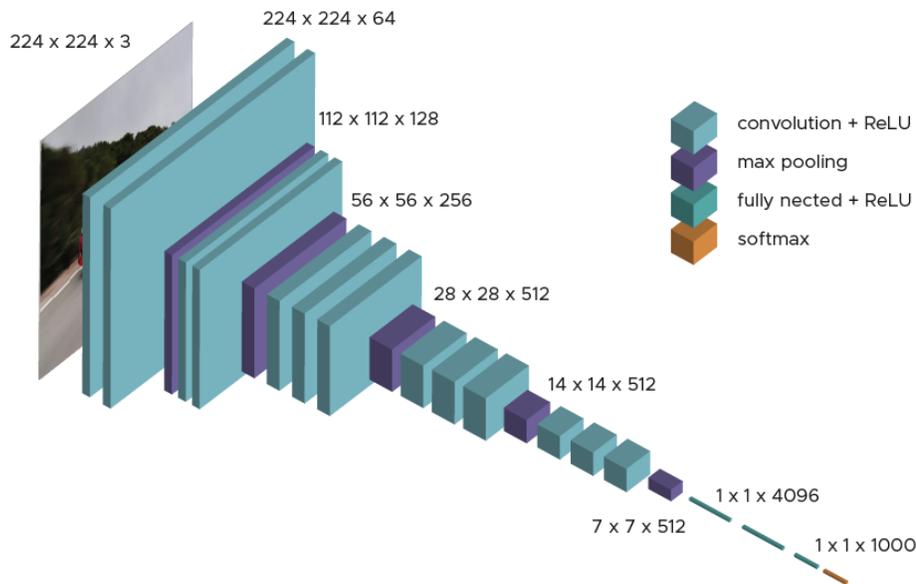


Figura 2: Ejemplo de arquitectura CNN. Fuente[10]

3.3. EDSR

El modelo EDSR [6] es una arquitectura basada en una CNN diseñada específicamente para abordar el problema de la superresolución de imágenes, es decir, aumentar la resolución de una imagen de baja resolución para obtener una versión de alta resolución.

El modelo EDSR se basa en la idea de utilizar una RN profunda para aprender una función de mapeo directo entre imágenes LR y HR. A diferencia de los métodos tradicionales de interpolación o técnicas de filtrado, el enfoque de aprendizaje profundo permite capturar relaciones más complejas y realizar una reconstrucción más precisa de los detalles de la imagen.

El modelo EDSR utiliza una arquitectura residual, lo que significa que utiliza bloques residuales que permiten que la red aprenda las diferencias entre la imagen de baja resolución y la imagen de alta resolución de destino. Cada bloque residual consiste en múltiples capas convolucionales y una conexión de salto *skip connection* que agrega la entrada original a la salida del bloque. Esta conexión de salto permite que la información original se preserve y ayude en la propagación del gradiente durante el entrenamiento.

Dichas capas residuales, junto con la capa de *upsampling* del modelo, están compuestas por capas de convolución y capas *shuffle*. Estas capas trabajan en conjunto para aumentar el tamaño y reconstruir la imagen, logrando así obtener la imagen SR.

Una vez entrenado, el modelo es capaz de tomar una imagen LR y generar una versión de alta resolución mediante la aplicación de la función de mapeo aprendida.

Todas las características comentadas se pueden observar esquematizadas en la Figura 3 donde se muestra la arquitectura del modelo.

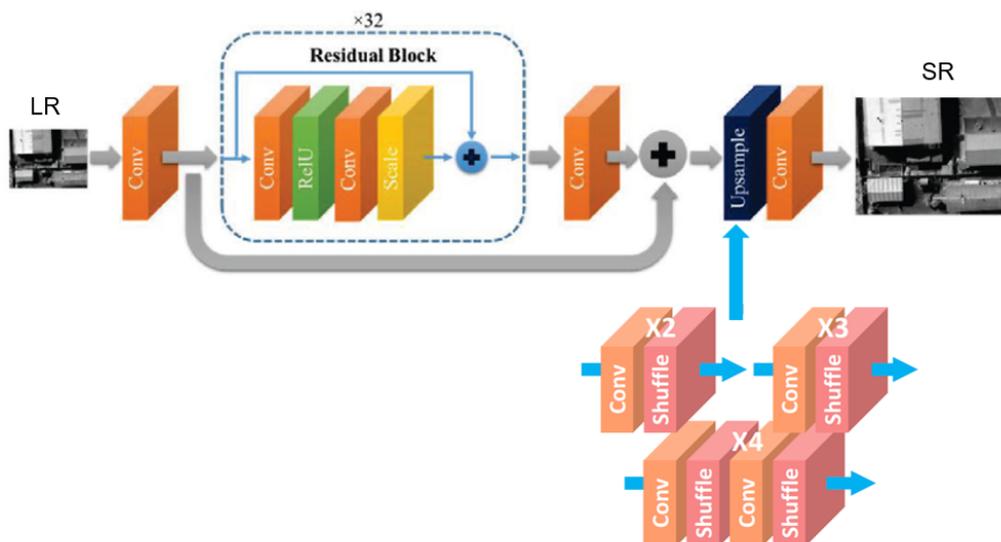


Figura 3: Arquitectura EDSR. Fuente: Elaboración propia

3.4. RDN

La (RDN) [7] es un modelo de red neuronal profunda utilizado para el problema de superresolución de imágenes.

La RDN se basa en la arquitectura de redes residuales, que permite el entrenamiento eficiente de RN profundas. Estas redes residuales están compuestas por bloques residuales, que son unidades básicas de procesamiento. Cada bloque residual tiene una conexión de salto *skip connection* que permite que la información se propague directamente desde la entrada a la salida del bloque. Esta conexión de salto evita problemas degradantes, como el desvanecimiento del gradiente, que pueden surgir en RN profundas.

La RDN extiende la arquitectura residual al introducir bloques residuales densos. En lugar de tener una única conexión de salto en cada bloque residual, la RDN conecta cada capa en un bloque con todas las capas siguientes. Esta estructura densa permite que la información fluya eficientemente a través de la red y facilita la fusión de características en diferentes niveles de abstracción.

Además de los bloques residuales densos, la RDN también utiliza capas de convolución para extraer características de la imagen de baja resolución. Estas capas convolucionales realizan la convolución de los datos de entrada con un conjunto de filtros, capturando patrones y detalles en diferentes escalas. La RDN utiliza múltiples capas convolucionales en su arquitectura para capturar características de manera progresiva y jerárquica.

La estructura comentada se puede observar esquematizada en la Figura 4, donde se muestra de forma gráfica el esquema del modelo y cómo están conectadas las diferentes capas entre sí.

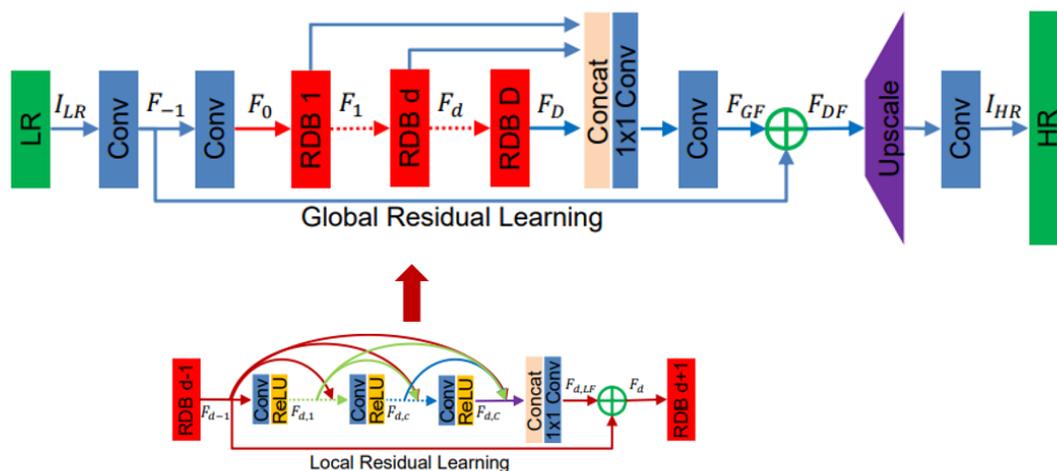


Figura 4: Arquitectura RDN. Fuente: Elaboración propia

3.5. Fase de entrenamiento

Una vez tenemos los modelos establecidos, es fundamental explicar la etapa de entrenamiento para aplicarla correctamente a los modelos que hemos expuesto anteriormente.

Dicho proceso de entrenamiento de una red neuronal se compone de dos etapas fundamentales: *Forward pass* (propagación hacia adelante) y *back-propagation* (retropropagación) [11]. Estos procesos están interconectados y permiten que la red neuronal aprenda a partir de los conjuntos de datos destinados al entrenamiento.

En la etapa de *forward pass*, los datos de entrenamiento se propagan a través de la red neuronal, capa por capa, desde la capa de entrada hasta la capa de salida, generando una predicción teniendo en cuenta, que cada neurona en una capa está conectada a las neuronas de la capa siguiente mediante conexiones ponderadas.

El proceso se inicia con la asignación de los datos de entrada a la capa de entrada de la red neuronal. Luego, se realiza un cálculo en cada neurona de la capa siguiente, donde se suma la ponderación de las conexiones multiplicada por la salida de las neuronas de la capa anterior. A esta suma se le aplica una función de activación, que introduce no linealidades en la red y le permite aprender representaciones más complejas. Este proceso se repite capa por capa hasta llegar a la capa de salida, donde se genera la predicción o salida de la red neuronal.

Durante la propagación hacia delante, se almacenan los valores de salida de cada neurona, ya que serán utilizados en la etapa de *back-propagation*.

Una vez que se ha generado la predicción de la red neuronal en la etapa de *forward pass*, se procede a compararla con el valor verdadero (etiqueta) correspondiente al dato de entrenamiento.

El objetivo del *back-propagation* es calcular la contribución de cada peso de la red en el error de la predicción, y utilizar esta información para ajustar los pesos y mejorar el rendimiento de la red. El proceso de *back-propagation* se inicia calculando el error entre la predicción de la red y el valor verdadero mediante una función de pérdidas. Luego, este error se propaga hacia atrás a través de la red, capa por capa, calculando el gradiente del error con respecto a los pesos de las conexiones. En cada capa, se calcula el gradiente mediante la regla de la cadena, multiplicando el gradiente recibido de la capa siguiente por la derivada de la función de activación de la capa actual. Estos gradientes se utilizan para actualizar los pesos de las conexiones mediante un algoritmo de optimización.

Este proceso de retropropagación se repite para cada dato de entrenamiento, y la red neuronal va ajustando gradualmente sus pesos para minimizar el error en las predicciones y mejorar su capacidad de generalización. En la Figura 5 se observa un ejemplo gráfico de como se produce el proceso del *forward* y *backward*.

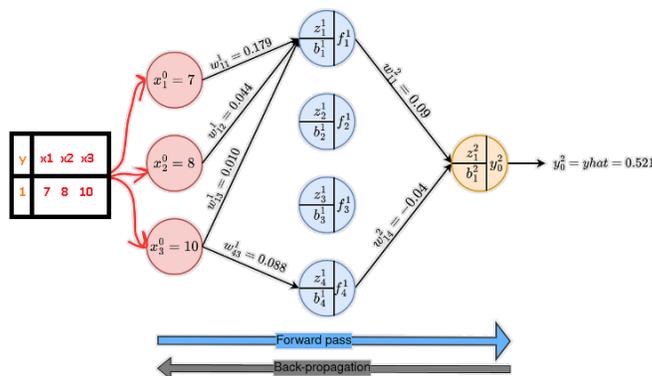


Figura 5: Proceso de *Forward* y *backward*. Fuente: Elaboración propia

3.6. Parámetros de entrenamiento

También es fundamental para realizar un buen entrenamiento de los modelos, optimizar al máximo los parámetros necesarios para entrenar la red, los cuales vamos a explicar a continuación.

1. Optimizador de la red:

Los optimizadores de la red son unos algoritmos que se utilizan para ajustar los parámetros del modelo con el fin de minimizar la función de pérdida o maximizar la función objetivo. El objetivo principal del optimizador es encontrar los valores óptimos para los pesos y sesgos de la red neuronal que permitan una mejor capacidad de predicción o generación de resultados deseados.

Para la aplicación de la superresolución de imágenes aplicada a redes neuronales, el optimizador *ADAM* [12] se destaca como una opción popular debido a su efectividad en el aprendizaje profundo y su capacidad para adaptar la tasa de aprendizaje durante el entrenamiento.

El algoritmo *ADAM* combina las ventajas del descenso de gradiente estocástico y el ajuste adaptativo de la tasa de aprendizaje. Utiliza estimaciones de momentos de primer y segundo orden del gradiente para actualizar los parámetros del modelo de manera más eficiente. Estos momentos estimados se utilizan para ajustar la tasa de aprendizaje para cada parámetro en función de su importancia y el historial de actualizaciones anteriores.

2. Tasa de aprendizaje:

La tasa de aprendizaje [13] es otro componente clave en el entrenamiento de redes neuronales para obtener imágenes de superresolución. La tasa de aprendizaje determina la magnitud de los ajustes realizados en los pesos y sesgos del modelo durante el entrenamiento. Es esencial elegir un valor adecuado para la tasa de aprendizaje, ya que esto puede afectar significativamente la capacidad de aprendizaje del modelo y la estabilidad del proceso de entrenamiento.

En el contexto de la superresolución de imágenes, una tasa de aprendizaje demasiado alta puede provocar oscilaciones y divergencia en el proceso de entrenamiento, lo que dificulta la convergencia del modelo. Por otro lado, una tasa de aprendizaje demasiado baja puede producir el estancamiento del entrenamiento en mínimos locales.

3. Función de pérdidas:

La función de pérdida se utiliza para cuantificar la discrepancia entre las predicciones del modelo y los valores reales de los datos de entrenamiento.

En el caso de la superresolución de imágenes, el objetivo es aumentar la resolución de una imagen LR para obtener una versión de mayor resolución. La elección de la función de pérdida es fundamental para medir la diferencia entre la imagen de salida generada por el modelo y la imagen HR de referencia.

La función de pérdida de error cuadrático medio (MSE) [14] es una opción común para problemas de superresolución de imágenes. El MSE calcula la diferencia promedio al cuadrado entre los píxeles de la imagen generada y los píxeles de la imagen de referencia. Esta función penaliza los errores grandes de manera más significativa que los errores pequeños.

4. Épocas y batch size:

Las épocas se refieren a los ciclos completos en los que se pasa todo el conjunto de datos de entrenamiento a través del modelo, mientras que el *batch size* determina cuántas muestras se procesan simultáneamente antes de actualizar los pesos. Ambos parámetros son importantes para el proceso de entrenamiento de una red neuronal y pueden influir en el rendimiento y la capacidad de generalización del modelo.

El número de épocas influye en la capacidad de capturar patrones complejos y evitar el sobreajuste. El tamaño del *batch size* afecta la velocidad de entrenamiento y el uso de memoria, pero debe equilibrarse para aprovechar los patrones en los datos. Es crucial encontrar un equilibrio óptimo entre ambos para obtener un buen rendimiento y generalización en el entrenamiento de la red neuronal.

Sección 2

4. Datos

4.1. Satélites y sensores

Los conjuntos de datos que se utilizaron a la hora de entrenar los modelos que se proponen eran conjuntos de imágenes satelitales de dos satélites, en concreto que son el Landsat-8 y Sentinel-2. También se emplearon imágenes de alta resolución espacial que se obtienen mediante el vuelo de avionetas que montan diferentes tipos de sensores.

4.1.1. Landsat-8

El satélite Landsat-8 es una misión espacial desarrollada y operada por la Administración Nacional Oceánica y Atmosférica de los Estados Unidos (NOAA) y el Servicio Geológico de los Estados Unidos (USGS). Fue lanzado el 11 de febrero de 2013 y se encuentra en órbita polar alrededor de la Tierra.

Landsat-8 [15] es el octavo satélite de las misiones Landsat las cuales están pensadas para la observación y monitoreo de la Tierra a largo plazo. Cabe destacar que este satélite tiene una cobertura global de la superficie terrestre. Esto implica que el satélite es capaz de capturar imágenes de cualquier lugar del mundo, junto a su ciclo de repetición de un mismo punto en la corteza terrestre que es de 16 días aproximadamente, permitiendo su uso en una amplia gama de aplicaciones como puede ser el monitoreo de procesos medioambientales, por ejemplo las variaciones en la línea de costa.

En cuanto, a los sensores que monta a bordo el satélite destaca el *Operational Land Imager* (OLI) se trata de un sensor óptico que captura imágenes de la superficie terrestre en nueve bandas espectrales. Estas bandas abarcan desde el espectro visible hasta el infrarrojo térmico, lo que permite la detección y análisis de diversos fenómenos terrestres. Aparte del sensor OLI el satélite también monta otro sensor llamado *Thermal Infrared Sensor* (TIRS) el cual mide la radiación térmica emitida por la Tierra en dos bandas espectrales térmicas.

