



Universidad Internacional de la Rioja (UNIR)

**Escuela Superior de Ingeniería y
Tecnología**

Master in Artificial Intelligence

**Enhancing Spatial Resolution of
Sentinel-2 Satellite Images Using
Multispectral Fusion**

Final Master's Project

Presented by: Julio Cesar Contreras Huerta

Supervised by: Miguel Angel Navarro Burgos

City: Valencia

Date: February 2024

TABLE OF CONTENTS

1. Introduction	1
1.1 Motivation	3
1.2 Problem Statement	3
1.3 Estructura del Trabajo	5
2. Contexto y estado del arte	8
2.1 Procesamiento de imágenes satelitales	8
2.1.1 Historia y avances tecnológicos (1970s-1990s)	8
2.1.2 Constelaciones y democratización del acceso (2000s-Presente) . .	8
2.1.3 Tendencias actuales y desafíos	10
2.1.4 Aplicaciones	11
2.1.4.1 Agricultura de precisión	11
2.1.4.2 Gestión de recursos naturales	11
2.1.4.3 Respuesta a emergencias	12
2.1.5 Desafíos	12
2.1.5.1 Resolución espacial y temporal	12
2.1.5.2 Ruido y efectos atmosféricos	12
2.1.5.3 Fusión de datos multisensor	12
2.1.5.4 Manejo de grandes volúmenes de datos	13
2.1.5.5 Accesibilidad y costo	13
2.2 Armonización de datos	13
2.2.1 La misión Sentinel-2	13
2.2.1.1 Resolución espacial	14

2.2.1.2	Resolución espectral	14
2.2.1.3	Resolución temporal	14
2.2.1.4	Resolución temporal	14
2.2.1.5	Niveles de procesamiento	15
2.2.2	Conjuntos de datos para superresolución	16
2.2.2.1	OpenImages	16
2.2.2.2	NAIP	17
2.2.2.3	Sen2Venus	18
2.2.2.4	Sentinel-2 degradado	18
2.2.3	Resolución Espacial Efectiva	19
2.2.3.1	Modulation Transfer Function (MTF)	21
2.2.3.2	MTF en Sentinel-2	22
2.2.3.3	MTF y fusión de imágenes	23
2.2.3.4	MTF y modelos de aprendizaje profundo	23
2.2.4	Corregistro espacial	23
2.2.4.1	Estrategias	24
2.2.4.2	Métodos de alineación	24
2.2.5	Relación entre bandas de Sentinel-2	25
2.2.5.1	Modelado de relaciones entre bandas	25
2.2.5.2	Evaluación de modelos de superresolución	26
2.3	Superresolución de imágenes satelitales	26
2.3.1	Propósito	27
2.3.1.1	Tareas	27
2.3.1.2	Retos actuales	27
2.3.2	Métodos tradicionales	27
2.3.2.1	Interpolación	28
2.3.2.2	Transformaciones de frecuencia	28

2.3.2.3	Enfoques iterativos	28
2.3.2.4	Limitaciones de los métodos tradicionales	28
2.3.3	Métodos basados en aprendizaje profundo	29
2.3.3.1	Superresolución de Imagen Única (SISR)	29
2.3.3.2	Superresolución Multi-Imagen (MISR)	32
2.3.3.3	Comparación entre SISR y MISR	32
2.3.4	Protocolo de Wald	32
2.3.4.1	Criterios del Protocolo de Wald	33
2.3.4.2	Procedimientos de evaluación	33
2.3.4.3	Métricas principales	33
3.	Objetivos y metodología de trabajo	35
3.1	Objetivo General	35
3.2	Objetivos Específicos	35
3.3	Metodología de trabajo	35
3.3.1	Obtención y procesamiento de los datos	36
3.3.1.1	Preprocesamiento de los datos	37
3.3.1.2	Imágenes sin nubes	37
3.3.1.3	Heterogeneidad espectral	38
3.3.1.4	Selección del área de aplicación	38
3.3.2	Metodología general para la fusión	38
3.3.2.1	Preparación de datos	38
3.3.2.2	Filtrado en el dominio de Fourier	39
3.3.2.3	Arquitectura del modelo y SPAN	39
3.3.3	Fusión x2: mejora de resolución intermedia	41
3.3.3.1	Ajustes en la preparación de datos	41
3.3.3.2	Componentes	42

3.3.3.3	Filtrado en el dominio de Fourier	43
3.3.3.4	Configuración del modelo	44
3.3.3.5	Función de pérdida y entrenamiento	46
3.3.4	Fusión x4: mejora de resolución a nivel fino	47
3.3.4.1	Ajustes en la preparación de datos	47
3.3.4.2	Definición de entrada, objetivo y salida	48
3.3.4.3	Filtrado en el dominio de Fourier	48
3.3.4.4	Configuración del modelo	48
3.3.4.5	Integración con el modelo de Fusión x2	51
3.3.4.6	Implementación del filtrado con <i>Hard Constraint</i>	51
3.3.4.7	Función de pérdida y entrenamiento	52
3.3.4.8	Análisis de la pérdida de testeo	54
3.3.4.9	Optimización de la inferencia mediante combinación de convoluciones	55
4.	Identificación de Requisitos	59
4.1	Trabajo Previo Realizado	59
4.1.1	Identificación del Problema	59
4.1.2	Contexto de Uso	59
5.	Descripción de la Herramienta Software Desarrollada	60
5.1	Detalles del Proceso de Desarrollo	60
5.1.1	Fases del Desarrollo	60
5.1.2	Diagramas de Flujo y Arquitectura	60
5.1.3	Capturas de Pantalla	60
6.	Evaluación	62
6.1	Usabilidad de la Herramienta	62

6.2 Aplicabilidad al Problema Propuesto	62
References	63

1. INTRODUCTION

A medida que el mundo se vuelve cada vez más interconectado, la posibilidad de observar y analizar nuestro entorno desde el espacio se ha convertido en una de las herramientas más importantes para enfrentar los desafíos globales. Desde esta perspectiva, son las misiones satelitales, a través de sus avanzados sistemas de sensores, las que permiten la recolección de información sobre la superficie terrestre, proporcionando la visión más detallada y extensa de los procesos de cambio en nuestro planeta. Entre las disciplinas clave que hacen posibles estas funciones, una de las más destacadas es la teledetección, que transforma los datos obtenidos por estos sensores en información crítica para la resolución de problemas ambientales y sociales. Esta disciplina proporciona análisis específicos sobre las transformaciones y características de determinadas áreas geográficas. Con los avances en la IA (Inteligencia Artificial), la capacidad analítica ha alcanzado un nuevo nivel, permitiendo que casi cualquier tipo de imagen satelital sea analizada con una precisión y eficiencia sin precedentes.

La resolución espacial es un parámetro fundamental para el análisis de imágenes de teledetección. Se define como la distancia mínima en el terreno que separa dos objetos independientes que pueden ser distinguidos. Los factores que la determinan incluyen la altitud, la distancia y la calidad de los instrumentos utilizados (Alparone et al., 2015). Un factor adicional importante relacionado con la resolución espacial es la Distancia de Muestreo en el Terreno (GSD), que es la porción de la superficie de la Tierra representada por cada uno de los píxeles (Lillesand et al., 2015).

Uno de los sensores más utilizados debido a su alta resolución espacial es el Sentinel-2 MSI, operado por la European Space Agency (ESA, 2019). Desde su lanzamiento en junio de 2015, ha proporcionado imágenes multiespectrales de acceso abierto, lo que ha generado un interés significativo en la comunidad científica y en diversas industrias, convirtiéndose en una herramienta de clave para la provisión de datos en diversas aplicaciones,

como la monitorización del uso del suelo, la detección de cambios y el análisis de vegetación. El sensor está equipado con 13 bandas espectrales distribuidas en tres resoluciones espaciales: 4 bandas espectrales con una resolución espacial de 10 m, 6 bandas espectrales con 20 m y 3 bandas espectrales con 60 m, diseñadas para la recolección de diversos parámetros topográficos (Lanaras et al., 2018). Esta diversidad espectral permite realizar estudios detallados sobre una amplia gama de fenómenos terrestres mediante la combinación de bandas específicas, desde la calidad del agua hasta el monitoreo de superficies nevadas.

Otro de los sensores que destaca por su mayor resolución es el National Agriculture Imagery Program (NAIP), que proporciona imágenes aéreas con una resolución de hasta 1 metro. Aunque su cobertura está limitada al territorio continental de los Estados Unidos, ofrece imágenes aéreas con una resolución espacial de hasta 1 metro, lo que lo convierte en una fuente importante para estudios que requieren un nivel de detalle superior. NAIP captura imágenes durante los meses de mayor actividad agrícola, de junio a agosto, y proporciona actualizaciones cada tres años (anteriormente cada cinco antes de 2009). Desde 2011, las imágenes incluyen bandas RGB y NIR de manera consistente, mejorando su calidad y su aplicabilidad en áreas como la planificación urbana y el monitoreo agrícola. Aunque NAIP y Sentinel-2 presentan enfoques y alcances diferentes, la combinación de imágenes de ambos sensores permite obtener una mayor riqueza de datos, lo que resulta clave para estudios que requieren tanto una cobertura extensa como un nivel de detalle excepcional.

Existen diversas técnicas para abordar la limitación de la resolución espacial y espectral en las imágenes satelitales dentro del campo de la teledetección, desde métodos tradicionales de fusión de imágenes hasta enfoques más recientes basados en inteligencia artificial. Entre estos últimos, el deep learning ha demostrado ser especialmente prometedor (Gargiulo et al., 2019). Una de las aplicaciones más efectivas de esta técnica es la superresolución, que busca aumentar la resolución espacial de las imágenes generando detalles adicionales a partir de datos existentes. Estas técnicas de machine learning per-

miten reconstruir imágenes de mayor resolución a partir de versiones de menor calidad, optimizando tanto la precisión espacial como la coherencia espectral.

mucho texto

1.1 Motivation

El desarrollo de técnicas para mejorar la resolución de las imágenes obtenidas por el sensor Sentinel-2 ha sido un desafío importante en la comunidad científica. En este proceso, arquitecturas avanzadas de deep learning como las Generative Adversarial Networks (GANs) y las redes neuronales convolucionales (CNNs) juegan un papel crucial, gestionando eficientemente los datos (Salgueiro Romero et al., 2020). Lograr un aumento en la resolución espacial, manteniendo la coherencia espectral y optimizando la calidad de la información, abre nuevas fronteras en el análisis de imágenes satelitales.

El acceso a imágenes gratuitas y de alta calidad, como las proporcionadas por Sentinel-2, es un recurso valioso para la investigación científica. Aunque existen otras formas de obtener imágenes de alta resolución, muchas veces estas no cubren áreas tan extensas o no cuentan con la temporalidad necesaria para ciertos estudios. Gracias a su cobertura global y su alta frecuencia de revisita, este sensor ofrece una ventaja significativa en este sentido. Mejorar su resolución permitiría realizar estudios más detallados sin necesidad de recurrir a sensores costosos, haciendo que los datos disponibles tengan un impacto aún mayor.

Además, [este trabajo busca aportar](#) al avance en la aplicación de IA en teledetección. Al utilizar arquitecturas innovadoras, esta investigación tiene el potencial de abrir nuevas posibilidades para mejorar la resolución de imágenes satelitales. Este enfoque no solo beneficia a la ciencia actual, sino que también puede inspirar futuras investigaciones, permitiendo la continuación del desarrollo de estas técnicas en nuevos contextos.

1.2 Problem Statement

A pesar de la disponibilidad de modelos avanzados de superresolución (Navarro & Sánchez, 2020; Salgueiro Romero et al., 2020), el principal desafío radica en cómo

garantizar que las imágenes resultantes no solo mejoren la resolución espacial, sino que también mantengan la coherencia espectral, un factor crucial en aplicaciones científicas. Para superar las limitaciones de resolución espacial de los sensores como Sentinel-2, en este trabajo se propone el uso de técnicas avanzadas de superresolución basadas en modelos de aprendizaje profundo. La superresolución se ha convertido en una herramienta poderosa para mejorar la calidad de las imágenes, **utilizando arquitecturas como las redes neuronales convolucionales (CNNs) y las Generative Adversarial Networks (GANs)** (**Creo que estás plateando mal ya no menciones CNN o GANS arriba ya lo mencionaste**), que permiten generar imágenes de alta resolución mientras se preserva la coherencia espectral y espacial. Estas técnicas, aplicadas a las imágenes multiespectrales de Sentinel-2, tienen el potencial de generar resultados comparables a los obtenidos con sensores de mayor resolución, como el NAIP, pero con la ventaja de trabajar con datos más accesibles.

En diversos estudios, se ha seguido el Protocolo Wald, un estándar en el campo de la teledetección, para validar la calidad de las imágenes generadas mediante superresolución. Este protocolo establece criterios rigurosos que garantizan que las imágenes mejoradas mantengan consistencia espectral y espacial en relación con las imágenes originales. En este trabajo, se propone un modelo de superresolución (SR) que transforma las bandas de 10m y 20m de Sentinel-2 en imágenes de mayor resolución (2.5m), preservando los detalles espectrales críticos para estudios precisos. El proceso de mejora incluye un paso intermedio de superresolución de 10m a 40m, lo que fortalece la robustez del modelo.

Además, el uso del Protocolo Wald asegura que las imágenes fusionadas conserven tanto la integridad espacial como la espectral, validando la calidad de las mismas. Este enfoque, que separa los procesos de fusión de imágenes y mejora de resolución mediante superresolución, permite que cada etapa trabaje de forma independiente, aumentando la versatilidad y aplicabilidad del método propuesto para diferentes tipos de sensores y aplicaciones.