Para finalizar las características, la resolución espacial de Landsat-8 varía según la banda espectral. Las bandas pancromáticas (banda 8) tienen una resolución de 15 metros, mientras que las bandas multiespectrales tienen una resolución de 30 metros. Las bandas térmicas tienen una resolución de 100 metros.

A continuación, en la Tabla 2 se muestra un resumen de todas las bandas de los dos sensores comentados anteriormente junto a sus respectivas longitudes de onda y

resoluciones.

Landsat-8		
	Logitud de onda (μm)	Resolución (m)
Banda-1 - Coastal Aerosol	0.435 - 0.451	30
Banda-2 - Blue	0.452 - 0.512	30
Banda-3 - Green	0.533 - 0.590	30
Banda-4 - Red	0.636 - 0.673	30
Banda-5 - NIR	0.851 - 0.879	30
Banda-6 - SWIR 1	1.566 - 1.651	30
Banda-7 - SWIR 2	2.107 - 2.294	30
Banda-8 - Pancromatic	0.503 - 0.676	15
Banda-9 - Cirrus	1.363 - 1.384	30
Banda-10 - TIR 2	10.60 - 11.19	100
Banda-11 - TIR 1	11.50 - 12.51	100

Tabla 2: Bandas del satélite Landsat-8

4.1.2. Sentinel-2

El satélite Sentinel-2 [16] es parte del programa Copernicus de la Agencia Espacial Europea (ESA) y proporciona datos e imágenes de media resolución de la superficie terrestre. Fue lanzado el 23 de junio de 2015 y más tarde se lanzó un satélite con las mismas características nombrando dichos satélites como Sentinel-2A y Sentinel-2B dichos satélites se encuentran en una órbita polar y se complementan reduciendo la resolución temporal de la misión debido a que cada uno está aproximadamente en el punto opuesto al otro en la órbita polar en la que se encuentran.

El objetivo principal de la misión del satélite Sentinel-2 es proporcionar observaciones precisas y actualizadas sobre la superficie terrestre en todo el mundo. Para llevar a cabo este objetivo, el satélite está equipado con un instrumento óptico de imágenes multispectrales de media resolución, el cual captura 13 bandas espectrales diferentes, lo que permite una amplia cobertura espectral y una mayor capacidad para monitorear diversos fenómenos como los cambios en la línea de costa.

Respecto a la resolución espacial, el Sentinel-2 ofrece una resolución de 10 metros para las bandas espectrales visibles e infrarrojas cercanas. También presenta bandas de 20 metros de resolución espacial para las bandas espectrales de longitud de onda más larga, por último el sensor también presenta otras bandas de 60 metros. Además, el satélite tiene una resolución temporal de unos 5 días aproximadamente.

A continuación, en la Tabla 3 se observa un resumen de todas las bandas de disponibles para las tomas captadas con el satélite Sentinel-2 junto a sus respectivas longitudes de onda y resoluciones.

Landsat-8		
	Logitud de onda (μm)	Resolución (m)
Banda-1 - Aerosol	0.43 - 0.45	60
Banda-2 - Blue	0.45 - 0.52	10
Banda-3 - Green	0.54 - 0.57	10
Banda-4 - Red	0.65 - 0.68	10
Banda-5 - Red edge 1	0.69 - 0.71	20
Banda-6 - Red edge 2	0.73 - 0.74	20
Banda-7 - Red edge 3	0.77 - 0.79	20
Banda-8 - NIR 1	0.78 - 0.90	10
Banda-8A - NIR 2	0.85 - 0.87	20
Banda-9 - Water vapour	0.93 - 0.95	60
Banda-10 - Cirrus	1.36 - 1.39	60
Banda-11 - SWIR 1	1.56 - 1.65	20
Banda-12 - SWIR 2	2.10 - 2.28	20

Tabla 3: Bandas del satélite Sentinel-2

4.1.3. Ortoimágenes ICV

Para el caso de la obtención de las imágenes de alta resolución espacial, el Institut Cartogràfic Valencià (ICV) realiza vuelos fotométricos anualmente con el objetivo de obtener ortoimágenes que tienen una resolución espacial subpíxel. Estas imágenes de alta resolución se utilizaron para formar el segundo conjunto de datos.

En los vuelos que capturaron las imágenes empleadas para formar el *dataset* propio, se empleó el sensor *ULTRACAM EAGLE MARK 3*, el sensor es capaz de captar cuatro bandas espectrales, que incluyen las bandas del espectro visible (RGB) y el infrarrojo cercano (NIR). En cuanto a su velocidad de captura es de 1,5 segundos por imagen.

También, es importante mencionar que se pueden usar diferentes distancias focales en los vuelos. En el caso de los vuelos ejecutados por el ICV, se empleó una distancia focal de 80 mm, lo que permitió obtener una resolución espacial de 25 cm al volar a una altura aproximada de 2000 metros.

4.2. Métodos en el procesamiento de las imágenes

Para la creación del conjunto de datos propio se realizaron diversas técnicas de preprocesado, dichas técnicas son el remuestreo de las imágenes obtenidas en los vuelos de avioneta y la reproyección del sistema de referencia (CRS) en las imágenes obtenidas por el satélite Sentinel-2.

4.2.1. Reproyección

La reproyección es el proceso de cambiar la proyección o sistema de coordenadas de una imagen georreferenciada. Para llevar a cabo este método se pueden aplicar diferentes interpolaciones matemáticas para obtener las coordenadas correspondientes a la proyección de destino, entre estas las más utilizadas son la interpolación del vecino cercano o la interpolación bilineal.

La interpolación bilineal utiliza una combinación lineal de los valores de los píxeles vecinos para estimar el valor de un píxel en una ubicación deseada. Para ello, considera los cuatro píxeles más cercanos al punto de destino y realiza una interpolación lineal en ambas direcciones (horizontal y vertical).

Más concretamente, dado un punto de destino en la nueva imagen proyectada, se calcula su posición en la imagen de origen. Una vez se han identificado los cuatro píxeles más cercanos al punto de destino en la imagen de origen, se calcula la contribución ponderada de cada píxel vecino en la interpolación bilineal, teniendo en cuenta que los píxeles más cercanos tienen un mayor peso en el cálculo y se realiza una interpolación lineal tanto en la dirección horizontal como en la dirección vertical para estimar el valor del píxel en la nueva ubicación.

Se repiten los pasos anteriores para cada punto de destino en la nueva imagen, obteniendo así la imagen reproyectada completa.

4.2.2. Remuestreo

El remuestreo de imagen, es el proceso de modificar la resolución espacial de una imagen digital.

Cuando se realiza el remuestreo de una imagen, se toman los píxeles originales y se redistribuyen en una nueva cuadrícula para adaptarse al nuevo tamaño deseado. Dependiendo de si se está aumentando o disminuyendo el tamaño de la imagen, se pueden utilizar diferentes algoritmos de remuestreo. Entre estos algoritmos se destacan la interpolación bilineal y el método del vecino más cercano.

En el algoritmo del vecino más cercano [17] se eligen aquellos píxeles cuya distancia entre su centro y el punto localizado en la nueva imagen sea el menor, es decir, el valor del píxel de la nueva celda será el mismo que el del píxel más próximo respecto a la imagen original. Esto se puede apreciar representado en la Figura 6, donde se observa que el píxel seleccionado tiene una mayor influencia por los píxeles amarillos que por los valores rojos, por lo tanto, se selecciona el valor de los píxeles amarillos.

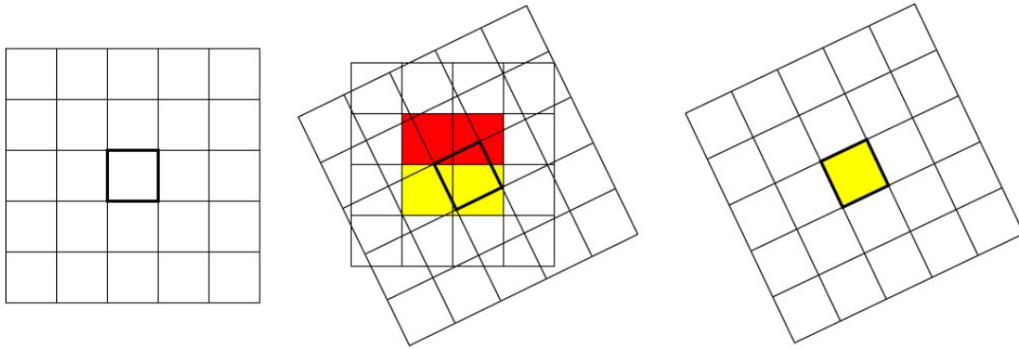


Figura 6: Remuestreo vecino más cercano. Fuente:[17]

Como ventajas de este algoritmo se destacan la no generación de valores intermedios, muy útil en rasters con valores asociados que no se desean modificar. Y desventajas de este método destaca la generación de movimientos de datos que causan contrastes entre valores, repetitividad en los valores y en ocasiones dicho algoritmo puede generar un aspecto escalonado.

En el método bilineal, el valor del píxel nuevo se obtiene a través de las distancias medias ponderadas de los valores de las cuatro celdas más próximas.

La Figura 7 se muestra un ejemplo de la aplicación del método bilineal donde se tiene en cuenta tanto la influencia de los píxeles rojos y amarillos, por lo tanto, se obtiene un valor intermedio entre estos valores que se representan en la Figura 7 con el color naranja.

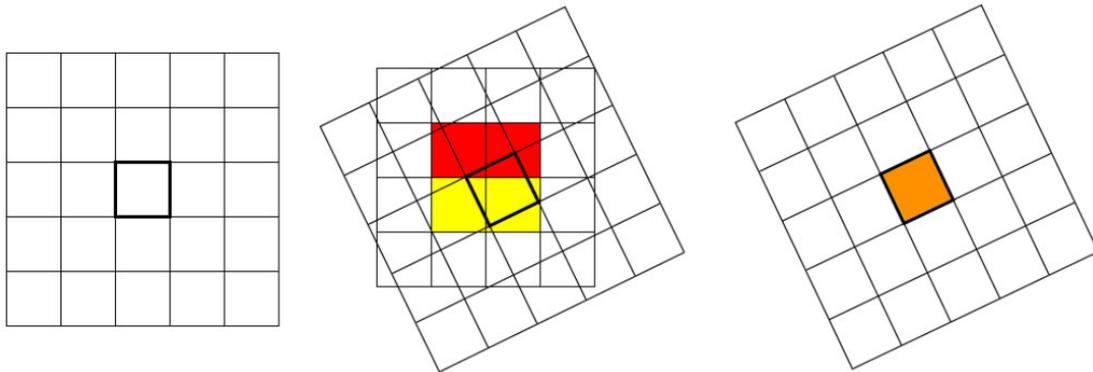


Figura 7: Remuestreo Interpolación bilineal. Fuente:[17]

Como ventajas se destacan la supresión en un porcentaje del efecto de valores escalonados típico en el algoritmo anterior y la generación de imágenes con menos contrastes visuales, es decir, se obtienen imágenes más claras y suavizadas. En cuanto

a las desventajas, se destaca la mayor complejidad a la hora de aplicar la interpolación, causando que la implementación de este método sea más costosa.

4.3. Bases de datos

El Trabajo Fin de Grado está compuesto por la implementación de dos bases de datos. Una es la bases de datos pública llamada OLI2MSI [8] formada por imágenes satelitales de Landsat-8 y Sentinel-2. Por otra parte, una base de datos propia compuesta por imágenes de satélite provenientes del Sentinel-2 y de diferentes vuelos fotométricos realizados con la finalidad de obtener ortoimágenes con gran resolución espacial, la cual se explicara como se generó y sus características en la sección 3 del documento.

4.3.1. OLI2MSI

El conjunto de datos OLI2MSI [8] se trata de un conjunto de datos que está compuesto por imágenes de los satélites Landsat-8 y Sentinel-2. Como se ha explicado anteriormente, los sensores multiespectrales de estos satélites tienen diferentes resoluciones espaciales para los diferentes canales o bandas. Por eso en este conjunto de datos se seleccionaron las bandas pertenecientes al espectro visible (RGB) y como se ha mostrado anteriormente en el caso del sensor OLI que monta Landsat-8 la resolución espacial es de 30 metros por píxel en dichas bandas, mientras que en el sensor de Sentinel-2 es de 10 metros por píxel.

Por lo tanto, se tomó como imágenes LR las capturadas por el sensor OLI que monta Landsat-8 y como imagen HR las captadas por el satélite Sentinel-2 con el objetivo de mejorar la resolución espacial de 30 metros por píxel a 10 metros por píxel.

Cabe destacar, que el conjunto de datos está formado por tomas que se tomaron en fechas coincidentes entre los dos satélites, evitando grandes variaciones en las tomas. En la Tabla 4 se muestra un pequeño resumen de algunas características relevantes de *dataset*.

Satélite	Landsat-8	Sentinel 2B
Bandas	RGB	RGB
Resolución espacial	30	10
Fecha	September 23, 2019	September 23, 2019

Tabla 4: Características del dataset OLI2MSI

Además, se seleccionaron las imágenes de nivel de procesamiento 1C para minimizar las diferencias en entre las condiciones atmosféricas y los cambios en el entorno entre los pares de imágenes.

Todo el conjunto de datos está parcheado a tamaños de 480x480 píxeles para las imágenes HR y 160x160 para las imágenes LR esta diferencia es debida al cambio de resolución espacial entre dichas imágenes que como ya se ha comentado anteriormente es de 30 metros por píxel para las imágenes LR y de 10 metros por píxel para las HR.

También, como se observa en las tablas 1 y 2 las bandas del RGB para el satélite Landsat-8 no tiene el mismo intervalo de longitud de onda para las diferentes bandas, aunque los intervalos son muy similares, por lo tanto, no tendrá mucha influencia a la hora de implementar el conjunto de datos.

Por otra parte, el conjunto de datos OLI2MSI está compuesto por 5325 imágenes, de las cuales se han destinado 4225 imágenes para la partición de entrenamiento, 1000 para la partición de validación y 100 para la partición de test, la selección de las particiones se puede apreciar en la Figura 8 del documento.



Figura 8: Partición del conjunto de datos OLI2MSI. Fuente: Elaboración propia

En cuanto a la región de la cual se tomaron las imágenes, como se puede observar en la Figura 9. Se trata de áreas coincidentes que están distribuidas en diferentes zonas, tales como lagos, bosques, zonas urbanas, entre otras. Además de las zonas de adquisición de las tomas, en la Figura 10 se puede observar algunos ejemplos de las imágenes que componen el conjunto de datos OLI2MSI donde en las imágenes superiores son las captadas con el satélite Landsat-8 y las de abajo las obtenidas con el Sentinel-2.

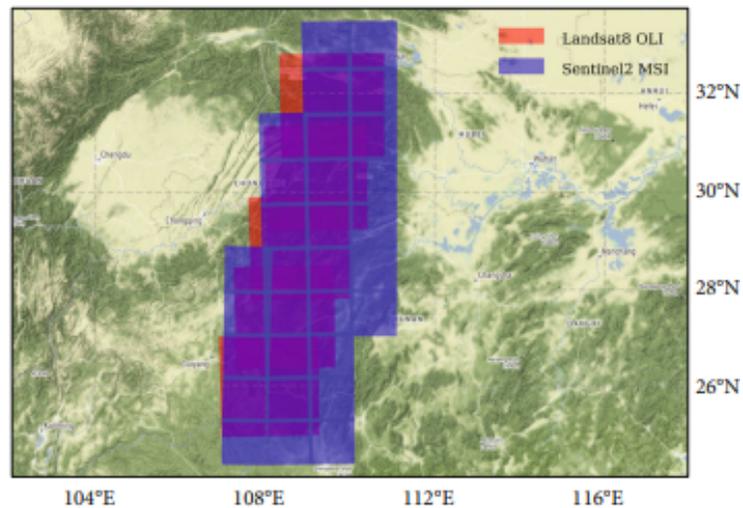


Figura 9: Huella del conjunto de datos OLI2MSI. Fuente:[8]

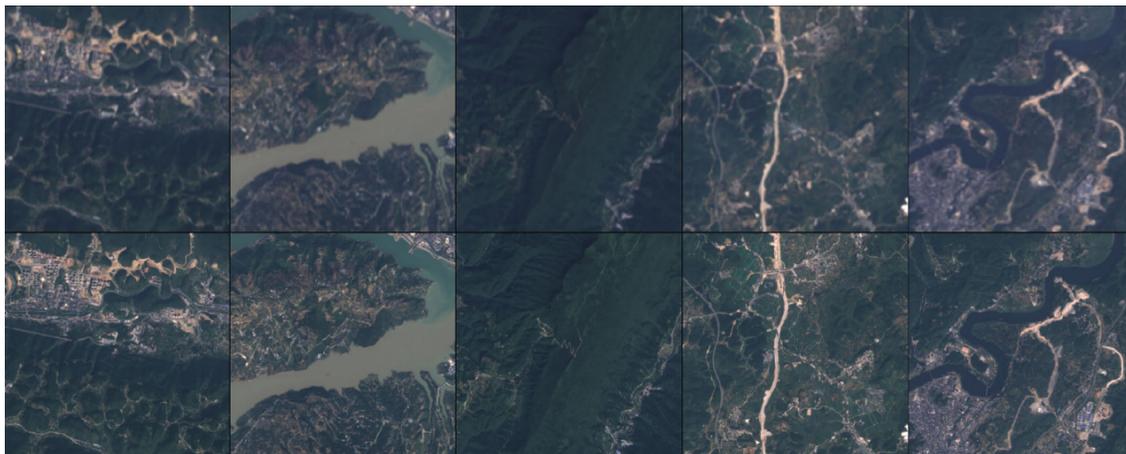


Figura 10: Ejemplo de algunos pares de imágenes del conjunto de datos OLI2MSI. Fuente:[8]

4.3.2. Base de datos propia

También se creó una base de datos propia donde se utilizaron imágenes satelitales captadas con el satélite Sentinel-2, así como imágenes de mayor resolución espacial adquiridas a través de vuelos planificados por el 'ICV' con el objetivo de obtener un conjunto de datos óptimo para mejorar la resolución de las imágenes Sentinel-2 en entornos costeros.

Para construir el conjunto de datos correctamente, se llevó a cabo un preprocesamiento que garantizara la correcta formación de pares de imágenes. En los conjuntos de datos destinados a la técnica de superresolución, es crucial que las imágenes em-

parejadas sean completamente coincidentes en términos de ubicación y fecha, a fin de evitar discrepancias que puedan afectar el entrenamiento del modelo.

El preprocesamiento constó de varios pasos. En primer lugar, se estableció un sistema de referencia consistente para todas las imágenes. Para lograr esto, se realizó una reproyección de las imágenes, utilizando una interpolación bilineal para preservar la calidad de los píxeles. Luego, se eliminaron las bandas no necesarias de las imágenes de Sentinel-2, conservando solo las bandas 'RGB'.

Posteriormente, se realizó un remuestreo de los píxeles en las imágenes capturadas durante los vuelos, para asegurar una relación de tamaño de píxel adecuada para el entrenamiento del modelo. El remuestreo se configuró en 2,5 metros por píxel, utilizando también una interpolación bilineal. Por último, se recortó la imagen de Sentinel-2 para que coincidiera con el tamaño de la imagen remuestreada. Con estos preprocesamientos, se logró preparar adecuadamente el conjunto de datos para el entrenamiento del modelo de superresolución.

A continuación, se muestran en la Figura 11 algunos ejemplos de como quedaron los pares en el conjunto de datos después teniendo en cuenta todas las características comentadas anteriormente.

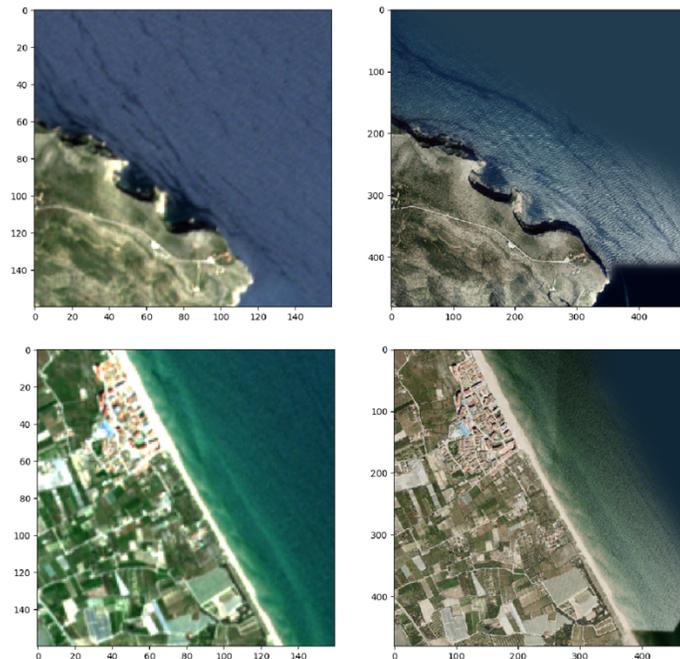


Figura 11: Ejemplo de algunos pares de imágenes del conjunto de datos propio. Fuente elaboración propia

Por último, el conjunto de datos propio está compuesto por 127 imágenes, de las cuales se han destinado 80 imágenes para la partición de entrenamiento, 20 para la partición de validación y 27 para la partición de test. Toda esta información la podemos observar esquematizada en la Figura 13.

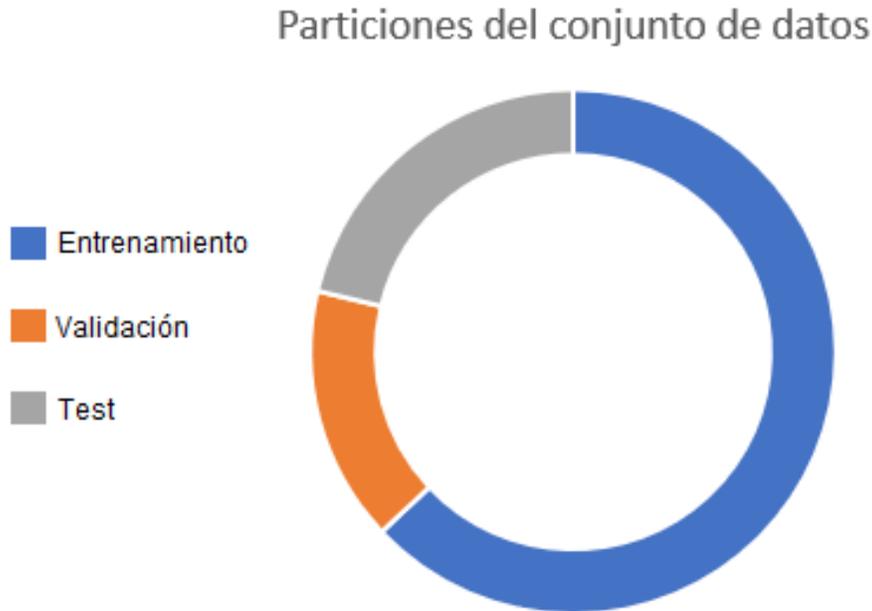


Figura 13: Partición del conjunto de datos propio. Fuente: Elaboración propia

4.3.3. Base de datos conjunta

También se optó por crear una base de datos conjunta combinando la base de datos OLI2MSI con el conjunto de datos creado. Este enfoque permitió tener una mayor variedad de imágenes satelitales para entrenar y evaluar los modelos de superresolución.

Los modelos utilizados son capaces de trabajar con diferentes factores de escala, lo que les permite procesar imágenes de distintas resoluciones espaciales. Esta capacidad de adaptación es importante para lograr una mejor superresolución de las imágenes.

El objetivo de fusionar los dos conjuntos de datos fue maximizar la información disponible y obtener imágenes superresueltas de mayor calidad. Al tener una base de datos conjunta, el modelo puede aprender de una variedad más amplia de imágenes satelitales y mejorar su capacidad para generar imágenes superresueltas, precisas y detalladas. Cabe destacar que las particiones de los conjuntos de datos se mantuvie-

ron, es decir, se cogieron las mismas particiones para cada conjunto de datos.

4.4. Entorno de desarrollo

Para llevar a cabo todas las tareas planteadas en este Trabajo de Fin de Grado, utilizaremos principalmente el entorno de desarrollo PyCharm de Python, junto con numerosas bibliotecas que ofrece Python para aplicaciones en el campo del aprendizaje profundo y la manipulación de conjuntos de datos.

4.4.1. Librerías más empleadas

Las librerías que más se han empleado para desarrollar el trabajo fin de grado son: Tiffle, Numpy y PyTorch.

1. Tiffle [18]: se trata de una librería destinada a la carga de todo tipo de imágenes *tiff* y almacenamiento de las imágenes en *arrays* de Numpy. Dicha librería fue creada por *Christoph Gohlke*.
2. Numpy [19]: Es una librería creada para el manejo de vectores y matrices multidimensionales, además de ofrecer una gran cantidad de funciones matemáticas para operar con ellas.
3. PyTorch [20]: Es la librería principal con la que se ha desarrollado el trabajo fin de grado. Es una biblioteca pública pensada para el desarrollo y entrenamiento de modelos en el campo del aprendizaje profundo.

Las características más importantes por la que se decidió utilizar esta librería es la amplia gama de herramientas y versatilidad que ofrece en comparación con otras librerías pensadas para el campo del aprendizaje profundo como Keras, que son más simples a la hora de realizar una implementación, pero limita mucho la cantidad de parámetros y variables que puedes modificar en la creación y entrenamiento de los modelos.

4.4.2. Otros software utilizados

Asimismo, aparte de Python se han empleado dos programas libres para el procesamiento y tratamiento de las imágenes satelitales, que son: SNAP y QGIS.

1. SNAP: *Sentinel Application Platform* (SNAP) es un software libre desarrollado por la Agencia Espacial Europea (ESA) con la finalidad de poder trabajar con imágenes de satélites y realizar múltiples procesos complejos a dichas imágenes. Entre dichos procesos de tratamiento de imágenes o procesos de teledetección se destacan la capacidad de manejo en las diferentes bandas multispectrales que presenta las imágenes satelitales y la posibilidad de realizar remuestreos entre otras aplicaciones.

2. QGIS: Se trata de un sistema de información geográfica (SIG) de código abierto, es un software con mucha potencialidad debido a la gran cantidad de opciones disponibles en todo tipo de dato o imagen geográfica.

Sección 3

5. Metodología

5.1. Diagrama de flujo

A continuación se muestra una representación visual del proceso seguido, mostrando la secuencia metodológica que se muestra en la Figura 14.

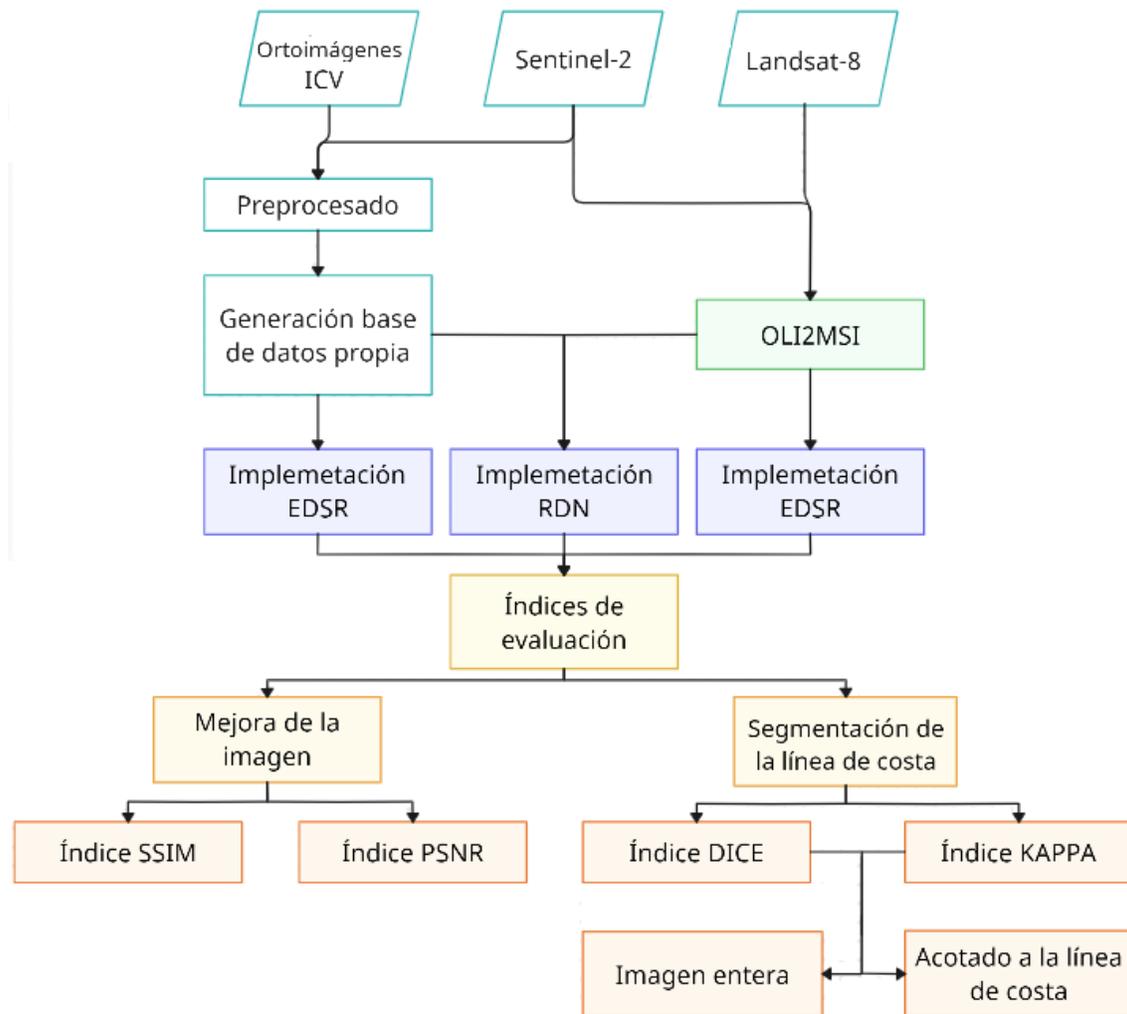


Figura 14: Diagrama de flujo metodológico. Fuente elaboración propia

5.2. Implementación de los modelos

Ahora se explicará con mayor detalle cómo fueron las implementaciones de cada modelo, qué parámetros de entrenamiento se seleccionaron y qué conjuntos de datos se emplearon.

5.2.1. EDSR aplicado en conjunto de datos OLI2MSI

Para abordar el primer objetivo específico correspondiente con obtener imágenes de superresolución utilizando la red neuronal EDSR en el conjunto de datos OLI2MSI, fue fundamental la selección de los parámetros correctos para alcanzar el objetivo de mejorar la resolución de las imágenes de entrada.

En primer lugar, fue crucial el factor de escala utilizado en el modelo EDSR y la relación entre el tamaño de los parches de entrada. En este caso, se estableció que los parches de las imágenes LR tienen dimensiones de 160x160 píxeles, mientras que los parches correspondientes de las imágenes de alta resolución HR tienen dimensiones de 480x480 píxeles. Dichos parches venían ya parcheados a estas dimensiones en el conjunto de datos OLI2MSI [8].

Por lo tanto, el factor de escala fue de 3, lo que significa que se esperaba obtener imágenes de alta resolución que sean tres veces más grandes que las imágenes de baja resolución con una resolución espacial máxima de 10 metros por píxel. Durante el entrenamiento del modelo, se tuvo en cuenta este factor de escala para capturar los detalles y las relaciones necesarias en la generación de superresolución.

En cuanto a la implementación del modelo EDSR, se utilizó una versión pre-entrenada obtenida de [21]. Esta elección se basó en la facilidad de carga del modelo y la aportación de buena documentación del mismo que permite el entrenamiento del modelo EDSR en el contexto de la superresolución de imágenes, aprovechando las funcionalidades y características específicas proporcionadas por el repositorio.

Capa	Entrada	Salida
Conv1	$3 \times H \times W$	$64 \times H \times W$
ResBlocks (num_blocks)	$64 \times H \times W$	$64 \times H \times W$
Conv2	$64 \times H \times W$	$256 \times H \times W$
Upsampling (factor=3)	$256 \times H \times W$	$256 \times 3H \times 3W$
Conv3	$256 \times 3H \times 3W$	$3 \times 3H \times 3W$

Tabla 5: Dimensiones de la red EDSR con factor de escala de *upsampling* de 3

La Tabla 5 muestra las dimensiones de entrada y salida de cada capa del modelo EDSR en esta implementación concreta para el conjunto de datos OLI2MSI.

Donde la capa inicial de convolución (Conv1) extrae características iniciales de la imagen de entrada, transformándola de dimensiones $3 \times H \times W$ a $64 \times H \times W$. A continuación, los bloques Residual (ResBlocks) capturan y fusionan características, manteniendo las dimensiones de entrada y salida en $64 \times H \times W$. Estos bloques permiten aprender representaciones más ricas y complejas.

Posteriormente, la capa de convolución (Conv2) incrementa la profundidad de las características a $256 \times H \times W$. Seguidamente, la capa de *upsampling* aumenta la resolución en un factor de 3, transformando las dimensiones a $256 \times H \times W$ a $256 \times 3H \times 3W$.

Finalmente, la última capa de convolución (Conv3) reduce la profundidad a $3 \times 3H \times 3W$ para obtener la imagen SR de salida.