Esta investigación no solo pretende avanzar en el campo de la superresolución de imágenes, sino también ofrecer una solución viable y accesible para aquellos proyectos que

requieren imágenes de alta resolución, pero que deben operar dentro de las limitaciones económicas y técnicas de los sensores actuales.

1.3 Estructura del Trabajo

- En el **primer capítulo**, se establecen el propósito y la motivación del trabajo. El problema abordado se centra en la necesidad de mejorar la resolución espacial de las imágenes satelitales, específicamente las del Sentinel-2, mediante técnicas de superresolución. El capítulo también describe la organización de la tesis, delineando las principales secciones y objetivos.
- El **segundo capítulo** profundiza en un análisis detallado del contexto, examinando el estado actual de las tecnologías de imágenes satelitales y la demanda de datos de alta resolución en diversos campos, como la monitorización ambiental y la planificación urbana. Se discuten los beneficios y desafíos de mejorar la resolución espacial mediante la fusión de datos multiespectrales y modelos de superresolución, con un enfoque en las limitaciones de las técnicas existentes.
- El **tercer capítulo** revisa el estado del arte en tecnologías de superresolución aplicadas a imágenes satelitales. El capítulo incluye una visión general de los métodos clave como el *pansharpening*, la inversión de modelos de imagen y los enfoques de aprendizaje profundo. También se destaca la importancia del *Protocolo de Wald* como un marco de validación para garantizar la consistencia espectral y espacial en las imágenes de alta resolución generadas.
- En el **cuarto capítulo**, se definen los objetivos generales y específicos del proyecto. El objetivo principal es desarrollar un modelo de superresolución capaz de mejorar las imágenes del Sentinel-2 de 10m y 20m de resolución a una resolución de 2.5m, utilizando una combinación de modelos basados en redes neuronales convolucionales (CNN) y técnicas de fusión de datos multiespectrales. La sección de metodología detalla el proceso seguido para el entrenamiento del modelo, la preparación del con-

junto de datos y la evaluación utilizando el *Protocolo de Wald*.

- **El quinto capítulo** presenta los requisitos funcionales y no funcionales del sistema de superresolución. Los requisitos funcionales incluyen la capacidad de manejar imágenes multiespectrales de diversas resoluciones y la generación de resultados consistentes y de alta calidad con una resolución de 2.5m. Los requisitos no funcionales se centran en la eficiencia, escalabilidad y robustez del sistema al manejar grandes conjuntos de datos.
- En **el sexto capítulo**, se explica en detalle el desarrollo del modelo de superresolución. Esto incluye el proceso de entrenamiento, el diseño de las arquitecturas de redes neuronales (Modelo de Fusión X2 y X4) y la integración de técnicas de aprendizaje profundo. Las limitaciones encontradas durante la fase de entrenamiento, como las restricciones de memoria y el rendimiento del modelo, se discuten junto con posibles optimizaciones.
- **El séptimo capítulo** evalúa el rendimiento del modelo desarrollado, comparándolo con técnicas existentes. La evaluación se basa en métricas tanto cualitativas como cuantitativas, centrándose en la fidelidad espacial y espectral de las imágenes de alta resolución generadas. El capítulo también incluye un análisis de la usabilidad del sistema y su impacto potencial en aplicaciones del mundo real, como la agricultura de precisión y el monitoreo del uso del suelo.
- **El octavo capítulo** concluye el trabajo, resumiendo las principales contribuciones de la tesis y proponiendo futuras direcciones de investigación. Las posibles mejoras incluyen el perfeccionamiento de la arquitectura de aprendizaje profundo, la exploración de otras técnicas de fusión y la aplicación del modelo a diferentes tipos de imágenes satelitales.
- **El Apéndice I** incluye imágenes de salida generadas por el modelo de superresolución, demostrando la resolución mejorada para varias bandas del Sentinel-2.

- **El Apéndice II** presenta los resultados de las pruebas de validación del *Protocolo de Wald*, mostrando el rendimiento del sistema en el mantenimiento de la consistencia espectral y espacial.
- **El Apéndice III** proporciona documentación detallada de las herramientas de software y bibliotecas utilizadas para la implementación del modelo.
- **El Apéndice IV** contiene el artículo de investigación asociado con esta tesis, enviado para su publicación.

2. CONTEXTO Y ESTADO DEL ARTE

2.1 Procesamiento de imágenes satelitales

El procesamiento de imágenes satelitales desempeña un papel clave en la observación terrestre y el monitoreo ambiental. A lo largo de su evolución, este campo ha integrado innovaciones tecnológicas, datos multiespectrales y métodos avanzados, como el aprendizaje profundo, para abordar problemas complejos. En esta sección, se analiza el contexto histórico, los avances tecnológicos, las aplicaciones actuales y las tendencias que enfrentan las tecnologías satelitales, destacando los desafíos que persisten en su desarrollo.

2.1.1 Historia y avances tecnológicos (1970s-1990s)

La evolución de los satélites de observación de la Tierra comenzó con hitos como el lanzamiento del *Landsat 1* en 1972, que introdujo imágenes multiespectrales y permitió avances en cartografía y gestión de recursos naturales. Este satélite marcó el inicio de un enfoque sistemático en la observación terrestre, integrando innovaciones como sensores multiespectrales y la adquisición de datos digitales (Wulder et al., 2022).

Durante las décadas de 1980 y 1990, los avances tecnológicos dieron lugar a sensores más avanzados, como los radiómetros y los radares de apertura sintética (SAR). Satélites como *SPOT* y *ERS* permitieron nuevas aplicaciones en meteorología, oceanografía y monitoreo de regiones polares. Por ejemplo, los sensores SAR demostraron ser efectivos para mapear regiones cubiertas por nubes o en oscuridad prolongada, mientras que los altímetros proporcionaron mediciones precisas de la topografía oceánica (Amani et al., 2022). Estos avances sentaron las bases para las capacidades multiesensor actuales.

2.1.2 Constelaciones y democratización del acceso (2000s-Presente)

El desarrollo de programas como *Copernicus* y los satélites *Sentinel* marcó un cambio significativo en la democratización del acceso a datos satelitales de alta calidad mediante modelos de datos abiertos. Esto impulsó aplicaciones innovadoras en agricultura de

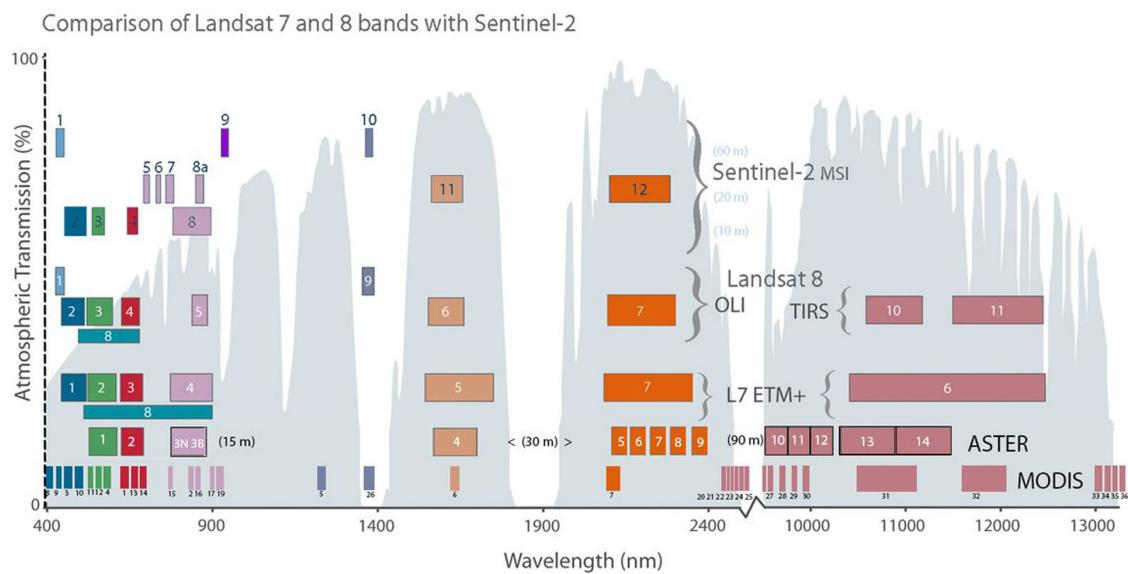
precisión, monitoreo ambiental y planificación urbana, mientras fomentó la participación ciudadana mediante iniciativas de *crowdsourcing* y sensores accesibles (Karagiannopoulou et al., 2022).

La integración de constelaciones masivas, como las gestionadas por el servicio de emergencias de *Copernicus*, también ha mejorado la respuesta a emergencias al proporcionar datos críticos en tiempo casi real. Estas iniciativas han optimizado la precisión y alcance de las acciones en todas las fases de gestión de desastres (Denis et al., 2016).

Por otro lado, las constelaciones privadas como *Planet Labs* han transformado la observación terrestre al combinar alta resolución temporal y espacial con análisis avanzados de *big data*. Estas plataformas han facilitado estudios detallados, como el monitoreo de la conectividad hídrica en regiones árticas (Cooley et al., 2017).

Figura 1.

Comparación de las bandas espectrales de satélites de observación terrestre.



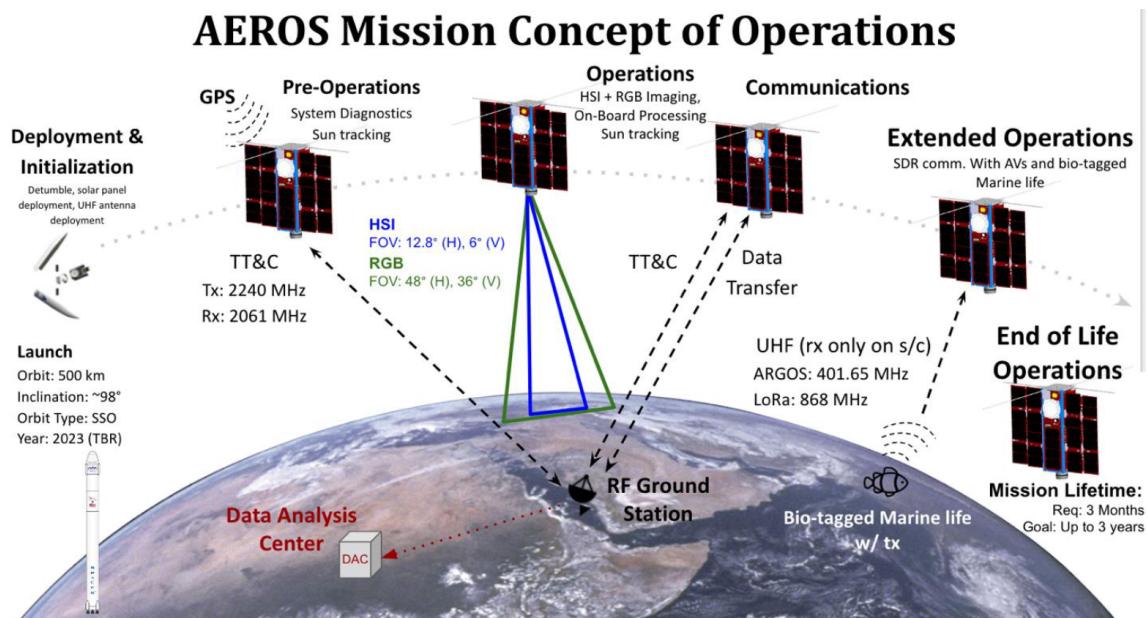
Nota. La figura ilustra cómo la diversidad de bandas espectrales en satélites como *Landsat-7*, *Landsat-8* y *Sentinel-2* contribuyen a aplicaciones específicas, como el monitoreo ambiental y la planificación urbana. Recuperado de Landsat (2016).

2.1.3 Tendencias actuales y desafíos

El futuro del procesamiento de imágenes satelitales está marcado por la miniaturización de satélites, como los CubeSats, y el surgimiento de constelaciones privadas. Los CubeSats, pequeños satélites modulares de 10 cm³, han ampliado el acceso a la observación terrestre al ofrecer alternativas económicas y de fácil implementación. Estas plataformas permiten un monitoreo continuo en regiones remotas y vulnerables, mejorando la capacidad de respuesta ante desastres naturales (Santilli et al., 2016).

Figura 2.

Diagrama de Operaciones de la Misión AEROS.



Nota. Este diagrama destaca cómo los *CubeSats*, como AEROS, contribuyen al monitoreo ambiental con capacidades avanzadas de adquisición y análisis de datos. Recuperado de Prendergast et al. (2022).

Sin embargo, estos avances presentan desafíos como la generación de desechos orbitales, las limitaciones en la miniaturización de sensores y la sostenibilidad financiera. Superar estos obstáculos requiere colaboraciones público-privadas y políticas que promuevan la maniobrabilidad y desorbitación segura (Denis et al., 2017).

2.1.4 Aplicaciones

2.1.4.1 Agricultura de precisión. La tecnología de imágenes satelitales, proporcionada por sistemas como Sentinel-2 y Landsat, permite monitorear el estado de los cultivos, detectar estrés hídrico y estimar rendimientos agrícolas. Índices espectrales como el NDVI (*Normalized Difference Vegetation Index*) identifican áreas de baja productividad y optimizan el uso de recursos, como agua y fertilizantes (ED Chaves et al., 2020; Griffiths et al., 2019).

Particularmente, un estudio en el sudeste asiático destacó cómo el uso combinado de Sentinel-2 y tecnologías como el SAR optimizó la gestión agrícola bajo diferentes condiciones. Por un lado, índices como el NDVI y el EVI (*Enhanced Vegetation Index*) permitieron reducir el desperdicio de agua en cultivos de arroz en un 20 %, además de ajustar los tiempos de cosecha para maximizar rendimientos. Por otro lado, el SAR facilitó el monitoreo de áreas agrícolas bajo condiciones de nubosidad persistente, complementando los datos ópticos de Sentinel-2 para mejorar el análisis de humedad del suelo y el seguimiento de cultivos en zonas con alta cobertura nubosa (Kaushik et al., 2021).

2.1.4.2 Gestión de recursos naturales. La monitorización de recursos naturales a través de datos satelitales proporciona una supervisión continua de cuerpos de agua, recursos minerales y zonas forestales. Esto facilita estrategias de conservación basadas en evidencia científica. Por ejemplo, el programa Copernicus ha sido clave para evaluar el impacto de la deforestación en la región amazónica, permitiendo la implementación de políticas de reforestación en áreas críticas (Rast et al., 2019).

Además, en la gestión de cuencas hidrográficas, la integración de datos satelitales con modelos geoespaciales ha optimizado la toma de decisiones. Un ejemplo notable es el uso combinado de imágenes ópticas y radar para monitorear la calidad del agua en el río Ganges, mejorando la identificación de fuentes de contaminación y el control de flujos hídricos en tiempo real (Caballero et al., 2019; Lisboa et al., 2024).

2.1.4.3 Respuesta a emergencias. Las imágenes satelitales de alta resolución, combinadas con datos en tiempo real, son fundamentales para responder a desastres naturales como inundaciones y terremotos. Por ejemplo, durante las inundaciones en el sudeste asiático en 2021, la combinación de datos de Planet Labs y Sentinel permitió identificar áreas afectadas en menos de 48 horas, optimizando la distribución de ayuda humanitaria (Matgen et al., 2020).

Una aplicación destacada es el uso de Sentinel-1 para evaluar la extensión de incendios forestales mediante tecnologías de radar, incluso bajo condiciones de humo denso. Este enfoque mejora la precisión del mapeo de áreas quemadas y facilita una planificación efectiva para la recuperación de los ecosistemas (Ajmar et al., 2017).

2.1.5 Desafíos

2.1.5.1 Resolución espacial y temporal. La precisión de las aplicaciones depende de la resolución espacial y temporal de los datos. Mientras que satélites como Sentinel-2 ofrecen una resolución espacial de 10 metros, esto puede ser insuficiente para análisis detallados, como la detección de cultivos individuales. En contraste, sistemas privados como PlanetScope proporcionan imágenes con resolución de hasta 30 cm por píxel, aunque su alto costo sigue siendo una barrera para su adopción generalizada (Schiavon et al., 2021).

2.1.5.2 Ruido y efectos atmosféricos. Los efectos atmosféricos, como la presencia de nubes y aerosoles, afectan la calidad de las imágenes ópticas, reduciendo su utilidad en ciertas aplicaciones. No obstante, tecnologías como el radar de apertura sintética (SAR) ofrecen soluciones robustas frente a estas limitaciones, mejorando el monitoreo en regiones con condiciones climáticas adversas (Caballero et al., 2019).

2.1.5.3 Fusión de datos multisensor. Combinar datos ópticos, térmicos y de radar sigue siendo un desafío técnico debido a las diferencias en resoluciones espaciales y espectrales. Sin embargo, estas herramientas están logrando avances importantes, particularmente en aplicaciones agrícolas donde se requiere combinar datos de humedad del

suelo con índices espectrales (Rast et al., 2019).

2.1.5.4 Manejo de grandes volúmenes de datos. La cantidad de datos generados por satélites modernos plantea retos significativos en almacenamiento y análisis, especialmente en regiones con recursos tecnológicos limitados. Este problema se agrava en estudios que requieren integración de múltiples fuentes (Matgen et al., 2020).

2.1.5.5 Accesibilidad y costo. Aunque programas como Copernicus han mejorado el acceso a datos satelitales, las imágenes de muy alta resolución, ofrecidas por empresas privadas, siguen siendo una barrera económica para su adopción generalizada (Denis et al., 2016).

2.2 Armonización de datos

La armonización de datos busca garantizar la coherencia entre imágenes satelitales con distintas resoluciones espaciales, espirituales y temporales. Este proceso es esencial para aplicaciones avanzadas en teledetección, como la superresolución y la fusión multi-espectral. La armonización establece la base para técnicas avanzadas, como la superresolución, que requieren datos consistentes para lograr precisión en tareas específicas. En esta sección, se analizan los aspectos fundamentales relacionados con las características de la misión Sentinel-2, los conjuntos de datos empleados en el contexto de la superresolución y las técnicas relevantes para mejorar la calidad y utilidad de las imágenes.

2.2.1 *La misión Sentinel-2*

La misión Sentinel-2 forma parte del programa Copernicus liderado por la Agencia Espacial Europea (ESA). Su objetivo principal es proporcionar datos multiespectrales de alta calidad para monitorear los cambios en la superficie terrestre, incluyendo agricultura, forestación, y recursos hídricos. Es clave en la gestión de emergencias y la vigilancia medioambiental gracias a su amplio alcance y frecuencia de revisita (Lanaras et al., 2018; Q. Wang et al., 2016).

2.2.1.1 Resolución espacial. Proporciona imágenes con resoluciones espaciales de 10, 20 y 60 metros, dependiendo de las bandas espectrales. Las bandas de 10 metros incluyen el visible (RGB), mientras que las bandas de 20 metros abarcan regiones como el infrarrojo cercano (NIR) y el infrarrojo de onda corta (SWIR). Sin embargo, estas bandas de 20 metros, aunque frecuentemente utilizadas para aplicaciones como el monitoreo del agua o la clasificación de cultivos, presentan limitaciones en cuanto al detalle espacial. Para abordar este desafío, se han propuesto soluciones basadas en redes neuronales convolucionales, como FUSE, que generan superresoluciones eficientes al combinar bandas de alta y baja resolución, asegurando la preservación de las características espectrales y estructurales (Gargiulo et al., 2019). Además, para aprovechar al máximo las capacidades del sensor, investigaciones recientes han explorado cómo las correlaciones entre bandas de diferente resolución pueden emplearse para mejorar las bandas de 20 y 60 metros a través de técnicas de superresolución basadas en aprendizaje profundo (Lanaras et al., 2018).

2.2.1.2 Resolución espectral. Sentinel-2 abarca 13 bandas espectrales, desde el ultravioleta hasta el infrarrojo de onda corta. Estas bandas permiten una caracterización detallada de la superficie terrestre, incluyendo la detección de cambios en la vegetación, la humedad del suelo y la calidad del agua (Q. Wang et al., 2016).

2.2.1.3 Resolución temporal. Compuesta por los satélites Sentinel-2A y Sentinel-2B, la misión ofrece un tiempo de revisita de 5 días, permitiendo una vigilancia casi continua, especialmente útil en estudios dinámicos. Se prevé el lanzamiento de Sentinel-2C en el futuro para ampliar la constelación (European Space Agency, 2024). Esto asegurará la continuidad en la adquisición de imágenes y mantendrá la calidad de los datos para una amplia gama de aplicaciones.

2.2.1.4 Resolución temporal. La misión Copernicus Sentinel-2 actualmente cuenta con tres satélites en órbita: Sentinel-2A, Sentinel-2B y, más recientemente, Sentinel-2C, lanzado el 4 de septiembre de 2024 (European Space Agency, 2024).

Con el lanzamiento de Sentinel-2C, que reemplazará gradualmente a Sentinel-2A, y el futuro Sentinel-2D que tomará el relevo de Sentinel-2B, se garantiza la continuidad de datos esenciales con una cobertura cada cinco días. Este avance fortalece el programa Copernicus en la observación de la Tierra y asegura la sostenibilidad de la misión más allá de 2035 con las futuras misiones Sentinel-2 Next Generation.

2.2.1.5 Niveles de procesamiento. Sentinel-2 ofrece imágenes en dos niveles: L1C y L2A (Ginting et al., 2024).

L1C. Datos en reflectancia TOA con correcciones geométricas y radiométricas, útiles para análisis generales, pero sin corrección atmosférica. Estas imágenes son adecuadas para estudios espaciales básicos, aunque pueden estar influenciadas por efectos atmosféricos.

L2A. Datos BOA corregidos atmosféricamente con Sen2Cor, ideales para estudios precisos como monitoreo de cultivos y agua. Estas imágenes eliminan los efectos de la atmósfera y el terreno, pero pueden presentar sobre correcciones en áreas con topografía compleja debido a limitaciones en el modelo de elevación digital (DEM).

Elección.

- **L1C:** Para análisis espaciales generales sin necesidad de corrección atmosférica.
- **L2A:** Para estudios que requieren alta precisión en reflectancia superficial corregida.

Tabla 1.

Comparación de bandas espectrales de Sentinel-2.

Designación de banda	Resolución (m)	Longitud de onda (nm)	Descripción
B1	60	443.9–442.3	Aerosols
B2	10	496.6–492.1	Blue
B3	10	560.0–559.0	Green
B4	10	664.5–665.0	Red
B5	20	703.9–703.8	Red Edge 1
B6	20	740.2–739.1	Red Edge 2
B7	20	782.5–779.7	Red Edge 3
B8	10	835.1–833.0	NIR
B8A	20	864.8–864.0	Red Edge 4
B9	60	945.0–943.2	Water vapor
B11	20	1613.7–1610.4	SWIR 1
B12	20	2202.4–2185.7	SWIR 2

Nota. Comparación de las bandas espectrales, resoluciones espaciales y longitudes de onda del satélite Sentinel-2. Adaptado de Gargiulo et al. (2019).

2.2.2 Conjuntos de datos para superresolución

La selección adecuada de conjuntos de datos desempeña un rol esencial en el desarrollo y rendimiento de algoritmos de superresolución. Estos conjuntos presentan una variedad de características, desde resolución espacial ypectral hasta diversidad geográfica, que impactan en la eficacia de los modelos. A continuación, se describen los principales conjuntos de datos utilizados en este ámbito, destacando sus aplicaciones y limitaciones.

2.2.2.1 OpenImages. Diseñado originalmente para tareas de visión por computadora, incluye una vasta colección de imágenes RGB que representan escenas cotidianas. Aunque su contenido no está relacionado directamente con la teledetección, este dataset es valioso para el preentrenamiento de modelos debido a su amplio volumen y accesibilidad bajo licencia CC-BY. Las imágenes de alta resolución (HR) alcanzan hasta 0.5 metros, mientras que las de baja resolución (LR) se generan sintéticamente mediante interpolación

para alcanzar dimensiones de 128×128 píxeles.

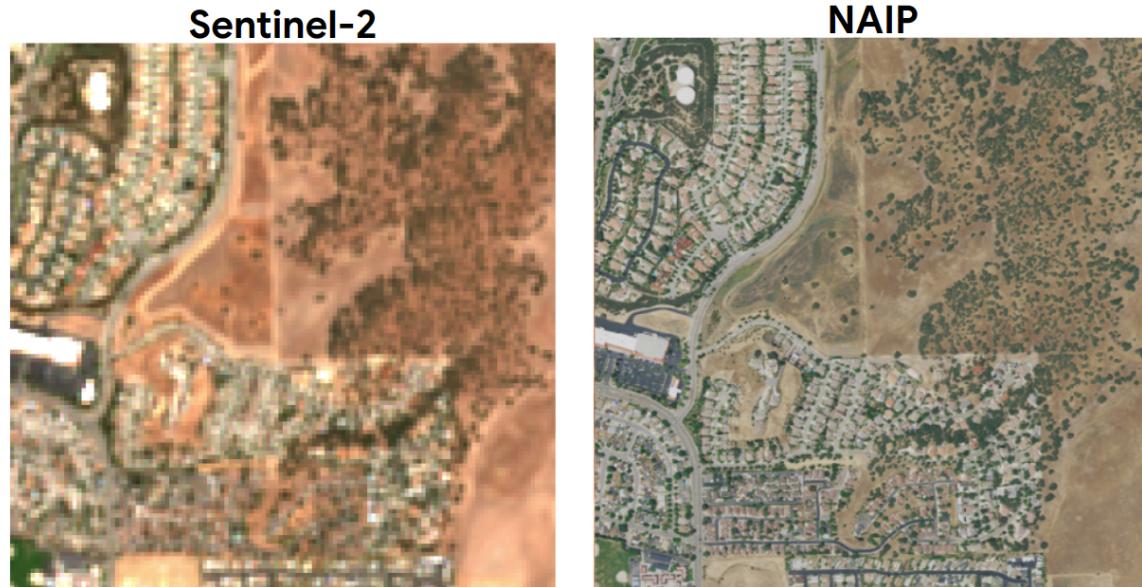
Pese a las diferencias espectrales y espaciales con respecto a los datos de teledetección, OpenImages ha demostrado ser útil en la transferencia de aprendizaje, permitiendo que modelos preentrenados en este conjunto se ajusten eficazmente a dominios específicos como Sentinel-2 (Gargiulo et al., 2019).

2.2.2.2 NAIP. El Programa Nacional de Imágenes de Agricultura (NAIP) recopila imágenes aéreas multiespectrales de alta resolución (0.6 metros) en los Estados Unidos. Estas imágenes, utilizadas principalmente para aplicaciones agrícolas, incluyen bandas espectrales como rojo, verde, azul e infrarrojo cercano (NIR). NAIP es un recurso abierto que proporciona datos actualizados regularmente, facilitando análisis detallados en agricultura de precisión y gestión de tierras.

En contextos de superresolución, el NAIP se emplea como referencia para comparar su alta resolución con las imágenes de Sentinel-2, permitiendo validar modelos que buscan mejorar la resolución espacial (Gargiulo et al., 2019).