En cuanto a los parámetros de entrenamiento, se seleccionaron los parámetros de entrenamiento siguientes: se entrenó el modelo durante 50 épocas con un *batch size* de uno, lo que permitió actualizar los pesos después de cada par de imágenes (LR y HR). Se implementó el optimizador *ADAM* con una tasa de aprendizaje de 0.0001 y se utilizó la función de pérdida MSE.

5.2.2. EDSR aplicado en *dataset* propio

Para abordar el segundo objetivo específico correspondiente con obtener imágenes de superresolución utilizando la red neuronal EDSR junto con el conjunto de datos propio.

El factor de escala y la relación entre los tamaños de los parches de entrada fueron aspectos cruciales en la implementación del modelo EDSR. En este caso, se decidió que los parches de las imágenes LR tengan dimensiones de 160x160 píxeles, mientras que los parches correspondientes de las imágenes HR tuvieran dimensiones de 640x640 píxeles

Por lo tanto, el factor de escala era de 4, lo que significa que se esperaba obtener imágenes SR que como máximo podrían alcanzar una resolución espacial de 2,5 metros por píxel.

En relación a la implementación del modelo EDSR, se tomó la decisión de utilizar la versión pre-entrenada mencionada previamente debido a las mismas consideraciones mencionadas anteriormente.

La Tabla 6 exhibe la misma estructura de la red EDSR presentada en el apartado previo, pero en esta ocasión se considera un factor de escala de 4 debido al cambio en el conjunto de datos utilizado.

Capa	Entrada	Salida
Conv1	$3 \times H \times W$ (kernel=3, stride=1)	$64 \times H \times W$ (kernel=3, stride=1)
ResBlocks (num_blocks)	$64 \times H \times W$	$64 \times H \times W$
Conv2	$64 \times H \times W$	$256 \times H \times W$
Upsampling (factor=4)	$256 \times H \times W$	$256 \times 4H \times 4W$
Conv3	$256 \times 4H \times 4W$	$3 \times 4H \times 4W$

Tabla 6: Dimensiones de la red EDSR con factor de escala de *upsampling* de 4

En cuanto a los parámetros de entrenamiento, se seleccionaron los parámetros de entrenamiento similares a los de la implementación anterior, con la única diferencia de que se disminuyeron el número de épocas a 30 debido a que los resultados eran más óptimos.

5.2.3. RDN aplicado al *dataset* conjunto

En esta nueva implementación, el objetivo siguió siendo mejorar la resolución espacial de las imágenes satelitales en zonas costeras. Sin embargo, se utilizó un conjunto de datos y un modelo diferentes a las implementaciones anteriores.

En este caso, se fusionaron los conjunto de datos OLI2MSI con el propio, creando un nuevo conjunto de datos compuesto por pares de imágenes LR y HR. Las imágenes LR podían provenir del satélite Landsat-8 o del satélite Sentinel-2, con una resolución espacial de 30 metros por píxel y 10 metros por píxel, respectivamente. Por otro lado, las imágenes HR podían ser imágenes del satélite Sentinel-2 o imágenes obtenidas a través de vuelos fotométricos para generar ortofotos con una resolución espacial cercana o incluso inferior a un metro por píxel.

En este caso, se decidió utilizar un factor de escala de 3 debido a que todas las imágenes del conjunto de datos OLI2MSI ya estaban parcheadas considerando este factor de escala. En cuanto al conjunto de datos propio, se realizó un remuestreo bilineal de las imágenes HR para que todos los parches que tuvieran el mismo tamaño.

Una vez que todos los parches estaban en el mismo tamaño, se procedió a entrenar el modelo teniendo en cuenta las capas del modelo RDN, las cuales se detallan en la Tabla 7.

Capa	Dimensión de entrada	Dimensión de salida
Capa de entrada	$3 \times H \times W$	$3 \times H \times W$
Convolución	$3 \times H \times W$	$64 \times H \times W$
Bloque RDB 1	$64 \times H \times W$	$64 \times H \times W$
Bloque RDB 2	$64 \times H \times W$	$64 \times H \times W$
\vdots	\vdots	\vdots
Bloque RDB 15	$64 \times H \times W$	$64 \times H \times W$
Convolución	$64 \times H \times W$	$64 \times H \times W$
Upsampling	$64 \times H \times W$	$3 \times 3H \times 3W$
Convolución	$3 \times 3H \times 3W$	$3 \times 3H \times 3W$
Capa de salida	$3 \times 3H \times 3W$	$3 \times 3H \times 3W$

Tabla 7: Dimensiones de las capas del modelo RDN con factor de escala 3.

La Tabla 7 muestra la estructura del modelo RDN utilizado para la superresolución de imágenes. La capa de entrada recibe la imagen LR representada como un tensor de dimensiones $3 \times H \times W$. Los tres canales de color RGB están representados por la dimensión 3, mientras que H y W representan la altura y el ancho de la imagen, respectivamente.

A continuación, se aplica una capa de convolución que tiene como objetivo procesar la imagen y extraer características relevantes. Esta capa convolucional transforma el tensor de entrada de $3 \times H \times W$ a un tensor de $64 \times H \times W$, aumentando la profundidad de los canales.

El modelo continúa con una serie de bloques Residual *Dense Blocks*, que son responsables de aprender representaciones más profundas y complejas de la imagen. Estos bloques residuales mantienen la misma dimensionalidad de $64 \times H \times W$ en su entrada y salida, lo que permite la propagación de los gradientes y evita la degradación de la información durante el entrenamiento.

Después de los bloques RDB, se aplica otra capa de convolución que mantiene la dimensionalidad de $64 \times H \times W$. Esta capa ayuda a consolidar las características aprendidas hasta este punto en el modelo.

A continuación, se aplica una capa de *upsampling* (aumento de muestreo) que aumenta la resolución de la imagen. Esta capa convierte el tensor de $64 \times H \times W$ a un tensor de $3 \times 3H \times 3W$, logrando así una imagen de mayor tamaño y resolución.

Finalmente, se aplica una última capa de convolución que mantiene la dimensionalidad de $3 \times 3H \times 3W$ en la salida del modelo. Esta capa realiza los ajustes finales en la imagen generada antes de ser entregada como resultado final de la superresolución.

En cuanto a los parámetros de entrenamiento, se utilizaron parámetros similares a la implementación anterior, con la única diferencia de que se aumentó el número de épocas a 75. Esto se debió a que el modelo mostró una buena progresión de aprendizaje en la época 50, por lo que se decidió aumentar el número de épocas para obtener mejores resultados.

5.3. Evaluación de la calidad de la imagen sintética de alta resolución generada

La evaluación de la calidad de las imágenes generadas es fundamental en la tarea de superresolución de imágenes. Dos índices ampliamente utilizados para ello son el Índice de Similitud Estructural (SSIM) y la Relación Señal-Ruido Pico (PSNR) los cuales se utilizaron para evaluar los resultados como se muestra en la siguiente sección.

5.3.1. SSIM

El SSIM [22] es una métrica que evalúa la similitud estructural entre una imagen generada y una imagen de referencia de alta resolución. Este índice tiene en cuenta aspectos como la luminancia, el contraste y la estructura. La luminancia se refiere a la información de brillo en la imagen, mientras que el contraste mide las diferencias entre los niveles de gris. La estructura se relaciona con la distribución espacial de los píxeles y su relación con los vecinos.

El SSIM se calcula comparando las características estructurales de la imagen generada y la imagen de referencia. Cuanto más cercano a 1 sea el valor del SSIM, mayor será la similitud estructural entre ambas imágenes y mejor será la calidad de la imagen generada. Un valor de SSIM cercano a 0 indica una baja similitud.

El SSIM se calcula utilizando la siguiente fórmula:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (5.1)$$

Donde:

- $SSIM$ representa el índice de similitud estructural entre las imágenes x e y .
- $\mu_x\mu_y$ son las medias de las imágenes x e y , respectivamente
- $\sigma_x^2\sigma_y^2$ son las varianzas de las imágenes x e y , respectivamente.
- σ_{xy} es la covarianza entre las imágenes x e y .
- c_1c_2 son constantes utilizadas para estabilizar la división cuando el denominador es cercano a cero.

5.3.2. PSNR

El PSNR [23] es un índice que evalúa la relación señal-ruido entre la imagen generada y la imagen de referencia. El PSNR se basa en la diferencia entre los valores de los píxeles y mide el nivel de ruido introducido durante el proceso de superresolución. Un valor alto de PSNR indica una menor cantidad de ruido y, por lo tanto, una mejor calidad de imagen.

El PSNR se calcula utilizando la diferencia cuadrada media entre los valores de píxeles de las dos imágenes y se expresa en decibelios (dB). Cuanto más alto sea el valor de PSNR, menor será la diferencia entre las imágenes y mejor será la calidad de la imagen generada. Teniendo en cuenta que para las imágenes tiene un valor de píxeles de 256, por lo tanto, el valor máximo del PSNR es de 48.13

PSNR se calcula utilizando la siguiente fórmula:

$$PSNR(x, y) = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}(x, y)} \right) \quad (5.2)$$

Donde:

- $PSNR$ representa la relación señal a ruido en la imagen entre x e y .
- MAX_I^2 es el valor máximo que puede tener un píxel en la imagen.
- MSE se calcula utilizando la siguiente fórmula:

$$\text{MSE} = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (5.3)$$

Donde:

- MSE representa el error cuadrático medio.
- m es el número de filas de las imágenes.
- n es el número de columnas de las imágenes.
- $I(i, j)$ es el valor del píxel en la imagen original en la posición (i, j) .
- $K(i, j)$ es el valor del píxel en la imagen reconstruida o procesada en la posición (i, j) .

5.4. Evaluación mediante la segmentación de la línea costera

Además del análisis de la mejora de la resolución espacial en las imágenes, se llevó a cabo un estudio adicional para cuantificar la efectividad de las técnicas de segmentación de la línea de costa utilizando otra red neuronal desarrollada por mi compañero Luis Correas Naranjo en su Trabajo de Fin de Grado.

En primer lugar, se calculó la máscara binaria correspondiente a cada imagen LR, HR y SR empleando la implementación del modelo de segmentación desarrollado por Luis. Esta red neuronal fue entrenada para detectar y segmentar de manera precisa la línea de costa en las imágenes. En la Figura 15 se observa un ejemplo del resultado de la segmentación al aplicar el modelo de segmentación desarrollado por Luis.

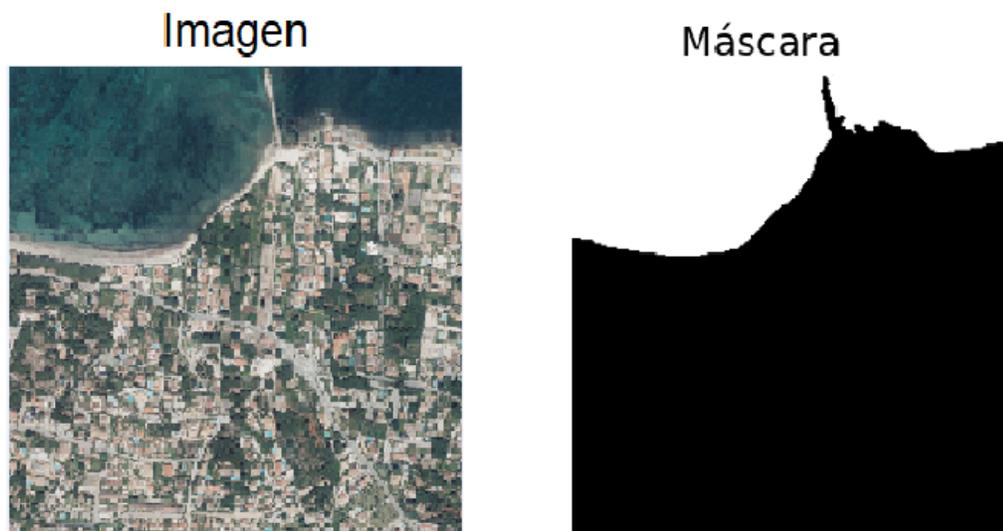


Figura 15: Ejemplo de máscara binaria obtenida mediante el modelo de segmentación. Fuente elaboración propia

En segundo lugar, se calcularon los índices Dice [24] y Kappa [25] para todas las máscaras binarias en el conjunto de prueba del *dataset* generado. Como se ha mencionado anteriormente, estos índices son medidas de evaluación ampliamente utilizadas en el campo de la segmentación de imágenes.

Por último, se obtuvo el promedio de estos índices en el conjunto de datos de prueba para tener una medida general de la efectividad de la segmentación de la línea de costa empleando la red neuronal desarrollada por Luis. Los resultados obtenidos se mostrarán en la sección siguiente del documento.

5.4.1. Matriz de confusión

La matriz de confusión es una matriz que muestra una comparativa entre las predicciones y los valores reales de un conjunto de datos, con el fin de evaluar el rendimiento de todo tipo de métodos o modelos de clasificación.

Cada columna de la matriz de confusión representa la información de una clase real, mientras que cada fila representa las predicciones obtenidas por la clasificación para una clase en particular, dichas características se representan claramente en la Figura 16.

		Valores reales	
Valores de predicción		Verdaderos positivos (TP)	Falsos Positivos (FP)
		Falsos Negativos (FN)	Verdaderos Negativos (TN)

Figura 16: Matriz de confusión. Fuente: Elaboración propia

- Los verdaderos positivos (TP) representan las instancias clasificadas correctamente como positivas.
- Los falsos positivos (FP) representan las instancias clasificadas incorrectamente como positivas.
- Los verdaderos negativos (TN) representan las instancias clasificadas correctamente como negativas.
- Los falsos negativos (FN) representan las instancias clasificadas incorrectamente como negativas.

La implementación de la matriz de confusión es muy utilizada para evaluar métodos de clasificación debido a que se pueden calcular multitud de métricas a partir de la misma. En este Trabajo Fin de Grado se ha obtenido el índice DICE.

5.4.2. DICE

El índice DICE es una métrica que representa la similitud entre dos máscaras binarias. Sus valores varían entre 0 y 1 donde un valor de 1 indica una perfecta superposición entre las dos máscaras, mientras que un valor 0 indica una discrepancia total en la superposición entre las máscaras.

Se obtiene con la fórmula:

$$DICE = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (5.4)$$

5.4.3. KAPPA

El coeficiente de Kappa se utiliza para medir la concordancia entre dos máscaras binarias y toma valores entre -1 y 1. Un valor 1 indica una concordancia perfecta y -1 indica una nula concordancia ente las muestras.

Se obtiene con la fórmula:

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (5.5)$$

Donde:

- κ representa el coeficiente de Kappa.
- p_o es la proporción observada de acuerdo entre las dos máscaras binarias.
- p_e es la proporción esperada de acuerdo entre las dos máscaras binarias por casualidad.

5.4.4. Cálculo de los índices acotados al área de la línea de costa

Debido a que la máscara generada en el proceso de segmentación de la línea de costa se aplicaba a nivel global de la imagen, al realizar los cálculos de los índices de evaluación, las métricas resultantes eran muy elevadas. Esto se debe a que los cambios solo se concentraban en el área costera. Por esta razón, se tomó la decisión de restringir el cálculo de estos índices únicamente a la zona de cambio, con el fin de obtener métricas más precisas y representativas de los cambios específicos en esa área.

Para acotar el cálculo de los índices en la zona de cambio, se siguieron los siguientes pasos:

En primer lugar, se cargó la máscara binaria correspondiente a cada imagen dentro del conjunto de prueba. A continuación, se obtuvo el *boundary* de la máscara de referencia, es decir, los bordes que delimitan la zona donde se produce el cambio de la imagen HR fijando así un área de cambio común entre las diferentes máscaras

(LR, HR y SR). Se tomó la decisión de fijar la imagen HR debido a que los cambios se producen por las diferencias al aplicar el modelo de segmentación en los diferentes tipos de imagen (LR, HR y SR) y por ello se consideró que la segmentación más correcta sería la de imagen con mayor resolución (HR).

Posteriormente, se generó una máscara aplicando un *buffer* de 5 píxeles a cada lado del *boundary*. Como se muestra en la figura inferior. Esto se realizó con el fin de acotar el cálculo de las métricas al área de cambio en las máscaras. Una vez establecido el *buffer*, se procedió a multiplicar la máscara producida con el área de cambio por las máscaras originales, obteniendo la máscara que muestra en la Figura 17.