Figura 3.

Comparación visual entre Sentinel-2 y NAIP.



Nota. La imagen muestra una comparación entre Sentinel-2 y NAIP. NAIP destaca por su mayor nivel de detalle, mientras que Sentinel-2 proporciona cobertura multiespectral útil para análisis globales y temporales. Elaboración propia.

2.2.2.3 Sen2Venus. El conjunto de datos Sen2Venus está diseñado específicamente para la mejora de resolución espacial de las bandas de Sentinel-2. Este dataset combina imágenes de Sentinel-2 con datos de referencia adquiridos por el satélite VENµS, que proporciona una resolución de 5 metros. Los datos, libres de nubes, abarcan 29 ubicaciones globales con un total de 132,955 parches de 256×256 píxeles.

Sen2Venus es clave para la armonización de datos multiespectrales, permitiendo pruebas directas de algoritmos de superresolución que buscan preservar la coherencia espectral y espacial (Michel et al., 2022).

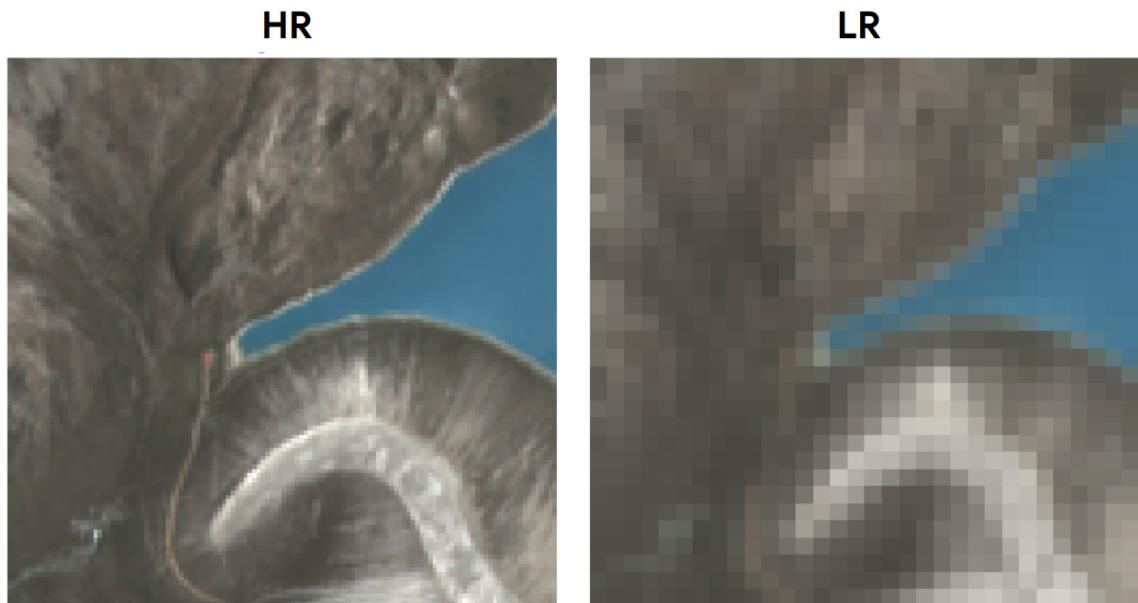
2.2.2.4 Sentinel-2 degradado. El conjunto Sentinel-2 degradado genera pares sintéticos de imágenes LR-HR en el dominio espectral de Sentinel-2. Las imágenes de alta resolución, originalmente de 10 metros, se degradan mediante un núcleo de convolución

para obtener pares con resoluciones de 40 y 10 metros, correspondientes a un factor de superresolución de 4.

Este dataset supera problemas comunes como desajustes geométricos y espectrales, proporcionando una base sólida para entrenar y evaluar modelos de superresolución desde cero. Sin embargo, las diferencias en la frecuencia espacial entre los objetos del suelo y los datos de Sentinel-2 representan un desafío adicional (Q. Wang et al., 2016).

Figura 4.

Ejemplo del conjunto de datos degradado Sentinel-2 (HR: 10m, LR: 40m).



Nota. La figura ilustra la degradación de una imagen de alta resolución (HR) a baja resolución (LR) mediante un núcleo de convolución. Este proceso se utiliza comúnmente en el desarrollo de algoritmos de superresolución para entrenar modelos capaces de restaurar detalles espaciales a partir de datos degradados. Elaboración propia.

2.2.3 Resolución Espacial Efectiva

La resolución espacial efectiva es un parámetro fundamental en la calidad de las imágenes satelitales, ya que refleja la capacidad del sistema de capturar detalles espaciales más allá de las limitaciones geométricas del píxel. Aunque a menudo se relaciona directamente con la resolución espacial nominal, la efectividad depende de factores como la calidad óptica del sensor, la calidad del procesamiento y las condiciones ambientales.

mente con el tamaño del píxel (GSD, *Ground Sampling Distance*), su definición práctica depende de factores como las propiedades ópticas del sistema y la dispersión de la energía luminosa en los píxeles, representada por la PSF (*Point Spread Function*).

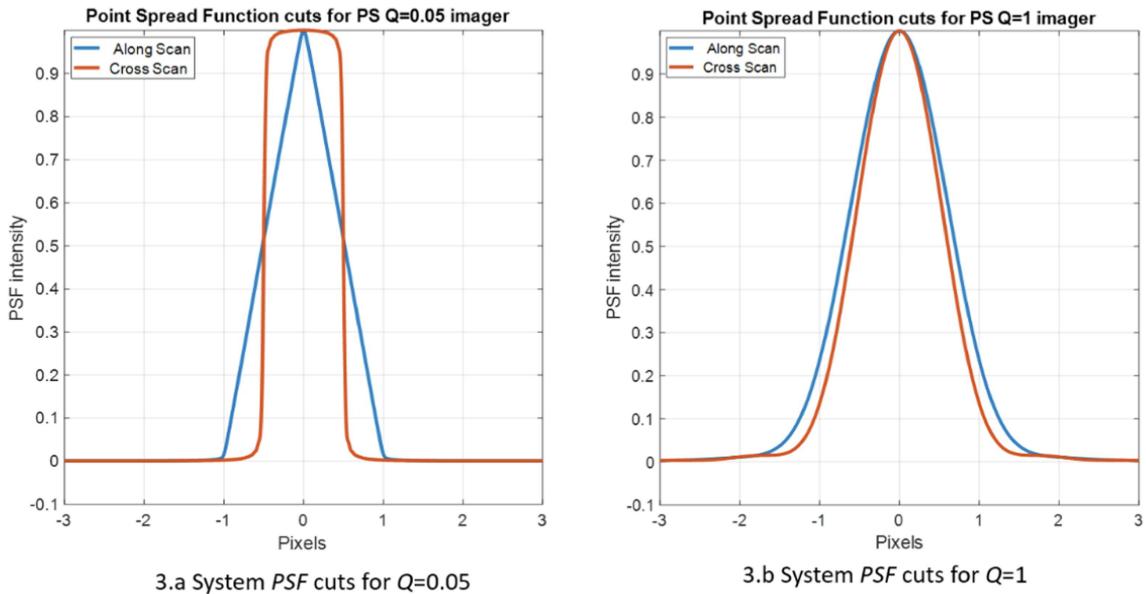
- **Propiedades ópticas del sistema:** Elementos como la apertura del lente y la longitud focal determinan la cantidad de detalles que pueden capturarse, influenciando directamente la resolución espacial efectiva (Valenzuela et al., 2022).
- **PSF (Point Spread Function):** Describe cómo la energía de una fuente puntual se distribuye entre varios píxeles. La PSF no solo modela las limitaciones geométricas, sino que también evalúa la calidad del sistema en términos de precisión óptica y electrónica (Pampanoni et al., 2024).

En este contexto, la resolución espacial efectiva puede describirse en términos de diferentes configuraciones de muestreo, basadas en el parámetro Q , que representa la relación entre el tamaño del píxel y la extensión de la PSF:

- **Bajo muestreo ($Q < 0.7$):** La señal se dispersa significativamente, perdiendo detalles espaciales críticos.
- **Muestreo óptimo ($0.7 \leq Q \leq 1.5$):** Logra un equilibrio entre el tamaño del píxel y la dispersión de la señal, maximizando la resolución efectiva.
- **Sobre-muestreo ($Q > 1.5$):** Aunque preserva los detalles, puede resultar en un uso ineficiente de los recursos de almacenamiento y procesamiento.

Figura 5.

Distribución de la PSF y su impacto en la resolución espacial efectiva.



Nota. Distribución de la PSF en las direcciones Along Scan y Cross Scan para condiciones de muestreo insuficiente ($Q = 0.05$) y óptimo ($Q = 1$), mostrando cómo se afecta la preservación de detalles espaciales. Recuperado de Valenzuela et al. (2022).

Este análisis subraya la importancia de optimizar el diseño del sistema para alcanzar un equilibrio entre resolución geométrica y calidad óptica. Este concepto conecta directamente con la evaluación de la MTF (*Modulation Transfer Function*), ya que esta permite comprender cómo las frecuencias espaciales se preservan en el proceso de captura de imágenes.

2.2.3.1 Modulation Transfer Function (MTF). Es una métrica clave para evaluar cómo un sistema óptico o sensor de imágenes preserva las frecuencias espaciales, desde patrones más amplios hasta detalles finos. Se determina mediante la Transformada de Fourier de la Función de Dispersión de Punto (PSF), y se define como:

$$\text{MTF}(f) = |\mathcal{F}\{\text{PSF}(x, y)\}|$$

Donde:

- f : frecuencia espacial medida en ciclos por unidad de distancia.
- \mathcal{F} : Transformada de Fourier de la PSF.
- $\text{PSF}(x, y)$: representación espacial de la dispersión de luz para una fuente puntual.

Un valor alto en la MTF indica una alta capacidad para transmitir detalles de alta frecuencia espacial, mientras que valores bajos sugieren una pérdida de resolución y disminución de nitidez. Estudios recientes han demostrado que las técnicas basadas en la MTF son fundamentales para mejorar la calidad de imágenes satelitales, como las de Sentinel-2, al incorporar modelos de degradación que reflejan fielmente la capacidad del sensor para capturar diferentes rangos de frecuencias espaciales (Hubble, 2021).

2.2.3.2 MTF en Sentinel-2. En el caso del satélite Sentinel-2, la *Modulation Transfer Function* (MTF) varía dependiendo de las bandas espectrales, reflejando diferencias en la capacidad de captura de detalles espaciales:

- **Bandas de 10 m:** Caracterizadas por una MTF alta, son ideales para capturar detalles finos, especialmente en aplicaciones como el monitoreo urbano y agrícola.
- **Bandas de 20 m:** Con una MTF moderada, logran un balance entre resolución espacial y área cubierta, siendo útiles para estudios de vegetación.
- **Bandas de 60 m:** Diseñadas para observaciones a gran escala, estas bandas priorizan patrones generales, lo que resulta en una MTF más baja pero adecuada para análisis globales.

Estudios recientes han demostrado que el diseño de Sentinel-2, que utiliza valores específicos de MTF adaptados a cada banda espectral, asegura una optimización de las aplicaciones según las necesidades de resolución espacial y espectral (Donike, 2022).

2.2.3.3 MTF y fusión de imágenes. En el proceso de fusión de imágenes de diferentes bandas y resoluciones, la *Modulation Transfer Function* (MTF) desempeña un papel esencial al balancear las frecuencias espaciales altas y bajas. Las bandas de mayor resolución contribuyen con detalles finos, mientras que las de menor resolución aportan información contextual, formando una representación integrada y coherente. La MTF garantiza que las propiedades espaciales y espectrales de cada banda se conserven durante la integración. Un análisis reciente demostró que métodos de fusión basados en MTF logran una alta fidelidad espacial ypectral al ajustar las diferencias inherentes en la resolución de sensores como Sentinel-2 y otros sistemas multiespectrales (Zhu & Su, 2024).

2.2.3.4 MTF y modelos de aprendizaje profundo. Los modelos de aprendizaje profundo, específicamente las redes neuronales convolucionales (CNNs), ofrecen una solución avanzada para compensar las limitaciones de la *Modulation Transfer Function* (MTF) en imágenes satelitales. Estos modelos permiten:

- Restaurar detalles perdidos en frecuencias espaciales altas, mejorando la nitidez de las imágenes.
- Combinar datos multiespectrales de manera coherente, integrando propiedades espaciales y espectrales.

Estas técnicas son fundamentales para superar las limitaciones ópticas en sensores como Sentinel-2. Por ejemplo, los modelos basados en CNN han sido diseñados para incorporar la degradación modelada por la MTF, logrando así una superresolución más precisa y una mejora general en la calidad de las imágenes, especialmente para aplicaciones de monitoreo terrestre y multiespectrales (Vasilescu et al., 2023).

2.2.4 Corregistro espacial

El corregistro espacial alinea geométricamente imágenes satelitales provenientes de sensores diferentes, como Sentinel-2 y NAIP, corrigiendo desplazamientos subpíxel

originados por diferencias geométricas. Este proceso es esencial en técnicas como la superresolución y la fusión de imágenes, garantizando coherencia espacial y evitando distorsiones que puedan afectar métricas espectrales como el NDVI o modelos predictivos en aplicaciones ambientales (Malczewska et al., 2023). A continuación, se describen las estrategias y métodos principales.

2.2.4.1 Estrategias. Dependiendo de la complejidad geométrica de los datos, se utilizan diferentes estrategias para el corregistro espacial, según Scheffler et al. (2017):

Enfoque global. Calcula un único vector de desplazamiento para toda la imagen, aplicando una transformación uniforme. Es eficiente en datos con variaciones geométricas mínimas, pero no adecuado para corregir distorsiones locales complejas.