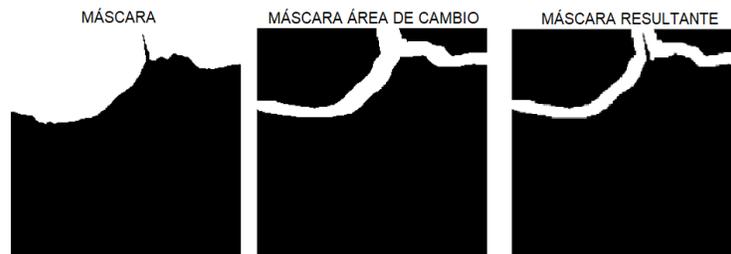


Figura 17: Proceso para obtener el área acotada en la zona de cambio. Fuente: Elaboración propia

Finalmente, se calcularon los índices comentados anteriormente en la zona acotada por el *buffer*. Para que los resultados reflejaran de manera precisa la similitud o diferencia entre las imágenes en la región de interés, evitando distorsiones generadas por otros factores fuera de esta zona. Dichos factores se observan en la siguiente sección del documento.

Sección 4

6. Resultados

6.1. Resultados OLI2MSI

Para evaluar los resultados de las imágenes de superresolución obtenidas mediante nuestro modelo, utilizaremos dos métricas ampliamente utilizadas en el campo de la generación de imágenes de alta resolución: el índice SSIM y el índice PSNR.

Se calcularon estos índices para la partición de prueba del conjunto de datos OLI2MSI, el cual consiste en imágenes con características similares a las de las imágenes utilizadas para el entrenamiento.

En primer lugar, se calcularon implementando el modelo EDSR entrenado únicamente con el conjunto de datos OLI2MSI. Donde los resultados que se obtuvieron fueron los siguientes:

	SSIM_OLI2MSI	PSNR_OLI2MSI
Media	0,9871	36,6244

Tabla 8: Valores medios de SSIM y PSNR con el modelo EDSR.

En segundo lugar, Utilizaremos el modelo RDN entrenado con la base de datos fusionada, ya que este modelo permite procesar tanto imágenes de Landsat-8 como imágenes de Sentinel-2 como entradas. A continuación, presentamos los resultados obtenidos:

	SSIM_OLI2MSI	PSNR_OLI2MSI
Media	0,9821	36,5840

Tabla 9: Valores medios de SSIM y PSNR con el modelo RDN.

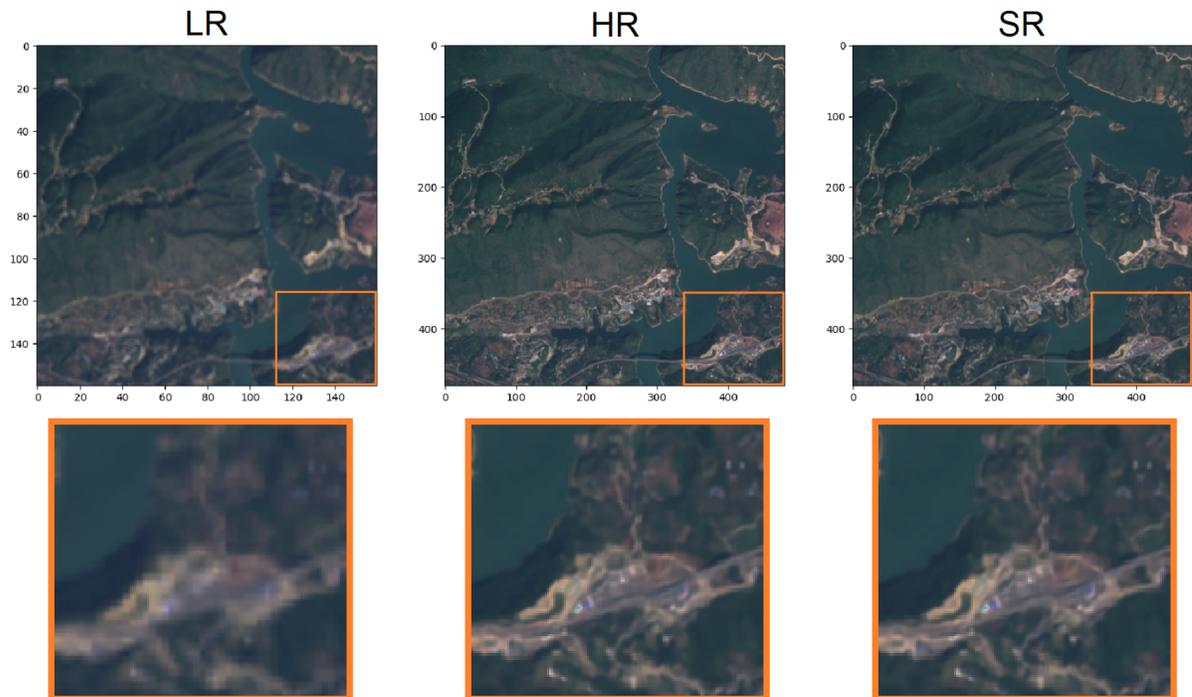
Los resultados obtenidos para ambas arquitecturas son altamente satisfactorios, las imágenes superresueltas muestran una similitud estructural muy cercana a la original y un nivel de ruido extremadamente bajo en ambos casos. Estos resultados demuestran la eficacia de los modelos en la tarea de superresolución para el conjunto de datos OLI2MSI.

Además, se aportan representaciones visuales de los resultados obtenidos aplicando la arquitectura RDN mediante las tres imágenes disponibles. Estas imágenes son: la imagen LR, la cual se obtiene a través del sensor OLI del satélite Landsat-8, como se mencionó anteriormente. Esta imagen LR es fundamental para comparar y evaluar el rendimiento de los otros dos tipos de imágenes.

Por otro lado, se encuentra la imagen HR, capturada por el satélite Sentinel-2. Esta imagen HR proporciona una mayor cantidad de detalles y resolución espacial en comparación con la imagen LR. Su utilización permite analizar de manera más precisa y detallada las características y elementos presentes en la escena capturada.

Finalmente, se muestra la imagen SR, la cual es el resultado obtenido a través del modelo entrenado con el conjunto de datos OLI2MSI. Esta imagen SR representa la imagen generada por el modelo a partir de las imágenes de entrada del conjunto de test presente en la base de datos OLI2MSI.

Estos tres tipos de imágenes (LR, HR y SR) se presentan gráficamente en las Figura 18, para facilitar la comparación visual y la evaluación de la mejora lograda mediante el modelo de generación de imágenes de alta resolución.



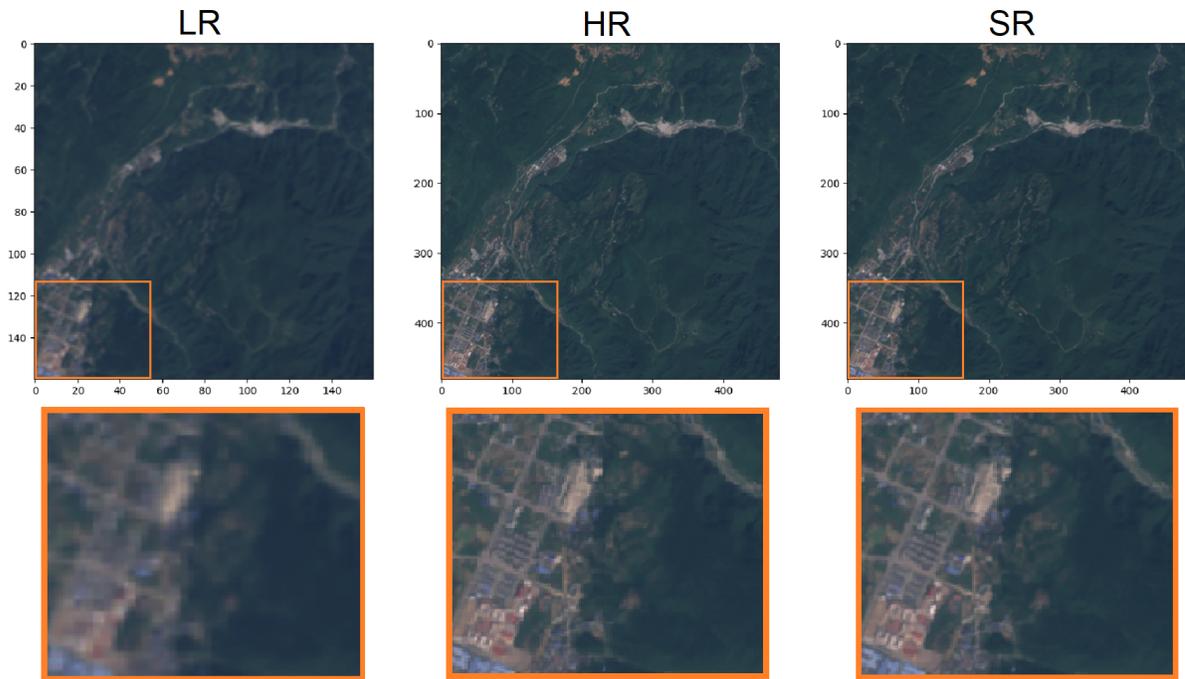


Figura 18: Resultado del conjunto de datos OLI2MSI. Fuente: Elaboración propia

Asimismo, se llevó a cabo una prueba utilizando una imagen descargada directamente del portal *EOBrowser* [26], un portal destinado a la visualización y descarga de todo tipo de imágenes satelitales. En este caso, se descargó una imagen de la zona costera de la Albufera en Valencia, capturada por el satélite Landsat-8. Con el objetivo de demostrar el funcionamiento del modelo en un contexto específico de zona costera y evaluar su rendimiento en dicho escenario.

En este caso, no podemos comparar la imagen con una imagen de Sentinel-2, ya que se trata de una imagen descargada. Sería bastante tedioso buscar una imagen que coincida en fecha y zona entre los dos satélites, aparte de otros factores como la no presencia de cuerpos nubosos en las tomas.

Sin embargo, podemos realizar una comparación visual entre las imágenes mostradas en la Figura 19. En esta comparación, es evidente que la imagen generada por el modelo tiene una mayor resolución espacial en comparación con la imagen de Landsat-8.

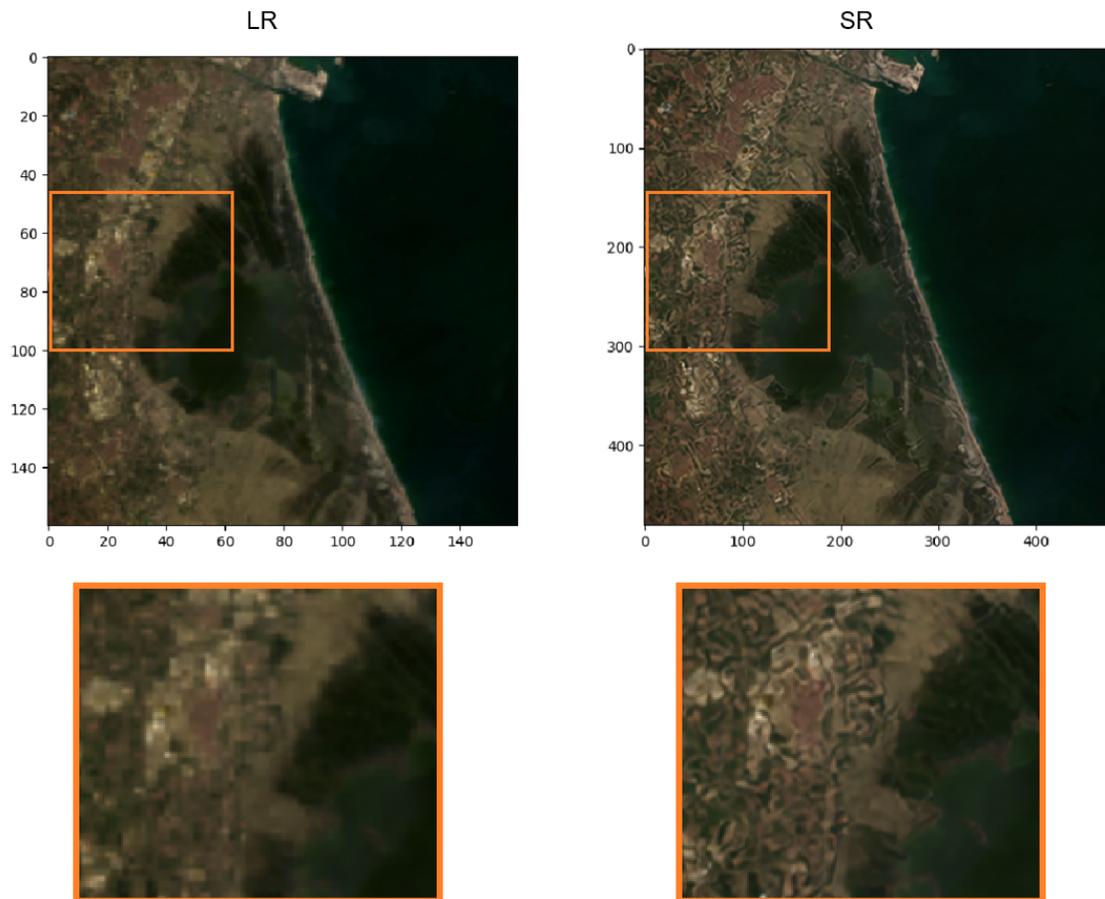


Figura 19: Prueba del modelo en una imagen de la Albufera. Fuente: Elaboración propia

6.2. Resultados Conjunto de datos propio

En cuanto al *dataset* generado, como se mencionó anteriormente en el apartado 3, también se destinó un porcentaje del conjunto de imágenes descargadas para llevar a cabo una validación posterior del modelo utilizando esta partición de prueba.

6.2.1. Análisis de imágenes superresueltas

También, se han calculado los índices PSNR y SSIM entre las imágenes del conjunto de datos propio utilizando el modelo RDN entrenado con la base de datos fusionada. Estas métricas (PSNR, SSIM) permiten evaluar el rendimiento del modelo RDN y su capacidad para generar imágenes SR basadas en las imágenes de entrada del satélite Sentinel-2.

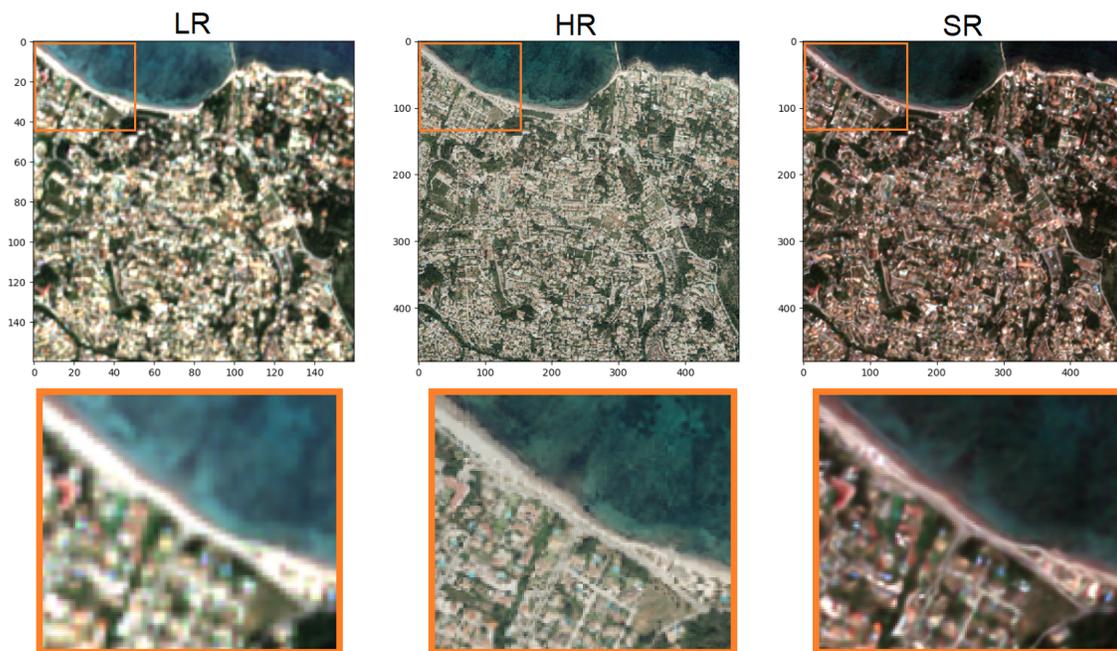
	SSIM	PSNR
Media	0,7514	31,1223

Tabla 10: Valores medios de SSIM y PSNR dataset propio

En este caso, los resultados obtenidos que se muestran en la Tabla 10, no alcanzan los niveles máximos obtenidos anteriormente, pero siguen indicando una mejora en la calidad de imagen al utilizar el modelo de superresolución en imágenes Sentinel-2 y resolución subpíxel. Estos resultados respaldan la eficacia del modelo en la tarea de mejorar la calidad de imágenes con limitaciones de resolución, lo cual es relevante para diversas aplicaciones en el ámbito de la teledetección y el análisis de imágenes.

Además de calcular las métricas mencionadas anteriormente, se proporcionan representaciones visuales de los resultados obtenidos. Estas representaciones visuales siguen la misma estructura que las figuras generadas para el *dataset* anterior. Sin embargo, hay una diferencia clave: en este caso, la imagen LR es obtenida con el satélite Sentinel-2, mientras que la imagen HR es obtenida a partir de los vuelos del ICV y remuestreadas a una resolución de 3,33 metros por píxel.

Estas representaciones visuales (Figura 20) permiten visualizar y comparar la calidad de las imágenes SR generadas por el modelo RDN en relación con las imágenes HR obtenidas a partir de los vuelos del ICV. Así, se puede evaluar visualmente la capacidad del modelo para mejorar la resolución y los detalles de las imágenes obtenidas del satélite Sentinel-2.



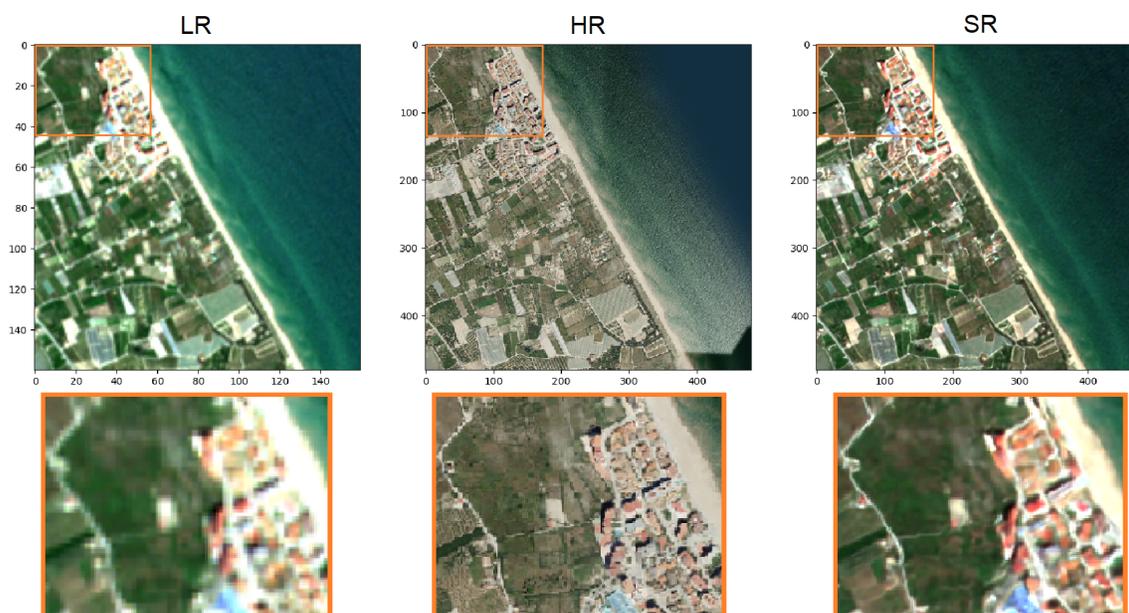


Figura 20: Resultados obtenidos con el conjunto de datos propio. Fuente: Elaboración propia

Observando las imágenes, se puede destacar una gran variación radiométrica entre las imágenes obtenidas. Esto se debe a que las imágenes de entrada LR y HR no tienen el mismo preprocesamiento y son capturadas por diferentes sensores, lo que produce estas variaciones.

Cabe destacar que este fenómeno afecta a los índices calculados anteriormente, disminuyendo su coherencia, lo cual indica que la mejora en la resolución espacial de las imágenes SR es evidente, pero no se refleja tan claramente en los índices de evaluación como en el caso del conjunto de datos OLI2MSI, debido a estas variaciones radiométricas.

Por otra parte, en el anexo II se muestran unos mapas comparativos donde se muestran algunos ejemplos donde las imágenes HR y LR se muestran de fondo y superpuesta se muestran imágenes SR correspondientes. En dichos mapas se puede observar con mayor calidad los cambios comentados.

6.2.2. Análisis de línea de costa

Como se ha comentado en la Sección anterior, también se realizó un análisis de resultados para cuantificar las variaciones en la precisión a la hora de detectar la línea de costa según la imagen que se segmentara, obteniendo de media el resultado que se muestran la Tabla 11.

Comparación	Índice Dice	Índice Kappa
HR vs SR	0,9778	0,9537
LR vs SR	0,9649	0,9403
LR vs HR	0,9596	0,9331

Tabla 11: Comparación de los índices Dice y Kappa para las diferentes combinaciones de imágenes.

Como se puede observar en la tabla anterior, las máscaras binarias generadas por el modelo RDN muestran una mayor similitud con las imágenes HR en comparación con las imágenes LR. Esto confirma que nuestro modelo funciona correctamente. Sin embargo, al calcular los índices para toda la máscara, los valores de las métricas resultan muy elevados, lo que puede distorsionar ligeramente los resultados. Por esta razón, se decidió calcular los índices únicamente en la zona acotada a la línea de costa, donde se producen los cambios.

Se adjuntan a continuación imágenes representadas en la Figura 21, que contienen las máscaras binarias generadas a partir de cada tipo de imagen para permitir un análisis visual de la segmentación realizada por el modelo en los tres casos.

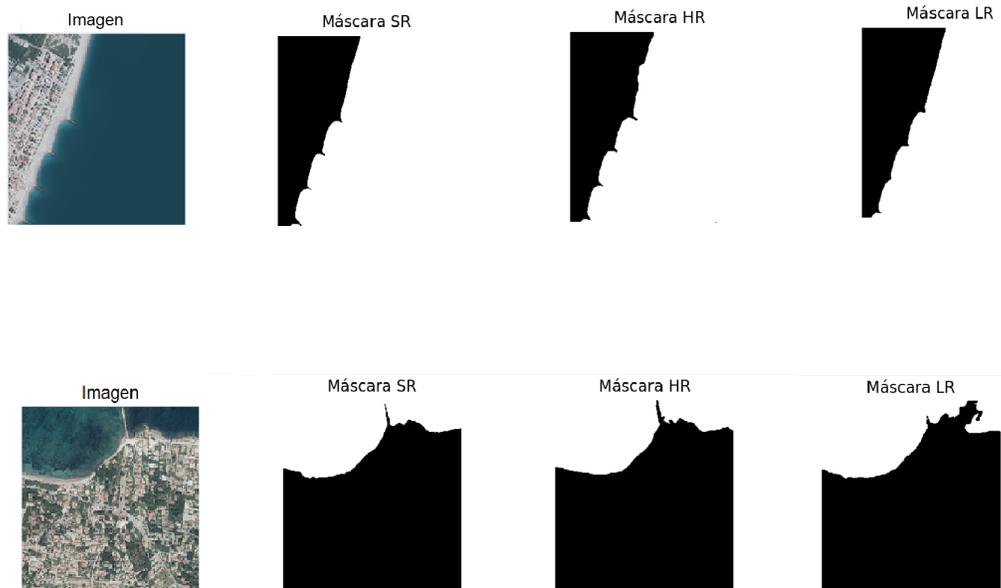


Figura 21: Mascaras de segmentación de la línea de costa. Fuente: Elaboración propia

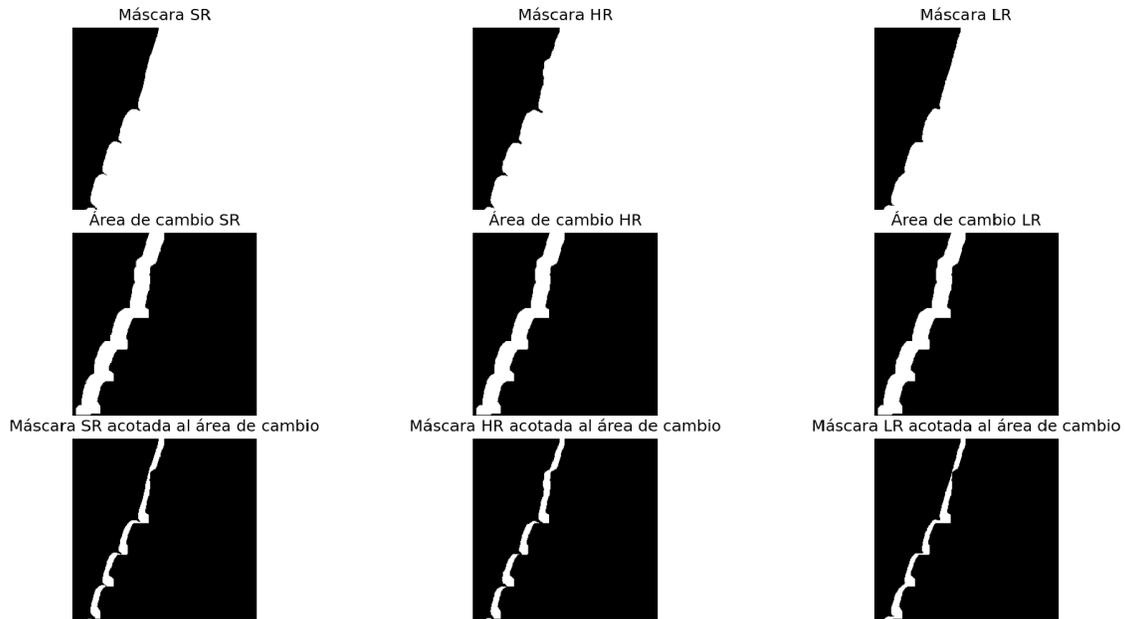
6.2.3. Análisis en la zona acotada a la línea de costa

Los resultados que se consiguieron de media para la partición de test en el conjunto de datos propio, aplicando la metodología comentada en la sección anterior, donde se explica como se realizó el cálculo de los índices únicamente en la zona acotada en la línea de costa anterior son los siguientes:

Comparación	Índice Dice	Índice Kappa
HR vs SR	0,8890	0,8846
LR vs SR	0,8803	0,8753
LR vs HR	0,8542	0,8480

Tabla 12: Comparación de los índices Dice y Kappa para las diferentes combinaciones de imágenes en la zona acotada.

A continuación, en la Figura 22 se observa el resultado de dos máscaras con la selección de la zona acotada, teniendo en cuenta que las métricas de evaluación solo se han calculado en la zona acotada por el *buffer*.



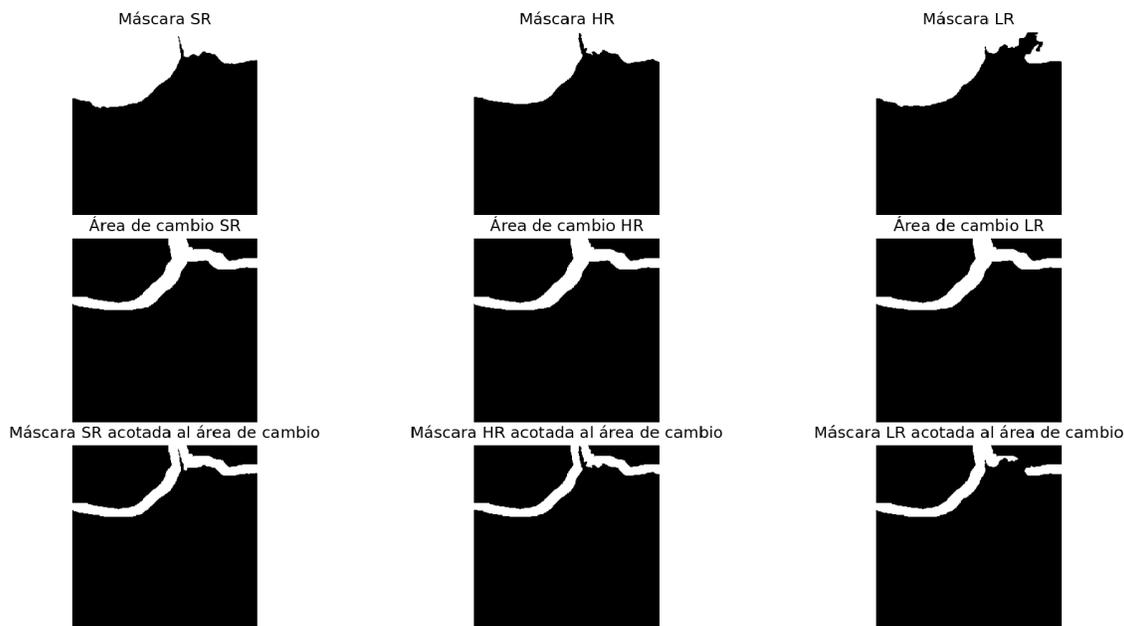


Figura 22: Segmentación de la línea de costa en la zona acotada. Fuente: Elaboración propia

Analizando los resultados de la tabla anterior, se puede afirmar que las imágenes HR y SR son las más similares, seguidas por las imágenes LR y SR. Por otro lado, las imágenes LR y HR son las menos parecidas. Esto indica que el modelo RDN utilizado tiene una mayor capacidad de mejorar la resolución en zonas costeras en comparación con las imágenes LR.

Sin embargo, al observar los índices entre las imágenes LRvsSR y HRvsSR, se puede notar que estas métricas son muy similares. Esto puede indicar que las imágenes SR no son tan similares a las HR. Hay dos factores cruciales que deben tenerse en cuenta para entender esta discrepancia.

En primer lugar, las imágenes HR y LR tienen diferentes procesos de normalización debido a que provienen de fuentes diferentes. Mientras que las imágenes LR ya vienen preprocesadas, las imágenes HR no tienen este preprocesado. Aunque ambas imágenes se han tratado y normalizado de la misma manera para tomarlos como entrada del modelo, esta variación en el preprocesado puede generar discrepancias en la interpretación de los colores, lo que resulta en algunas discrepancias en los resultados.

Por otra parte, las imágenes LR y HR no coinciden exactamente en la hora de captura. Esto puede provocar pequeñas variaciones en el estado de la marea en la zona costera, lo que también puede generar discrepancias entre las imágenes.

Teniendo todos los factores anteriores y observando los resultados, se puede afir-

mar que las imágenes SR producidas por el modelo son imágenes con una base sólida y precisa para realizar análisis y estudios de la línea de costa.

Además de todo lo comentado anteriormente, se ha llevado a cabo la georreferenciación de las imágenes SR, LR y HR con el fin de poder realizar una comparativa entre ellas. Esta comparativa se muestra visualmente en el anexo adjunto, donde se presentan diferentes mapas que ilustran la comparación entre los diferentes tipos de imágenes. Observando dicha comparativa, es posible afirmar que las imágenes SR tienen una resolución espacial mayor que las imágenes LR. Esta mejora en la resolución se puede apreciar tanto en las zonas costeras como en las zonas más interiores de las imágenes. Aunque las imágenes SR no alcanzan la misma nitidez que las imágenes HR.

6.3. Análisis comparativo entre otros artículos científicos

En el artículo original [27] que describe el conjunto de datos OLI2MSI y lo aplica a diferentes arquitecturas, obtuvieron un valor de 0,946 para la métrica SSIM y un valor de 37,68 para el índice PSNR aplicando el modelo EDSR.

Con el modelo EDSR implementado, se lograron unos valores de 0,9871 para el SSIM y un valor de 36,6244 para el PSNR. Estos resultados indican una mejora significativa en el SSIM en comparación con el enfoque presentado en el *paper*. Además, el alto valor de PSNR sugiere una buena calidad de imagen reconstruida.

Por lo tanto, el modelo EDSR obtenido ha demostrado un rendimiento superior al modelo EDSR descrito en el *paper* original en términos de similitud estructural y una muy baja discrepancia en el PSNR.

En cuanto a la comparativa del conjunto de datos propio, no se ha encontrado ningún artículo que se enfoque directamente en la superresolución de imágenes satelitales con un enfoque para obtener imágenes SR a tan baja resolución en zonas costeras.

Sin embargo, sí existen algunos artículos donde se consigue la mejora de la resolución de las imágenes Sentinel-2, como se menciona en [28] o [29], donde la mejora alcanzada en las imágenes satelitales tiene índices SSIM y PSNR adecuados. No obstante, no es posible realizar una comparativa directa debido a que no se aplican las mismas arquitecturas y tampoco se trabaja en zonas costeras.

Es importante tener en cuenta que las zonas costeras pueden afectar notablemente la superresolución debido a las variaciones constantes en el estado de la playa. Por lo tanto, aunque no se haya encontrado una comparativa directa en zonas costeras. Se puede afirmar que los resultados obtenidos en el conjunto de datos propio son buenos, considerando las dificultades asociadas a la implementación.

Sección 5

7. Conclusiones

En conclusión, en este Trabajo Fin de Grado se han desarrollado modelos funcionales aplicando RN. Más concretamente, se han conseguido desarrollar dos modelos.

El modelo EDSR se implementó con éxito en la base de datos pública OLI2MSI, logrando resultados muy positivos al obtener imágenes equivalentes a las de Sentinel-2 a partir de tomas de Landsat-8. En esta implementación se obtuvieron buenas precisiones en los índices PSNR donde se consiguió de media 36,6244 y SSIM donde se consiguió de media 0,9871.

Dichos índices fueron utilizados para cuantificar la correlación entre las imágenes generadas y sus etiquetas, Y observando los valores conseguidos se confirma que el método es adecuado para mejorar la resolución espacial en imágenes de Landsat-8.