Enfoque local. Divide la imagen en una cuadrícula densa, detectando y corrigiendo desplazamientos específicos en cada región. Es ideal para imágenes multitemporales o multisensor con desalineaciones por nubosidad, topografía u otros factores.

Fusión de técnicas. Métodos avanzados, como AROSICS, combinan correlación de fase y validación mediante algoritmos como RANSAC para filtrar errores y lograr precisión subpíxel, incluso en imágenes con alta heterogeneidad.

2.2.4.2 Métodos de alineación. Se han desarrollado diferentes algoritmos para abordar desafíos específicos del corregistro espacial, mejorando métricas clave como PSNR y SSIM en aplicaciones de fusión de imágenes y superresolución:

AROSICS. Software de código abierto basado en correlación de fase en la frecuencia espacial. Detecta y corrige desplazamientos subpíxel de manera automática, siendo robusto frente a variaciones atmosféricas y dinámicas de la superficie terrestre (Scheffler et al., 2017).

Framework MuRA. Utiliza redes neuronales como SuperPoint+SuperGlue para identificar correspondencias entre imágenes. Posteriormente, aplica transformaciones polinómicas derivadas de RANSAC, alcanzando precisión subpíxel en imágenes de alta resolución (Deshmukh et al., 2023).

ShiftNet. Integra el registro explícito de imágenes en la etapa de cálculo de pérdida, prediciendo desplazamientos subpíxel entre imágenes de superresolución (SR) y alta resolución (HR). Este enfoque mejora la calidad del corregistro en tareas de superresolución, optimizando la alineación espacial y la precisión espectral (Deudon et al., 2020).

2.2.5 Relación entre bandas de Sentinel-2

Las bandas espectrales de Sentinel-2 presentan resoluciones espaciales de 10, 20 y 60 metros, diseñadas para capturar diferentes niveles de detalle espacial y espectral. Mientras que las bandas de 10 m incluyen el visible (RGB) y el infrarrojo cercano (NIR), las de 20 m abarcan el infrarrojo de onda corta (SWIR) y bordes rojos. Estas diferencias permiten una caracterización detallada de la superficie terrestre, pero también plantean un desafío al integrar datos con resoluciones distintas.

Estudios recientes han demostrado la existencia de correlaciones significativas entre las bandas de 10 m y 20 m, que pueden explotarse para mejorar la resolución espacial de las bandas de menor detalle. Por ejemplo, la banda NIR (B8, 10 m) presenta fuertes correlaciones con las bandas SWIR (B11 y B12, 20 m), lo que facilita el uso de información de alta resolución de B8 para refinar las características espaciales de B11 y B12 (Lanaras et al., 2018).

2.2.5.1 Modelado de relaciones entre bandas. El proceso de integración entre bandas de distintas resoluciones se modela comúnmente mediante un enfoque de degradación:

$$I_{LR} = \delta(I_{HR}) + \epsilon \quad (2.1)$$

Donde:

- I_{LR} : Imagen de baja resolución generada.
- I_{HR} : Imagen de alta resolución tomada como referencia.
- δ : Transformación que incluye degradación espacial y ajustes espectrales.
- ϵ : Ruido asociado al sensor y al procesamiento.

Este modelo captura la relación entre bandas de diferente resolución y guía el desarrollo de técnicas para superresolución. En particular, las redes neuronales convolucionales (CNNs) han demostrado ser efectivas para aprender estas relaciones de manera no lineal, mejorando métricas como RMSE (Error Cuadrático Medio Relativo) y SAM (Ángulo Espectral Medio) en la superresolución de bandas de 20 m a 10 m.

2.2.5.2 Evaluación de modelos de superresolución. Los modelos de superresolución entrenados en conjuntos sintéticos (por ejemplo, degradando resoluciones de 40→20 m) han mostrado una alta generalización al aplicarse en datos reales (20→10 m). Esto asegura que las características espaciales y espectrales preservadas durante el entrenamiento se mantengan consistentes en casos reales. Por ejemplo, un modelo basado en CNN logró reducir en un 30 % el RMSE y mejorar un 15 % el SAM en bandas SWIR tras aplicar superresolución (Lanaras et al., 2018).

2.3 Superresolución de imágenes satelitales

La superresolución consiste en mejorar la resolución espacial de las imágenes mediante el uso de algoritmos avanzados que reconstruyen información perdida debido a limitaciones geométricas o de adquisición (He et al., 2023). Este enfoque es particularmente relevante en imágenes satelitales, donde las limitaciones tecnológicas imponen compromisos entre resolución espacial,pectral y temporal.

2.3.1 Propósito

La superresolución busca aumentar el nivel de detalle espacial de las imágenes mediante reconstrucciones precisas que aprovechan patrones existentes en los datos. Este proceso mejora la calidad visual y permite análisis más detallados en aplicaciones como teledetección, monitoreo ambiental y vigilancia.

2.3.1.1 Tareas. La superresolución es esencial en tareas como:

Recuperación espectral. Reconstrucción de imágenes hiperespectrales desde datos RGB o multiespectrales.

Colorización de imágenes. Generación de versiones en color de imágenes en escala de grises.

Compresión espectral. Recuperación de imágenes completas a partir de datos submuestreados, como en sistemas CASSI.

2.3.1.2 Retos actuales. Entre los principales desafíos se incluyen:

- Mejorar la generalización para diferentes sensores y condiciones de adquisición.
- Incrementar la robustez frente a ruido, degradación y condiciones ambientales adversas.
- Reducir la complejidad computacional para permitir implementaciones más eficientes.

2.3.2 Métodos tradicionales

Rohith y Kumar (2021) describen las bases de los métodos tradicionales para la superresolución, los cuales comenzaron utilizando técnicas matemáticas para extraer información faltante a partir de modelos simples y determinísticos. Estas técnicas incluyen interpolación, transformaciones de frecuencia y métodos iterativos.

2.3.2.1 Interpolación. La interpolación es uno de los métodos más básicos utilizados para mejorar la resolución espacial de imágenes. Consiste en ajustar los datos originales mediante funciones continuas y muestreo en intervalos más finos. Entre las técnicas más comunes destacan:

- **Interpolación de vecino más cercano:** replica píxeles adyacentes, siendo rápida pero con baja calidad visual.
- **Interpolación bilineal:** calcula promedios ponderados de los píxeles cercanos, mejorando la calidad pero generando suavizados excesivos.
- **Interpolación bicúbica:** utiliza valores de 16 píxeles adyacentes para un suavizado superior, siendo una de las más usadas debido a su equilibrio entre complejidad y calidad.

2.3.2.2 Transformaciones de frecuencia. Estas técnicas operan en el dominio de la frecuencia, realizando manipulaciones matemáticas para mejorar los detalles espaciales antes de volver al dominio espacial. Las principales herramientas incluyen:

- **Transformada de Fourier:** permite realizar detalles espaciales al eliminar aliasing espectral en imágenes.
- **Métodos basados en wavelets:** preservan detalles finos mediante transformaciones multiescala.

2.3.2.3 Enfoques iterativos. Los métodos iterativos, como la proyección sobre conjuntos convexos (POCS) y la retroproyección iterativa, optimizan la diferencia entre imágenes simuladas de baja resolución y las observadas. Aunque efectivos, presentan desafíos como la lentitud en la convergencia y la sensibilidad a parámetros iniciales.

2.3.2.4 Limitaciones de los métodos tradicionales. A pesar de su simplicidad, los métodos tradicionales tienen limitaciones significativas:

- Carecen de capacidad para modelar patrones contextuales y espectrales complejos.
- Son altamente sensibles a errores de registro y ruido en los datos.
- No consideran degradaciones específicas ni condiciones ambientales adversas, lo que restringe su aplicabilidad.

2.3.3 *Métodos basados en aprendizaje profundo*

El aprendizaje profundo (Deep Learning, DL) ha transformado la superresolución de imágenes al permitir el aprendizaje jerárquico de representaciones complejas directamente a partir de los datos. A diferencia de los métodos tradicionales, que enfrentan desafíos para modelar las relaciones entre imágenes de baja resolución (LR) y alta resolución (HR), los algoritmos de DL utilizan abstracciones de alto nivel para conectar ambos dominios de manera efectiva. Esto ha resultado en avances significativos en áreas como teledetección, vigilancia y análisis médico Rohith y Kumar (2021).

La superresolución basada en DL se clasifica en dos enfoques principales: la Superresolución de Imagen Única (SISR) y la Superresolución Multi-Imagen (MISR). Ambos buscan aumentar la resolución espacial, pero difieren en sus métodos, requisitos de entrada y aplicaciones.

2.3.3.1 Superresolución de Imagen Única (SISR). La Superresolución de Imagen Única (SISR) consiste en reconstruir una imagen de alta resolución (HR) a partir de una sola imagen de baja resolución (LR). Este enfoque es ampliamente utilizado debido a su simplicidad y versatilidad en aplicaciones como vigilancia, análisis médico y procesamiento de imágenes satelitales. La SISR ha evolucionado significativamente con la introducción del aprendizaje profundo, superando las limitaciones de los métodos tradicionales Anwar et al. (2020).

Redes Neuronales Convolucionales (CNN). Las CNN son fundamentales en la superresolución de imágenes, y su aplicación marcó un avance significativo en el campo. El modelo SRCNN Dong et al. (2014, 2015) introdujo una arquitectura básica con tres

capas convolucionales, demostrando la viabilidad del aprendizaje profundo para realizar superresolución. Posteriormente, se desarrollaron arquitecturas más avanzadas:

- **VDSR (Very Deep Super-Resolution):** Introducido por Kim et al. (2016), utiliza redes convolucionales profundas con aprendizaje residual para acelerar la convergencia y mejorar métricas como PSNR y SSIM.
- **EDSR (Enhanced Deep Super-Resolution):** Propuesto por Lim et al. (2017), elimina la normalización por lotes para reducir la complejidad computacional y mejorar la calidad visual.
- **FSRCNN y ESPCN:** Modelos como FSRCNN Dong et al. (2016) y ESPCN Shi et al. (2016) optimizan el uso de la memoria y mejoran la velocidad de procesamiento, especialmente para aplicaciones en tiempo real.

Redes Generativas Adversarias (GAN). Las GAN han revolucionado la superresolución al emplear un generador para producir imágenes superresolucionadas (HR) y un discriminador que evalúa su autenticidad comparándolas con imágenes reales. Modelos como SRGAN (Ledig et al., 2017) y ESRGAN (X. Wang et al., 2018) destacan por generar imágenes visualmente realistas. Sin embargo, estas redes pueden introducir artefactos visuales, lo que las hace más adecuadas para aplicaciones donde la calidad percibida es prioritaria sobre métricas cuantitativas (Creswell et al., 2018).

Redes basadas en Atención y GAN combinadas En aplicaciones más recientes, se han combinado técnicas basadas en atención y GAN para explotar sus fortalezas complementarias. Por ejemplo, modelos híbridos logran preservar detalles finos y evitar artefactos generados por GAN puras (J. Wang et al., 2021).

Autoencoders (AE). Los autoencoders son útiles en SISR debido a su capacidad para aprender representaciones latentes significativas y manejar datos multiespectrales. Por ejemplo, REDNet (Mao et al., 2016) emplea bloques residuales y conexiones de salto para preservar detalles esenciales, logrando reconstrucciones precisas en tareas como

el pan-sharpening. Estos modelos son particularmente efectivos en escenarios donde las relaciones entre las bandas espectrales requieren una modelización detallada.

Mecanismos de Atención. Los mecanismos de atención han mejorado significativamente las capacidades de las redes de superresolución al resaltar características importantes y dependencias de largo alcance. Redes como RCAN (Zhang et al., 2018), SwinIR (Liang et al., 2021), y otras arquitecturas de vanguardia (Cao et al., 2023; Choi et al., 2023; Hui et al., 2019; Kong et al., 2022; Liu et al., 2020; Muqeet et al., 2020; H. Wang et al., 2023) utilizan mapas de atención para capturar información contextual, asegurando la continuidad de detalles y texturas. Sin embargo, el uso de atención convencional puede aumentar significativamente la complejidad del modelo.

Mecanismos de Atención Ligeros La atención ligera permite obtener mejoras en rendimiento sin aumentar la complejidad computacional. Mapas de atención generados sin necesidad de parámetros adicionales (Choe & Shim, 2019; Haase & Amthor, 2020) han demostrado ser efectivos, maximizando la eficiencia en modelos con recursos limitados. Este enfoque es particularmente relevante para aplicaciones en tiempo real y sistemas con restricciones de hardware.

El Modelo Propuesto: SPAN SPAN (Red de Atención Sin Parámetros Rápida) combina simplicidad estructural y eficiencia computacional en la superresolución de imágenes. Utiliza Bloques de Atención Sin Parámetros Rápidos (SPAB) para extraer características de niveles progresivamente más altos mediante convoluciones con núcleos fijos y conexiones residuales. Las funciones de activación simétricas generan mapas de atención directamente desde las características extraídas, eliminando la necesidad de parámetros adicionales, lo que simplifica la red y acelera la inferencia sin sacrificar precisión (Wan et al., 2024b).

El diseño de SPAN resuelve limitaciones comunes en mecanismos de atención, preservando magnitudes y direcciones de las características clave y mitigando pérdidas de

información en regiones con detalles de alto nivel menos prominentes.

2.3.3.2 Superresolución Multi-Imagen (MISR). La MISR, también conocida como Multi-Frame Super-Resolution, utiliza múltiples imágenes de la misma escena para reconstruir una imagen HR. Este enfoque aprovecha la redundancia temporal y espacial para reducir ruido y compensar las limitaciones de una sola imagen (Deudon et al., 2020; Kawulok et al., 2021).

Técnicas de fusión. En MISR, las imágenes se fusionan tras la extracción de características. Métodos como HighRes-Net (Deudon et al., 2020) combinan imágenes temporalmente cercanas en el espacio de características para producir representaciones mejoradas, que luego se procesan con flujos de trabajo SISR.

Ventajas y desafíos. MISR ofrece ventajas como la mejora de la calidad espacial y temporal mediante datos redundantes. Sin embargo, enfrenta desafíos como la sincronización de múltiples capturas y las diferencias atmosféricas entre imágenes Salvetti et al. (2020).

2.3.3.3 Comparación entre SISR y MISR. Mientras que SISR es adecuada para escenarios con una única imagen disponible, MISR aprovecha múltiples imágenes de la misma escena para mejorar la robustez frente al ruido y los detalles. SISR es computacionalmente más sencilla, pero MISR ofrece mejores resultados al aprovechar información redundante. La elección entre ambos enfoques depende de las características del conjunto de datos y las necesidades específicas de la aplicación.

2.3.4 Protocolo de Wald

Tras analizar los métodos más relevantes para la superresolución, es esencial evaluar la calidad de las imágenes generadas. En este contexto, el Protocolo de Wald, una metodología ampliamente aceptada en el ámbito de la fusión multiespectral y pancromática (Wald et al., 1997), se establece como el estándar principal de evaluación. Este protocolo

proporciona un marco objetivo para comparar y validar diferentes enfoques en términos de coherencia espacial y espectral, fundamentándose en dos criterios principales:

2.3.4.1 Criterios del Protocolo de Wald.

- **Consistencia:** Las imágenes superresueltas deben parecerse a las imágenes originales cuando son degradadas a su resolución inicial, asegurando que la información original no se pierda. Por ejemplo, bandas de mayor resolución pueden degradarse y compararse con las de resolución inferior.
- **Síntesis:** Las imágenes generadas deben aproximarse a imágenes ideales de alta resolución, simulando los resultados obtenidos por sensores avanzados. Este criterio se valida con métricas como PSNR, SSIM, SAM y ERGAS.

2.3.4.2 Procedimientos de evaluación. El protocolo define dos enfoques principales para la validación:

- **Resolución reducida (RR):** Las imágenes originales se degradan para generar datos simulados de baja resolución, que sirven como entrada para los métodos de superresolución. Las imágenes originales actúan como referencia para calcular métricas como RMSE, SSIM y SAM.
- **Resolución completa (FR):** Utiliza directamente las imágenes originales como entrada y salida, validando la calidad mediante métricas sin referencia como QNR, D_λ (distorsión espectral) y D_s (distorsión espacial).

Ambos enfoques son complementarios, midiendo respectivamente la coherencia a diferentes escalas y la calidad a la resolución más alta.

2.3.4.3 Métricas principales. Para analizar la calidad de las imágenes generadas, se emplean métricas ampliamente aceptadas en el procesamiento de imágenes satelitales:

- **PSNR (Peak Signal-to-Noise Ratio):** Evalúa la relación señal-ruido para medir la fidelidad de la reconstrucción.

- **SSIM (Structural Similarity Index):** Analiza la similitud estructural entre las imágenes originales y las generadas.
- **SAM (Spectral Angle Mapper):** Mide la precisión espectral calculando el ángulo entre vectores espectrales.
- **ERGAS (Erreur Relative Globale Adimensionnelle de Synthèse):** Resume la distorsión global, integrando coherencia espacial y espectral.

Optimización. El filtrado en el dominio de Fourier puede integrarse para equilibrar frecuencias y optimizar la calidad final, asegurando estándares de precisión tanto en pruebas controladas como en aplicaciones prácticas.

Este protocolo establece un marco objetivo para comparar y validar métodos de superresolución en términos de calidad y desempeño.

3. OBJETIVOS Y METODOLOGÍA DE TRABAJO

3.1 Objetivo General

Desarrollar un modelo de superresolución para mejorar la resolución espacial de imágenes Sentinel-2 de 10 y 20 metros a 2.5 metros, utilizando redes neuronales convolucionales y técnicas de fusión multiespectral.

3.2 Objetivos Específicos

- 1. Explorar e implementar técnicas avanzadas de fusión de imágenes:** Mejorar la resolución espacial y espectral de datos multiespectrales de Sentinel-2, preservando las características originales.
- 2. Desarrollar un modelo de fusión de imágenes basado en aprendizaje profundo:** Crear un modelo que combine datos multiespectrales para generar imágenes de mayor resolución espacial manteniendo la integridad espectral.
- 3. Aplicar el Protocolo de Wald para la validación:** Utilizar el Protocolo de Wald para asegurar que las imágenes super-resueltas mantengan consistencia con las imágenes originales en términos de propiedades espaciales y espirales.
- 4. Evaluar el rendimiento del modelo propuesto:** Analizar cómo la superresolución mejora la detección de inundaciones mediante índices de humedad y métricas de rendimiento.

3.3 Metodología de trabajo

Este trabajo emplea un enfoque de fusión en dos etapas: **Fusión x2** y **Fusión x4**, para mejorar la resolución espacial de imágenes satelitales Sentinel-2. Estas etapas permiten una transición progresiva desde resoluciones originales de *10 y 20 metros* a una resolución final de *2.5 metros*, manteniendo la integridad espectral y espacial de los datos.

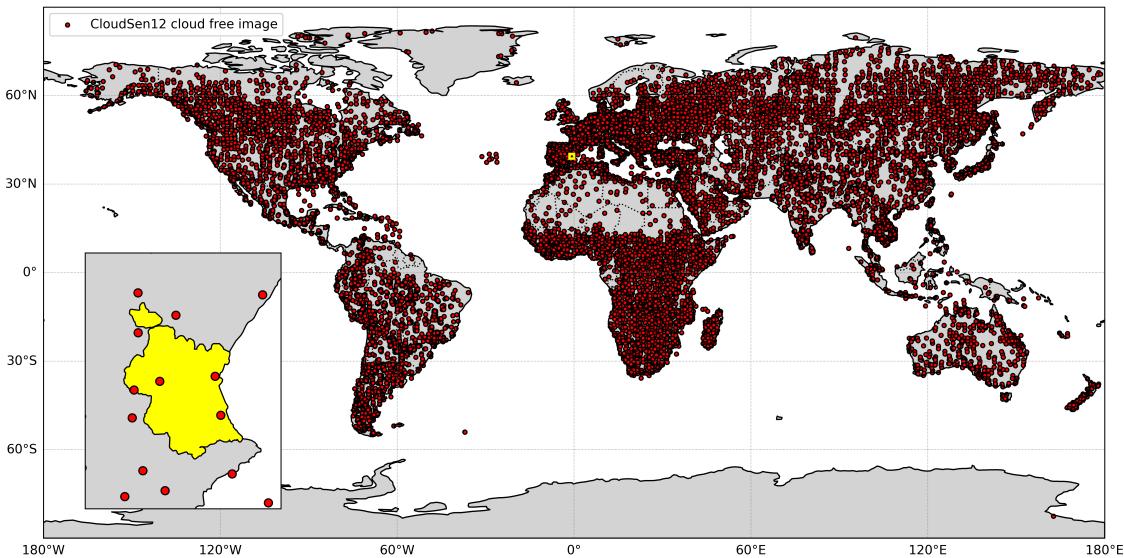
Las bandas multiespectrales, con resoluciones originales diferentes, se procesan de manera diferenciada en cada etapa de fusión. Este enfoque asegura que la metodología capture y preserve tanto los detalles espaciales como las características espectrales del entorno observado.

3.3.1 Obtención y procesamiento de los datos

Para el desarrollo de este proyecto, se utiliza el conjunto de datos *CloudSEN12* de Aybar et al. (2024), un dataset global diseñado para la comprensión semántica de nubes y sombras de nubes en imágenes Sentinel-2. Este conjunto es accesible a través de Hugging Face, siguiendo los pasos indicados en este [Colab público](#) e incluye un total de 49,400 imágenes en formato *Level-1C* (TOA), de las cuales se seleccionan 10,440 imágenes libres de nubes.

Figura 6.

Área de aplicación: ciudad de Valencia



Nota. Se muestran imágenes sin nubes del proyecto CloudSEN12 en zonas de alta heterogeneidad espectral que enriquecen el modelo, destacando la ciudad de Valencia como área de aplicación para evaluar la mejora en la detección de inundaciones mediante super-resolución ante su reciente vulnerabilidad a eventos extremos como un DANA (octubre, 2024). Elaboración propia.

3.3.1.1 Preprocesamiento de los datos. Las imágenes, que contienen bandas de 10 y 20 metros, provienen del conjunto de datos *CloudSEN12*, donde fueron descargadas con una resolución uniforme de 10 metros, utilizando el método de remuestreo de vecino más cercano (*nearest neighbor*) desde Google Earth Engine (*GEE*). Este proceso asegura la consistencia espacial necesaria para la integración multiespectral.

3.3.1.2 Imágenes sin nubes. El uso de imágenes libres de nubes garantiza que el modelo analice exclusivamente las características reales de la superficie terrestre. Esto elimina la necesidad de reconstrucciones adicionales, comunes en áreas con nubes, y mejora la fiabilidad del análisis multiespectral, como señala Li et al. (2023).

3.3.1.3 Heterogeneidad espectral. La selección de regiones con alta diversidad espectral, evitando áreas homogéneas como océanos y zonas polares, enriquece el conjunto de datos. Este enfoque fortalece la capacidad del modelo para capturar una mayor variedad de patrones espectrales, siguiendo recomendaciones de estudios como Jiao et al. (2023).

3.3.1.4 Selección del área de aplicación. La ciudad de Valencia se selecciona como área de aplicación para evaluar cómo la superresolución mejora la detección de inundaciones. Esta región, vulnerable a eventos extremos como un DANA (octubre, 2024), representa un escenario adecuado para analizar el impacto de esta metodología en tareas prácticas de monitoreo y gestión de desastres.

3.3.2 Metodología general para la fusión

En ambas etapas de fusión se siguen procedimientos similares en cuanto a la preparación de datos, el filtrado en el dominio de Fourier y el uso de mecanismos de atención mediante SPAN (*Self-parameterized Attention Network*). A continuación, se describen los elementos clave de la metodología aplicada en el proceso.

3.3.2.1 Preparación de datos. La preparación de datos garantiza la alineación y normalización de las bandas de diferentes resoluciones, facilitando su integración como entrada en el modelo. La normalización se realiza dividiendo los valores de los píxeles entre *10,000*, ajustando el rango de datos y mejorando la estabilidad durante el entrenamiento.

Durante el entrenamiento, para fomentar la generalización y prevenir el *overfitting*, se aplica un muestreo aleatorio que selecciona porciones distintas de las imágenes en cada época. Los lotes de datos se reorganizan en cada iteración, evitando que el modelo dependa de un orden específico y fortaleciendo su capacidad de generalización.

En la fase de prueba, se extrae una porción central de cada imagen y los datos se presentan en un orden fijo, garantizando evaluaciones consistentes y reproducibles, libres de la variabilidad introducida por el muestreo aleatorio.

Bandas de referencia. Las bandas de 10 metros, correspondientes a las regiones espectrales del espectro visible (RGB) y NIR (bandas 2, 3, 4 y 8), sirven como referencia para enriquecer la información espacial y espectral en ambas etapas de fusión. Según el requerimiento de cada etapa, estas bandas se usan en su resolución original o se escalan para integrarse con otras bandas.

Bandas ajustadas. Las bandas de 20 metros (bandas 5, 6, 7, 8A, 11 y 12) aportan información clave del borde rojo y el infrarrojo de onda corta. Su resolución viene ajustada a 10 metros mediante interpolación bilineal con antialiasing, asegurando una integración precisa en cada etapa de fusión.

3.3.2.2 Filtrado en el dominio de Fourier. Para separar los componentes de alta y baja frecuencia y optimizar el procesamiento de la imagen, se emplea un filtro en el dominio de Fourier. Se utiliza un filtro Butterworth de orden 6, ajustado según el factor de escala específico de cada etapa, para definir un radio de frecuencia en la máscara de paso bajo. Este procedimiento preserva los detalles esenciales mientras se atenúan las frecuencias no deseadas, garantizando una representación equilibrada entre la estructura global y los detalles finos.

3.3.2.3 Arquitectura del modelo y SPAN. El modelo de super-resolución se basa en una arquitectura de red neuronal convolucional que incorpora bloques de atención auto-parametrizados SPAB (*Self-parameterized Attention Blocks*). Estos bloques permiten al modelo enfocarse en áreas de la imagen que requieren mayor detalle, mejorando la calidad de la imagen fusionada.

Dentro de cada SPAB, se utilizan módulos convolucionales personalizados denominados Conv3XC. Estos módulos están diseñados para capturar características complejas de la imagen mediante una serie de operaciones convolucionales y conexiones residuales. La estructura del bloque Conv3XC es la siguiente:

1. **Primera convolución:** Convierte las características de entrada a un espacio inter-

medio de mayor dimensionalidad.

2. **Segunda convolución:** Aplica una convolución con un kernel de tamaño 3x3 para capturar patrones locales y detalles finos.
3. **Tercera convolución:** Reduce la dimensionalidad de las características al número deseado de canales de salida.
4. **Conexión residual:** Una convolución 1x1 que conecta directamente la entrada con la salida del bloque, facilitando el flujo de gradientes y mejorando la capacidad de aprendizaje.

Estas operaciones permiten al modelo aprender representaciones profundas y detalladas de las imágenes, capturando tanto la información global como los detalles locales. Las activaciones utilizadas incluyen funciones como *SiLU* y *Leaky ReLU*, que contribuyen a una mejor representación de las características no lineales de la imagen.