El modelo RDN, por otro lado, fue entrenado utilizando imágenes de diferentes resoluciones y logró obtener resultados satisfactorios en la tarea de superresolución en zonas costeras. Aunque los índices de calidad SSIM y PSNR en este caso no fueron tan altos como en el ejemplo anterior, en este caso fueron de un SSIM igual a 0,7514 y un PSNR de 31,1223 de media, Aun así fueron considerados aceptables para las imágenes generadas a partir de Sentinel-2 como entrada del modelo. Además, este modelo demostró una alta precisión también en los índices para el *dataset* OLI2MSI arrojando unos valores de SSIM 0,9821 y PSNR 36,5840 de media, lo que lo convierte en un modelo funcional para trabajar con imágenes de diferentes resoluciones como entrada.

Por otro lado, los valores que se obtuvieron al hacer el análisis de detección de la línea de costa entre los tres tipos de imágenes (HR, LR y SR) son positivos debido a que los índices DICE Y KAPPA son más elevados tanto calculando dichos índices para las imágenes completas como acotando el cálculo de los índices a la línea de costa. Más concretamente, se observa que para ambos casos los índices son mayores en el caso SRvsHR respecto a los índices obtenidos en el caso SRvsLR esto demuestra que la segmentación de línea de costa en las imágenes SR es más precisa que en las imágenes LR.

Por lo tanto, se podría afirmar que todos los objetivos específicos marcados en el inicio de este Trabajo Fin de Grado se han cumplido con éxito. Aunque sí que se podrían llevar a cabo algunas mejoras.

Una posible mejora consiste en aplicar una normalización similar a la de las imágenes satelitales de nivel de procesado 1C en el preprocesado de las imágenes obtenidas a partir de vuelos de avioneta. Esto ayudaría a que las imágenes fotogramétricas tengan características similares a las imágenes satelitales, como los efectos atmosféricos y las interpretaciones de color. Dicha normalización permitiría homogeneizar las imágenes, facilitando su análisis y comparación.

Otra mejora sería en la creación del conjunto de datos propio. Dado que la disponibilidad de imágenes de alta resolución coincidentes en fecha y ubicación con las imágenes de Sentinel-2 era bastante limitada. Para abordar esto, se podrían buscar imágenes de alta resolución en otros portales de descarga, como en el portal de descarga del Instituto Geográfico Nacional (IGN) [30] o incluso en portales de descarga de otros países. Esto permitiría obtener un mayor número de imágenes y proporcionaría al modelo una mayor robustez para trabajar con este tipo de imágenes.

Sección 6

8. Presupuestos

En esta sección se desglosarán los costes necesarios para poder replicar las técnicas de aprendizaje profundo que se han detallado a lo largo de todo el trabajo fin de grado. Para clarificar el entendimiento de los presupuestos vamos a dividirlo en diferentes partidas: partida de personal, partida de gastos informáticos.

8.1. Partida de personal

El personal involucrado en este Trabajo Fin de Grado consta de tres personalidades, de las cuales habrá que tener en cuenta tanto sus costes por hora como sus horas de dedicación a la realización del Trabajo Fin de Grado. El grupo de dichas personalidades está compuesto por el co-tutor del trabajo fin de grado con cargo de investigador en la rama de teledetección, el tutor del Trabajo Fin de Grado con cargo como coordinador de proyectos en investigación en el campo del aprendizaje profundo y un estudiante en el grado de ingeniería geomática y topografía.

Descripción	Horas	Precio (€/h)	Coste
Investigador	35	25	875,00
Investigador coordinador	30	30	900,00
Estudiante	300	7.5	2250,00
Total			4025,00

Tabla 13: Costes de personal

8.2. Partida de gastos informáticos

En cuanto a los gastos informáticos, se tendrá en cuenta el valor del hardware que se ha utilizado para llevar a cabo la realización del Trabajo Fin de Grado. No se contabilizarán la parte correspondiente al software, ya que todos los programas utilizados son de libre acceso y no implicaron costos adicionales.

Durante el proceso de creación y depuración del código, así como la redacción de la memoria, utilicé un ordenador portátil personal. Este equipo fue suficiente para realizar las tareas relacionadas con el desarrollo y la depuración del código, así como para la redacción de la documentación correspondiente.

Sin embargo, para las tareas que requerían un mayor poder de cálculo y recursos computacionales, como el entrenamiento y la evaluación del modelo, recurrí a los

servidores de computación disponibles en el grupo CVBLab. Estos servidores proporcionaron la capacidad de procesamiento necesaria para llevar a cabo estas tareas de manera eficiente, superando las limitaciones de mi propio ordenador portátil.

Descripción	Coste sin IVA (€)	Periodo de amortización (meses)	Intervalo amortizado (meses)	Coste imputable sin IVA (€)
ASUS VivoBook S14 S433EA-EB1152	710,21	24	6	177,55
IVA (21%)				37,29
TOTAL				214,84

Tabla 14: Coste de hardware

8.3. Total

Por lo tanto, el coste total sumando las dos partidas del presupuesto será de 4239,84 €

Referencias

- [1] S. Farsiu, M.D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution, 2004.
- [2] Saeed Anwar, Salman Khan, and Nick Barnes. A deep journey into super-resolution: A survey, 2020.
- [3] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks, 2015.
- [4] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks, 2016.
- [5] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.
- [6] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution, 2017.
- [7] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image super-resolution, 2018.
- [8] Junwei Wang, Kun Gao, Zhenzhou Zhang, Chong Ni, Zibo Hu, Dayu Chen, and Qiong Wu. Multisensor remote sensing imagery super-resolution with conditional gan. *Journal of Remote Sensing*, 2021, 2021.
- [9] Keiron O’Shea and Ryan Nash. An introduction to convolutional neural networks, 2015.
- [10] DataScientest. Vgg: ¿qué es este modelo? daniel te lo cuenta todo. <https://datascientest.com/es/vgg-que-es-este-modelo-daniel-te-lo-cuenta-todo>, Fecha de acceso: 2023-05-26.
- [11] Xiao Zhou, Weizhong Zhang, Zonghao Chen, Shizhe Diao, and Tong Zhang. Efficient neural network training via forward and backward propagation sparsification, 2021.
- [12] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

- [13] Pedro Carvalho, Nuno Lourenço, and Penousal Machado. Evolving learning rate optimizers for deep neural networks, 2021.
- [14] ScienceDirect. Squared error loss. <https://www.sciencedirect.com/topics/computer-science/squared-error-loss>. Accedido el 8 de junio de 2023.
- [15] NASA Goddard Space Flight Center. Landsat 8, Fecha desconocida. URL <https://landsat.gsfc.nasa.gov/satellites/landsat-8/>.
- [16] Agencia Espacial Europea (ESA). Sentinel-2, Fecha desconocida. URL <https://sentinel.esa.int/web/sentinel/missions/sentinel-2>.
- [17] GIS and Beers. Métodos de remuestreo de imágenes raster gis. <http://www.gisandbeers.com/metodos-remuestreo-de-imagenes-raster-gis/>, Fecha de acceso: 23 de mayo de 2023.
- [18] PyPI The Python Package Index. tiff file. <https://pypi.org/project/tiff file/>. Accedido el 8 de junio de 2023.
- [19] NumPy Contributors. Numpy. <https://numpy.org/>, . Accedido el 8 de junio de 2023.
- [20] PyTorch Contributors. Pytorch. <https://pytorch.org/>, . Accedido el 8 de junio de 2023.
- [21] Gabriel Gouvine. Super-resolution networks for pytorch. <https://github.com/Coloquinte/torchSR>, 2021.
- [22] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. doi: 10.1109/TIP.2003.819861.
- [23] Fernando A. Fardo, Victor H. Conforto, Francisco C. de Oliveira, and Paulo S. Rodrigues. A formal evaluation of psnr as quality measurement parameter for image segmentation algorithms, 2016.
- [24] Ilmonteux. Segmentation metrics. <https://ilmonteux.github.io/2019/05/10/segmentation-metrics.html>, 2019. Accedido el 11 de junio de 2023.
- [25] Fisterra. Medidas de concordancia - Índice kappa. <https://www.fisterra.com/formacion/metodologia-investigacion/medidas-concordancia-indice-kappa/>, Accedido el 11 de junio de 2023.

- [26] Sentinel hub eo browser. <https://apps.sentinel-hub.com/eo-browser/>. Accedido el 8 de junio de 2023.
- [27] Junwei Wang, Kun Gao, Zhenzhou Zhang, Chong Ni, Zibo Hu, Dayu Chen, and Qiong Wu. Multisensor remote sensing imagery super-resolution with conditional gan. *Journal of Remote Sensing*, 2021:1–11, 09 2021. doi: 10.34133/2021/9829706.
- [28] Luis Salgueiro Romero, Javier Marcello, and Verónica Vilaplana. Super-resolution of sentinel-2 imagery using generative adversarial networks. *Remote Sensing*, 12(15), 2020. ISSN 2072-4292. doi: 10.3390/rs12152424. URL <https://www.mdpi.com/2072-4292/12/15/2424>.
- [29] Nicolas Latte and Philippe Lejeune. Planetscope radiometric normalization and sentinel-2 super-resolution (2.5 m): A straightforward spectral-spatial fusion of multi-satellite multi-sensor images using residual convolutional neural networks. *Remote Sensing*, 12, 07 2020. doi: 10.3390/rs12152366.
- [30] Centro de descargas del instituto geográfico nacional (ign). <https://centrodedescargas.cnig.es/CentroDescargas/index.jsp>. Accedido el 12 de junio de 2023.

ANEXO I

RELACIÓN DEL TRABAJO CON LOS OBJETIVOS DE DESARROLLO SOSTENIBLE DE LA AGENDA 2030

Objetivos de desarrollo sostenibles	Alto	Medio	Bajo	No procede
ODS 1. Fin de la pobreza.				×
ODS 2. Hambre cero.				×
ODS 3. Salud y bienestar.			×	
ODS 4. Educación de calidad.				×
ODS 5. Igualdad de género.				×
ODS 6. Agua limpia y saneamiento.				×
ODS 7. Energía asequible y no contaminante.				×
ODS 8. Trabajo decente y crecimiento económico.			×	
ODS 9. Industria, innovación e infraestructuras.		×		
ODS 10. Reducción de las desigualdades.				×
ODS 11. Ciudades y comunidades sostenibles.		×		
ODS 12. Producción y consumo responsables.				×
ODS 13. Acción por el clima.			×	
ODS 14. Vida submarina.		×		
ODS 15. Vida de ecosistemas terrestres.		×		
ODS 16. Paz, justicia e instituciones sólidas.				×
ODS 17. Alianzas para lograr objetivos.				×

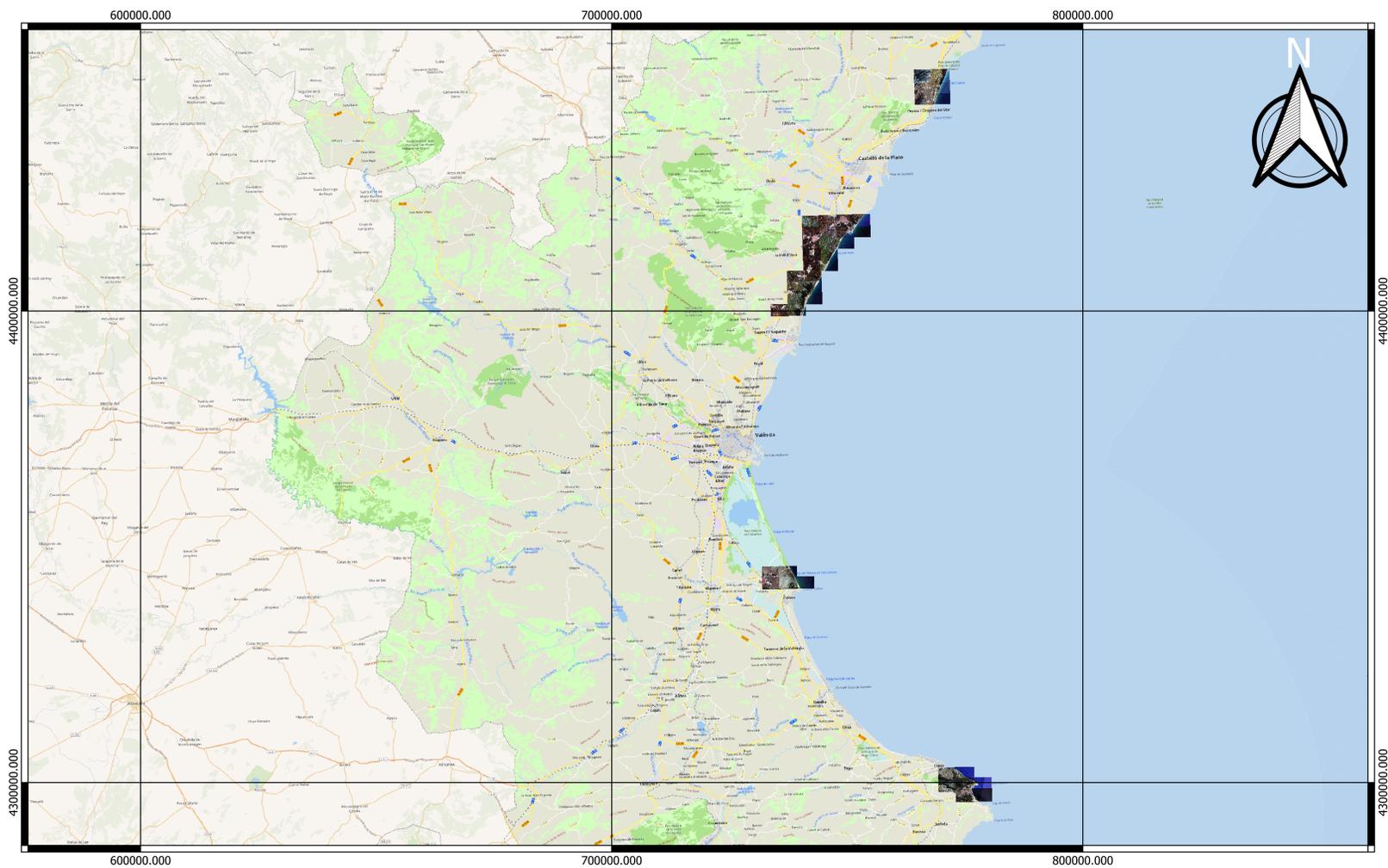
ODS 9: Se considera que tiene un impacto medio debido a que el proyecto se basa en técnicas de aprendizaje profundo con el fin de acometer un objetivo aplicando un abordaje innovador del problema.

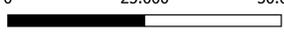
ODS 11: El Trabajo Fin de Grado ayuda a las ciudades y comunidades sostenibles que estén en zonas costeras y tengan una gran influencia por el mantenimiento y estado de sus playas, ya que con el monitoreo continuo del estado de sus playas es de gran ayuda para saber cómo actuar en diversas situaciones y mantener las zonas costeras en buen estado.

ODS 14: Las variaciones del nivel del mar pueden afectar a ecosistemas marinos produciendo alteraciones en la fauna y flora marina en las zonas donde se produzcan dichas variaciones, por eso el mapeo de los cambios es fundamental para actuar ante estos sucesos.

ODS 15: Las variaciones del nivel del mar también afecta a todos los ecosistemas terrestres que se encuentren en zonas costeras, produciendo grandes daños en dichos ecosistemas, por eso el mapeo de los cambios es importante para actuar ante estos sucesos.