La Figura 7 ilustra la arquitectura general del modelo y el flujo interno dentro de un bloque SPAB.

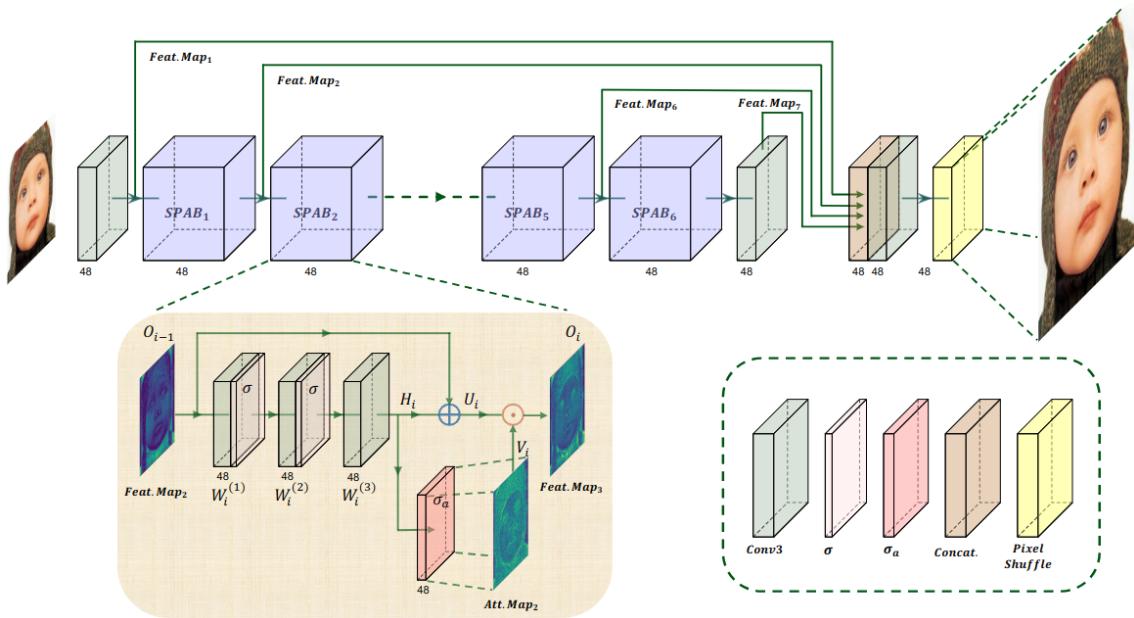


Figura 7. Arquitectura del modelo SPAN, que muestra la secuencia de bloques de atención auto-parametrizados (SPAB) y los módulos convolucionales personalizados (Conv3XC) utilizados para extraer características a través de convoluciones y conexiones residuales. La sección ampliada ilustra el flujo interno dentro de un bloque SPAB, donde la atención se genera de forma libre de parámetros mediante funciones de activación simétricas, maximizando la representación de características relevantes sin añadir parámetros adicionales. Este esquema destaca cómo el modelo integra mapas de características y mapas de atención para mejorar la resolución de la imagen de entrada. Recuperado de Wan et al. (2024a).

3.3.3 Fusión x2: mejora de resolución intermedia

La primera etapa, Fusión x2, se enfoca en duplicar la resolución espacial, permitiendo que las bandas de 20 metros alcancen una resolución de 10 metros. Este proceso sienta las bases para una posterior mejora en la etapa de Fusión x4, unificando la mayoría de las bandas en una resolución común y enriqueciendo la información disponible.

3.3.3.1 Ajustes en la preparación de datos.

Bandas de 10 metros. Estas bandas se redimensionan a una resolución espacial de 20 metros mediante interpolación bilineal con antialiasing. Este ajuste iguala su resolución con las bandas de 20 metros, permitiendo una integración coherente en el modelo y asegurando que los detalles característicos de las bandas de alta resolución se conserven.

Bandas de 20 metros. Inicialmente, se reducen a una resolución de 40 metros utilizando interpolación bilineal con antialiasing. Posteriormente, se reescalan nuevamente a 20 metros aplicando el mismo método. Este proceso de reducción y ampliación facilita que las bandas mantengan coherencia espacial con las bandas de 10 metros ajustadas, preparando los datos para su procesamiento conjunto en el modelo.

3.3.3.2 Componentes.

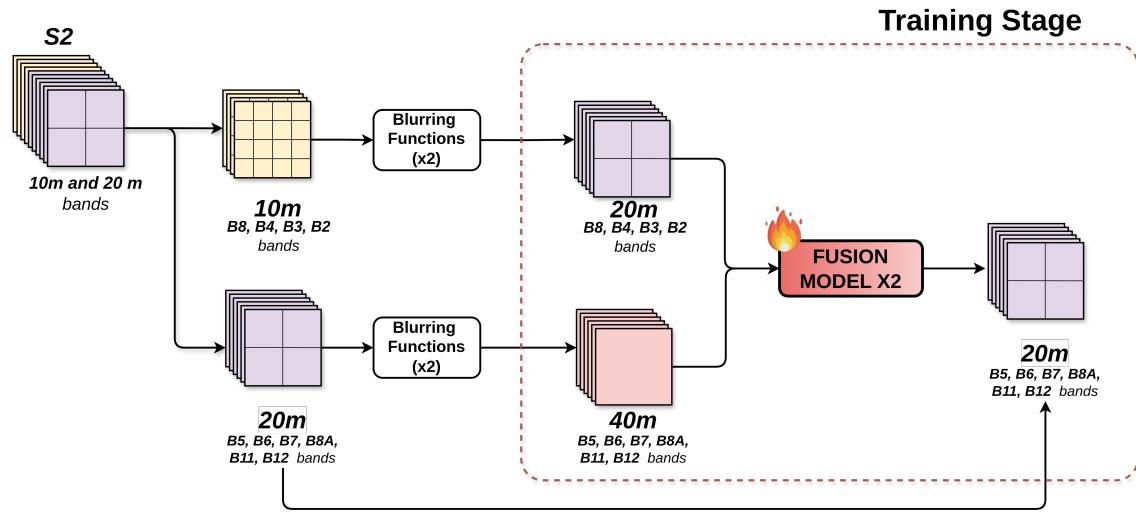
Entrada. Concatenación de las bandas de 10 metros escaladas a 20 metros y las bandas de 20 metros ajustadas. Esta combinación proporciona al modelo una fuente consistente de detalles espaciales y espectrales.

Objetivo. Las bandas originales de 20 metros sirven como referencia para el aprendizaje supervisado, permitiendo al modelo reconstruir los detalles espectrales.

Salida. Una representación fusionada que combina las características espectrales de las bandas de 20 metros con los detalles espaciales de las bandas de referencia, mejorando la resolución y el detalle de la imagen.

Figura 8.

Proceso de Fusión x2 en Superresolución



Nota. Esquema que muestra la preparación de datos y el modelo de Fusión x2, donde las bandas de entrada son ajustadas en resolución para generar una salida unificada de 20 metros, combinando detalles espaciales y espectrales. Elaboración propia.

3.3.3.3 Filtrado en el dominio de Fourier. Se aplica un filtro Butterworth de orden 6 con un factor de escala correspondiente a la duplicación de la resolución. El proceso incluye:

- **Transformación FFT:** Las imágenes interpoladas y las generadas por el modelo se transforman al dominio de Fourier mediante la Transformada Rápida de Fourier.
- **Aplicación de máscaras de frecuencia:** Se aplica una máscara de paso bajo a la imagen interpolada y una máscara de paso alto complementaria a la imagen de superresolución, integrando las frecuencias bajas y altas de manera equilibrada.
- **Transformación inversa FFT:** Se realiza la transformación inversa para regresar al dominio espacial, obteniendo una versión filtrada que sirve como salida optimizada del modelo.

Este filtrado asegura un balance adecuado entre coherencia espacial y detalle, optimizando el rendimiento en la fase de super-resolución.

3.3.3.4 Configuración del modelo. El modelo se configura con parámetros específicos para la Fusión x2:

- **Número de bloques SPAB:** 10 bloques para lograr un equilibrio entre rendimiento y complejidad computacional.
- **Canales de características:** 48 canales para capturar detalles finos en las características de la imagen.
- **Factor de escala:** 2, correspondiente a la duplicación de la resolución espacial.
- **Entradas y salidas:** 10 canales de entrada (bandas ajustadas) y 6 canales de salida (bandas fusionadas).

Se entrenaron diferentes configuraciones, variando su complejidad y capacidad para evaluar su desempeño. Las configuraciones utilizadas son las siguientes:

```
1      cnn_lightweight = {  
2          "in_channels": 10,  
3          "out_channels": 6,  
4          "feature_channels": 24,  
5          "upscale": 1,  
6          "bias": True,  
7          "train_mode": True,  
8          "num_blocks": 6,  
9      }  
10  
11      cnn_small = {  
12          "in_channels": 10,  
13          "out_channels": 6,  
14          "feature_channels": 48,  
15          "upscale": 1,
```

```
16         "bias": True,
17
18         "train_mode": True,
19
20         "num_blocks": 16,
21
22     }
23
24
25     cnn_medium = {
26
27         "in_channels": 10,
28
29         "out_channels": 6,
30
31         "feature_channels": 72,
32
33         "upscale": 1,
34
35         "bias": True,
36
37         "train_mode": True,
38
39         "num_blocks": 20
40
41     }
42
43
44     cnn_expanded = {
45
46         "in_channels": 10,
47
48         "out_channels": 6,
49
50         "feature_channels": 96,
51
52         "upscale": 1,
53
54         "bias": True,
55
56         "train_mode": True,
57
58         "num_blocks": 24,
59
60     }
61
62 }
```

En estas configuraciones, se varían los parámetros como el número de canales de características (`feature_channels`) y el número de bloques SPAB (`num_blocks`). El parámetro `upscale` se establece en 1, ya que en esta implementación el modelo opera en la misma resolución de entrada y salida, enfocándose en mejorar la calidad y detalle de la imagen a través del procesamiento interno del modelo.

El objetivo de utilizar diferentes tamaños de modelo es evaluar cómo la capacidad del modelo afecta al rendimiento en la tarea de super-resolución. Modelos más pequeños,

como el *lightweight*, tienen menos parámetros y requieren menos recursos computacionales, pero pueden tener una capacidad limitada para capturar detalles finos. Por otro lado, modelos más grandes, como el *expanded*, tienen mayor capacidad para aprender representaciones complejas, pero a costa de un mayor tiempo de entrenamiento y requerimientos de hardware.

3.3.3.5 Función de pérdida y entrenamiento. Para el entrenamiento del modelo, se utiliza la función de pérdida L1, que mide la diferencia absoluta media entre la salida del modelo y las bandas originales de 20 metros. Esta función es adecuada para preservar detalles y minimizar artefactos en la imagen resultante, siendo especialmente efectiva en tareas de super-resolución donde la precisión en los valores de los píxeles es crucial.

El modelo se entrena utilizando un optimizador AdamW con una tasa de aprendizaje inicial de 1×10^{-4} y un decaimiento de peso de 1×10^{-2} , lo que ayuda a prevenir el sobreajuste y a mejorar la generalización del modelo. Además, se emplea un programador de tasa de aprendizaje cíclico (CyclicLR) para ajustar dinámicamente la tasa de aprendizaje durante el entrenamiento, favoreciendo una convergencia más rápida y estable.

```
1     optimizer = torch.optim.AdamW(
2             self.parameters(),
3             lr=self.learning_rate,
4             weight_decay=1e-2)
5
6     scheduler = torch.optim.lr_scheduler.CyclicLR(
7             optimizer,
8             base_lr=1e-5,
9             max_lr=1e-3,
10            step_size_up=1000)
```

El entrenamiento se realiza con un tamaño de lote de 4 y durante 50 épocas. Se aplica una estrategia de parada temprana (*early stopping*) basada en la pérdida de validación para evitar el sobreentrenamiento. Estas configuraciones permiten al modelo aprender de manera eficiente y efectiva, logrando una reducción significativa en la pérdida y mejorando

la calidad de las imágenes de super-resolución generadas.