ANEXO II
MAPAS



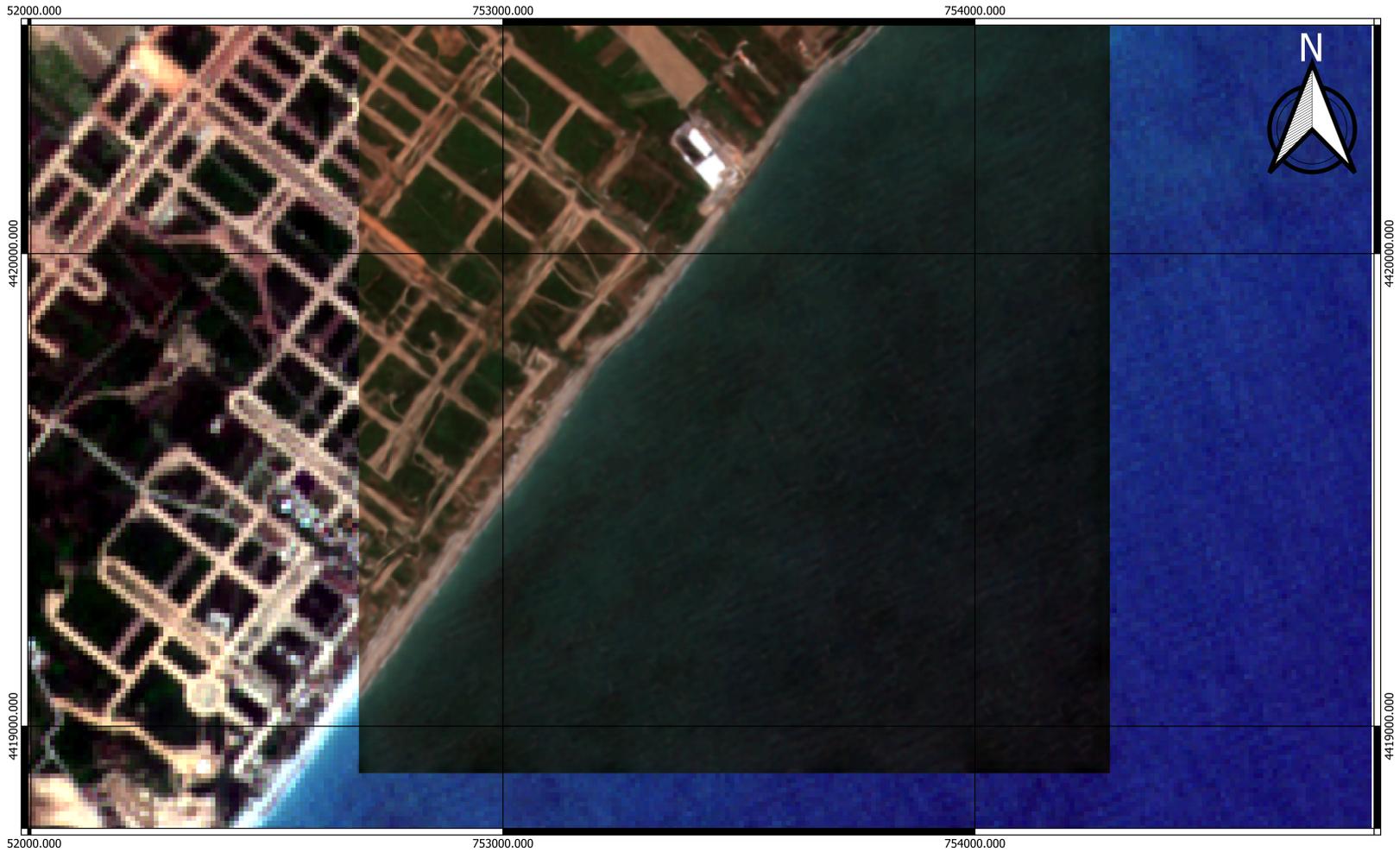
<p>Escala: 70000</p> <p>0 25.000 50.000 m</p> 	<p>Sistema de referencia:</p> <p>EPSG: 25830</p>	<p>Título:</p> <p>Mapa huella del conjunto de datos propio</p>	<p>Autor:</p> <p>Alejandro Diéguez López</p>	<p>Fecha:</p> <p>11/06/2023</p>	 <p>ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA GEODÉSICA CARTOGRÁFICA Y TOPOGRÁFICA</p>
--	---	---	---	--	--



Escala: 1/7000	Sistema de referencia:	Título:	Autor:	Fecha:	
	EPSG: 25830	Mapa de imágenes SR_1 vs LR_1	Alejandro Diéguez López	11/06/2023	



<p>Escala: 1/10000</p> <p>0 250 500 m</p> 	<p>Sistema de referencia:</p> <p>EPSG: 25830</p>	<p>Título:</p> <p>Mapa de imágenes SR_1 vs HR_1</p>	<p>Autor:</p> <p>Alejandro Diéguez López</p>	<p>Fecha:</p> <p>11/06/2023</p>	 <p>ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA GEODÉSICA CARTOGRÁFICA Y TOPOGRÁFICA</p>
--	---	--	---	--	--



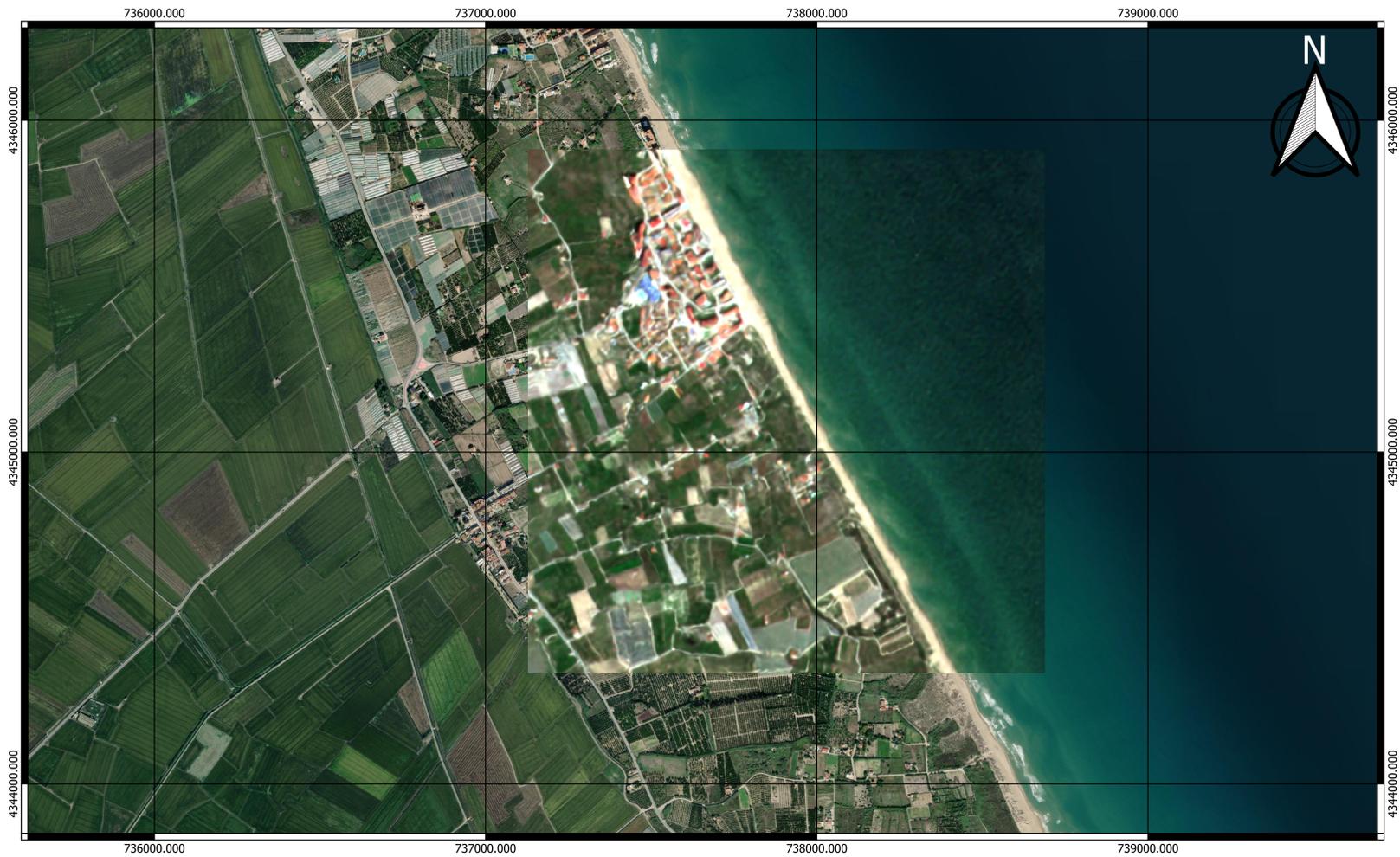
Escala: 1/7000	Sistema de referencia:	Título:	Autor:	Fecha:	
	EPSG: 25830	Mapa de imágenes SR_3 vs LR_3	Alejandro Diéguez López	11/06/2023	



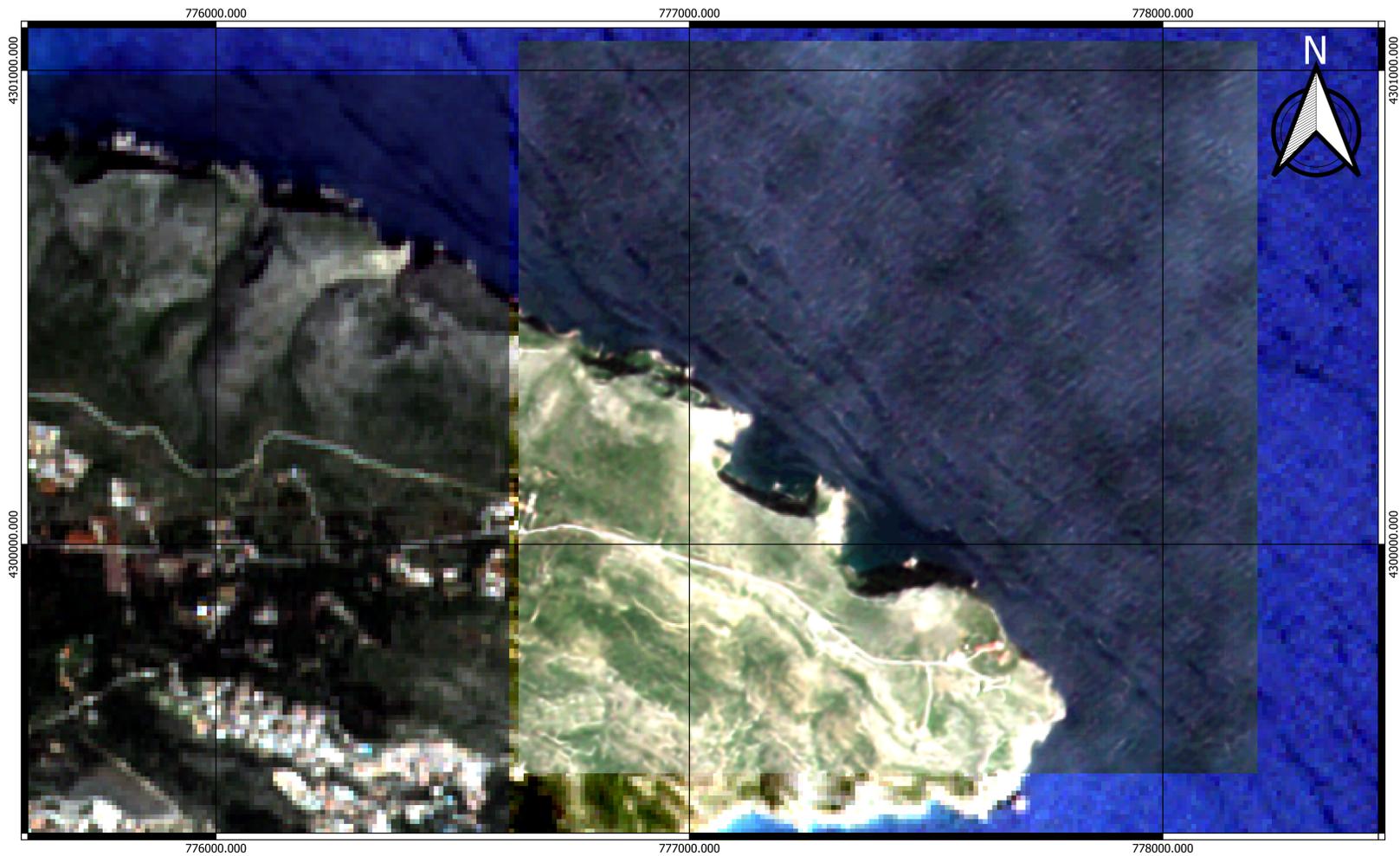
Escala: 1/10000	Sistema de referencia:	Título:	Autor:	Fecha:	 <small>ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA GEODÉSICA CARTOGRÁFICA Y TOPOGRÁFICA</small>
0 250 500 m 	EPSG: 25830	Mapa de imágenes SR_3 vs HR_3	Alejandro Diéguez López	11/06/2023	



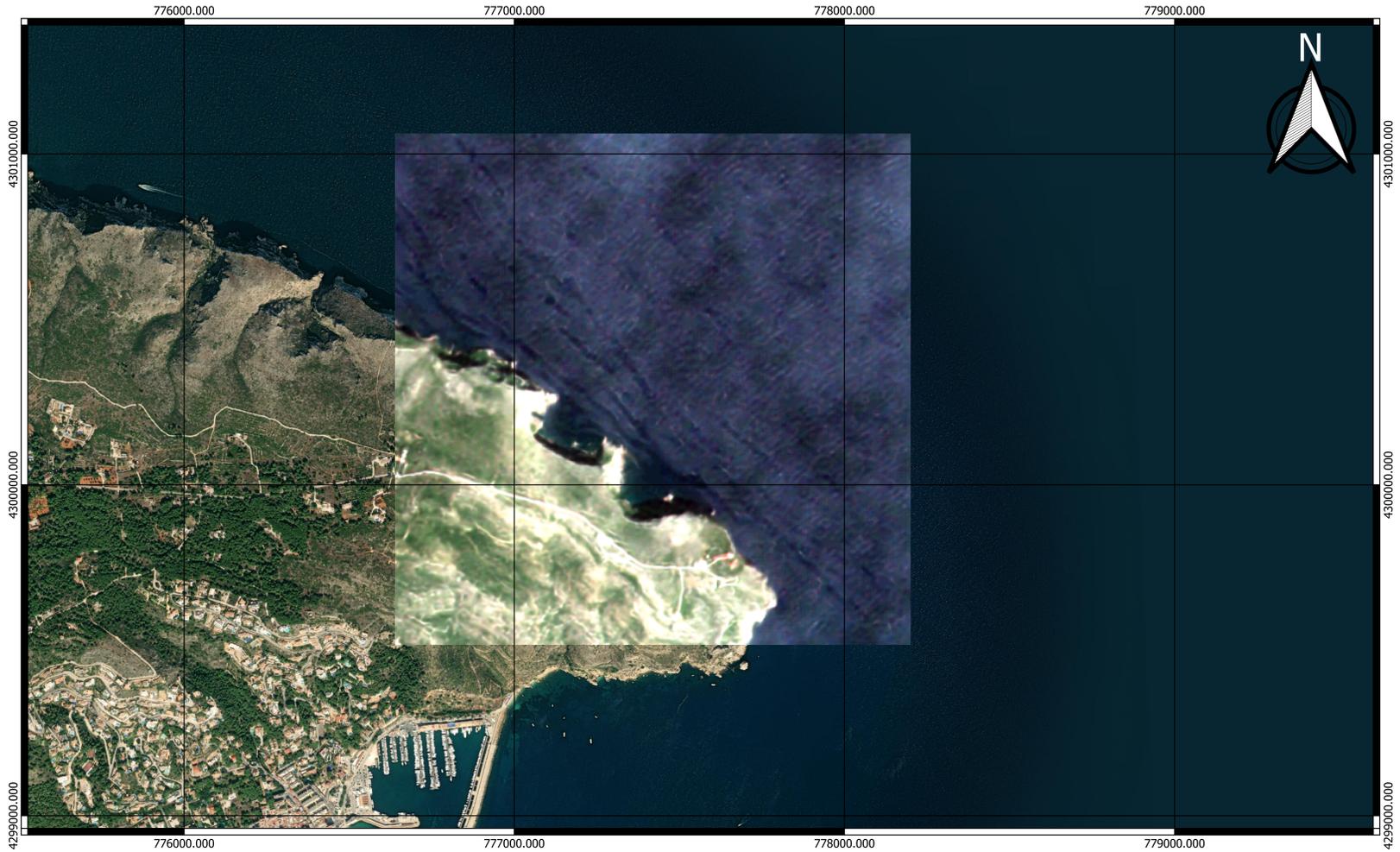
Escala: 1/7000	Sistema de referencia:	Título:	Autor:	Fecha:	 <small>ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA GEODÉSICA CARTOGRAFÍA Y TOPOGRAFÍA</small>
	EPSG: 25830	Mapa de imágenes SR_4 vs LR_4	Alejandro Diéguez López	11/06/2023	



<p>Escala: 1/10000</p> <p>0 250 500 m</p> 	<p>Sistema de referencia:</p> <p>EPSG: 25830</p>	<p>Título:</p> <p>Mapa de imágenes SR_4 vs HR_4</p>	<p>Autor:</p> <p>Alejandro Diéguez López</p>	<p>Fecha:</p> <p>11/06/2023</p>	 <p>ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA GEODÉSICA CARTOGRÁFICA Y TOPOGRÁFICA</p>
--	---	--	---	--	--



Escala: 1/7000	Sistema de referencia:	Título:	Autor:	Fecha:	
	EPSG: 25830	Mapa de imágenes SR_2 vs LR_2	Alejandro Diéguez López	11/06/2023	



Escala: 1/10000	Sistema de referencia:	Título:	Autor:	Fecha:	
0 250 500 m 	EPSG: 25830	Mapa de imágenes SR_2 vs HR_2	Alejandro Diéguez López	11/06/2023	