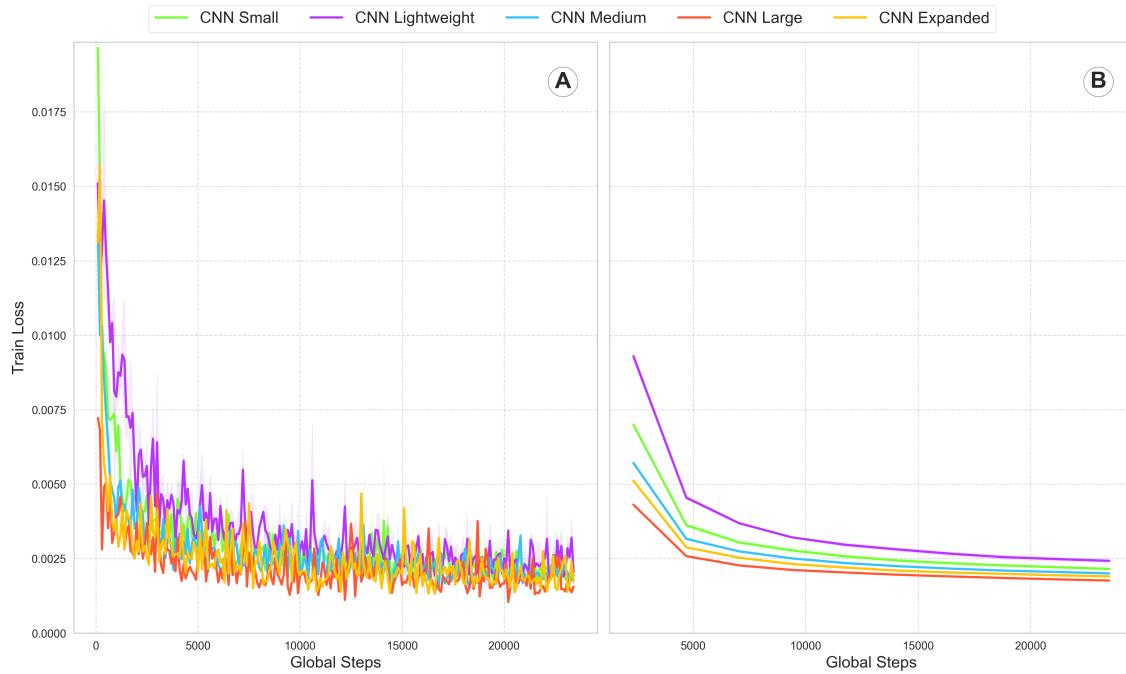


Figura 9. Comparación de la pérdida de entrenamiento para distintas configuraciones del modelo de fusión en super-resolución. En el panel A (izquierda), se observa la pérdida en función de los pasos globales, mostrando la estabilidad de cada configuración a lo largo del entrenamiento. En el panel B (derecha), la pérdida se presenta por épocas, donde las arquitecturas más complejas (Medium, Expanded, Large) convergen más rápidamente y logran menores valores de pérdida, lo que sugiere una mayor capacidad para capturar detalles en la tarea de super-resolución.

3.3.4 Fusión x4: mejora de resolución a nivel fino

La segunda etapa, Fusión x4, se enfoca en cuadruplicar la resolución espacial, obteniendo imágenes de 10 metros a partir de bandas originalmente de 40 metros. Esta fase aprovecha los resultados de la Fusión x2 para alcanzar una mayor fidelidad y detalle en la imagen final.

3.3.4.1 Ajustes en la preparación de datos.

Bandas de 10 metros. Se utilizan sin cambios, proporcionando detalles espaciales críticos que enriquecen la entrada del modelo.

Bandas de 20 metros. Se someten al siguiente proceso:

- **Reducción a 40 metros:** Se reducen a una resolución de 40 metros mediante interpolación bilineal con antialiasing.
- **Ampliación a 10 metros:** Posteriormente, se reescalan a una resolución de 10 metros utilizando el mismo método, permitiendo su integración con las bandas de 10 metros originales.

Bandas de 40 metros. Las bandas ajustadas a 40 metros se interpolan directamente a 10 metros mediante interpolación bilineal con antialiasing, proporcionando una representación adicional que enriquece la información espectral.

3.3.4.2 Definición de entrada, objetivo y salida.

- **Entrada:** Concatenación de las bandas de 40 metros interpoladas a 10 metros y las bandas de 10 metros originales, resultando en 10 canales de entrada.
- **Objetivo:** La salida generada por el modelo de Fusión x2 aplicado a las bandas de 20 metros, proporcionando una referencia refinada para el aprendizaje.
- **Salida:** Una imagen de super-resolución que combina efectivamente las características espectrales y espaciales, alcanzando una resolución de 10 metros con detalles refinados.

3.3.4.3 Filtrado en el dominio de Fourier. Se utiliza un filtro Butterworth de orden 6 ajustado para reflejar el factor de escala de 4. El proceso sigue los mismos pasos descritos anteriormente, adaptándose al nuevo factor de escala.

3.3.4.4 Configuración del modelo. Para la Fusión x4, se entrenan diferentes configuraciones del modelo CNNSR con el objetivo de evaluar su rendimiento y determinar la arquitectura más adecuada para la tarea. Estas configuraciones varían en términos de complejidad y capacidad, ajustando parámetros clave como el número de bloques SPAB y los canales de características.

Para la Fusión x4, se realizan ajustes específicos en el modelo:

- **Número de bloques SPAB:** 20 bloques para capturar detalles más finos y profundizar el aprendizaje.
- **Canales de características:** 72 canales para manejar la complejidad adicional de la tarea.
- **Factor de escala:** 1 (la salida es de la misma resolución que la entrada, pero con mayor detalle).
- **Entradas y salidas:** 10 canales de entrada y 6 canales de salida.

Se entrenaron diferentes configuraciones, variando su complejidad y capacidad para evaluar su desempeño. Las configuraciones utilizadas son las siguientes:

```
1      cnn_lightweight = {  
2          "in_channels": 10,  
3          "out_channels": 6,  
4          "feature_channels": 24,  
5          "upscale": 1,  
6          "bias": True,  
7          "train_mode": True,  
8          "num_blocks": 6,  
9      }  
10  
11      cnn_small = {  
12          "in_channels": 10,  
13          "out_channels": 6,  
14          "feature_channels": 48,  
15          "upscale": 1,  
16          "bias": True,  
17          "train_mode": True,  
18          "num_blocks": 16,  
19      }  
20  
21      cnn_medium = {
```

```
22         "in_channels": 10,
23
24         "out_channels": 6,
25
26         "feature_channels": 72,
27
28         "upscale": 1,
29
30         "bias": True,
31
32         "train_mode": True,
33
34         "num_blocks": 20,
35
36     }
37
38
39     cnn_expanded = {
40
41         "in_channels": 10,
42
43         "out_channels": 6,
44
45         "feature_channels": 96,
46
47         "upscale": 1,
48
49         "bias": True,
50
51         "train_mode": True,
52
53         "num_blocks": 24,
54
55     }
56
57
58     cnn_large = {
59
60         "in_channels": 10,
61
62         "out_channels": 6,
63
64         "feature_channels": 150,
65
66         "upscale": 1,
67
68         "bias": True,
69
70         "train_mode": True,
71
72         "num_blocks": 36,
73
74     }
75
76
77 }
```

Estas configuraciones permiten explorar cómo la complejidad del modelo afecta el rendimiento en la tarea de Fusión x4, desde modelos más ligeros hasta arquitecturas más profundas y con mayor capacidad de representación. Al evaluar estas variantes, es posi-

ble identificar un equilibrio óptimo entre el rendimiento y los recursos computacionales requeridos.

3.3.4.5 Integración con el modelo de Fusión x2. El modelo de Fusión x2 se utiliza para generar el objetivo que servirá como referencia en el entrenamiento de Fusión x4. Se cargan los pesos del modelo de Fusión x2 y se congela su entrenamiento:

```

1      # Cargar el modelo de Fusión x2
2
3      weights = torch.load(fusion_model, map_location="cpu")
4
5      fusionx2_model = utils.load_model(model_name="cnn",
6      model_size="medium")
7
8      fusionx2_model.load_state_dict(weights)
9
10     fusionx2_model.eval()
11
12     for param in fusionx2_model.parameters():
13
14         param.requires_grad = False

```

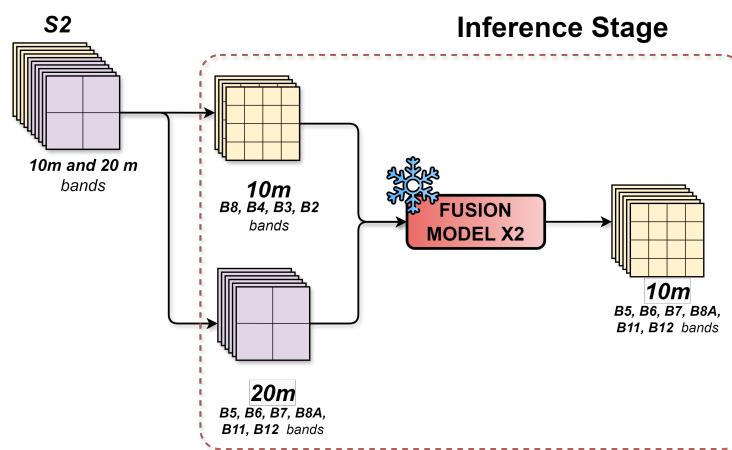


Figura 10. Diagrama del proceso de Fusión x2 en super-resolución. Las bandas de entrada a 20 metros y las de mayor resolución, preprocesadas con funciones de desenfoque, se integran en el modelo de Fusión x2. Este modelo produce una salida de 10 metros de resolución, combinando la información espacial de las bandas de alta resolución con la riqueza espectral de las de menor resolución. Este paso intermedio sirve para unificar las bandas en una resolución común y optimizar la información para la etapa de Fusión x4.

3.3.4.6 Implementación del filtrado con *Hard Constraint*. Se utiliza un *Hard Constraint* basado en una capa convolucional con pesos predefinidos que implementa el

filtro Butterworth de orden 6. Esto asegura que los componentes de baja frecuencia de la imagen interpolada se mantengan durante el procesamiento:

```

1      self.hard_constraint = CNNHardConstraint(
2          filter_method="butterworth",
3          filter_hyperparameters={"order": 6},
4          scale_factor=2,
5          in_channels=6,
6          out_channels=[0, 1, 2, 3, 4, 5],
7          device=device,
8      )
9

```

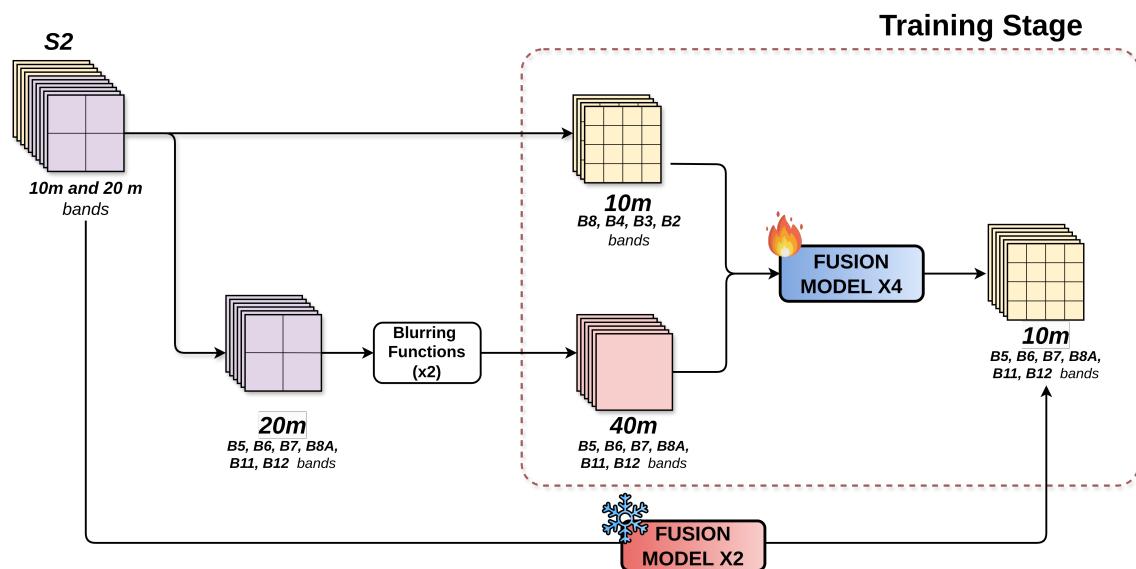


Figura 11. Diagrama del proceso de Fusión x4 en super-resolución. Se muestra la preparación de datos, en la que se aplican funciones de desenfoque (*Blurring Functions*) para reducir la resolución inicial de las bandas. Las bandas de 20 metros y las de mayor resolución (preprocesadas por el modelo de Fusión x2) son combinadas mediante el modelo de Fusión x4, que genera una salida final de 10 metros de resolución. Este modelo permite aumentar la fidelidad espacial de las bandas, integrando información de diferentes resoluciones para obtener una imagen más detallada y precisa.

3.3.4.7 Función de pérdida y entrenamiento. Para el entrenamiento del modelo de Fusión x4, se utiliza la *superloss*, una combinación de funciones de pérdida que incluye L1, LPIPS y CLIP. Esta función permite al modelo aprender no solo diferencias pixel a

pixel, sino también características perceptuales y semánticas de alto nivel, mejorando la calidad visual y la coherencia del contenido en la imagen resultante.

Los componentes de la *superloss* son:

- **L1 Loss** (L_{L1}): Calcula la diferencia absoluta media entre los píxeles de la imagen generada y la imagen objetivo, enfocándose en la precisión de bajo nivel.
- **LPIPS Loss** (L_{LPIPS}): Evalúa la similitud perceptual entre las imágenes utilizando características de redes neuronales profundas, capturando la calidad visual desde una perspectiva humana.
- **CLIP Loss** (L_{CLIP}): Utiliza el modelo CLIP para comparar características semánticas de alto nivel, asegurando que la imagen generada mantenga coherencia en términos de contenido.

La función de pérdida total se define como:

$$\text{SuperLoss} = \lambda_1 \cdot L_{L1} + \lambda_2 \cdot L_{LPIPS} + \lambda_3 \cdot L_{CLIP}$$

Donde los pesos $\lambda_1 = 17.7474$, $\lambda_2 = 0.8778$ y $\lambda_3 = 2.4049$ se establecen para equilibrar la contribución de cada término.

El modelo se entrena utilizando PyTorch Lightning, con el mismo optimizador AdamW y programador de tasa de aprendizaje cíclico (CyclicLR) descritos en la sección de Fusión x2 (ver Sección 3.3.3.5). Esto asegura una convergencia efectiva y un desempeño óptimo del modelo. Se utiliza una tasa de aprendizaje inicial de 1×10^{-4} y un tamaño de lote de 4, entrenando el modelo durante 50 épocas. Al igual que en Fusión x2, se implementa una estrategia de parada temprana basada en la pérdida de validación para evitar el sobreentrenamiento.

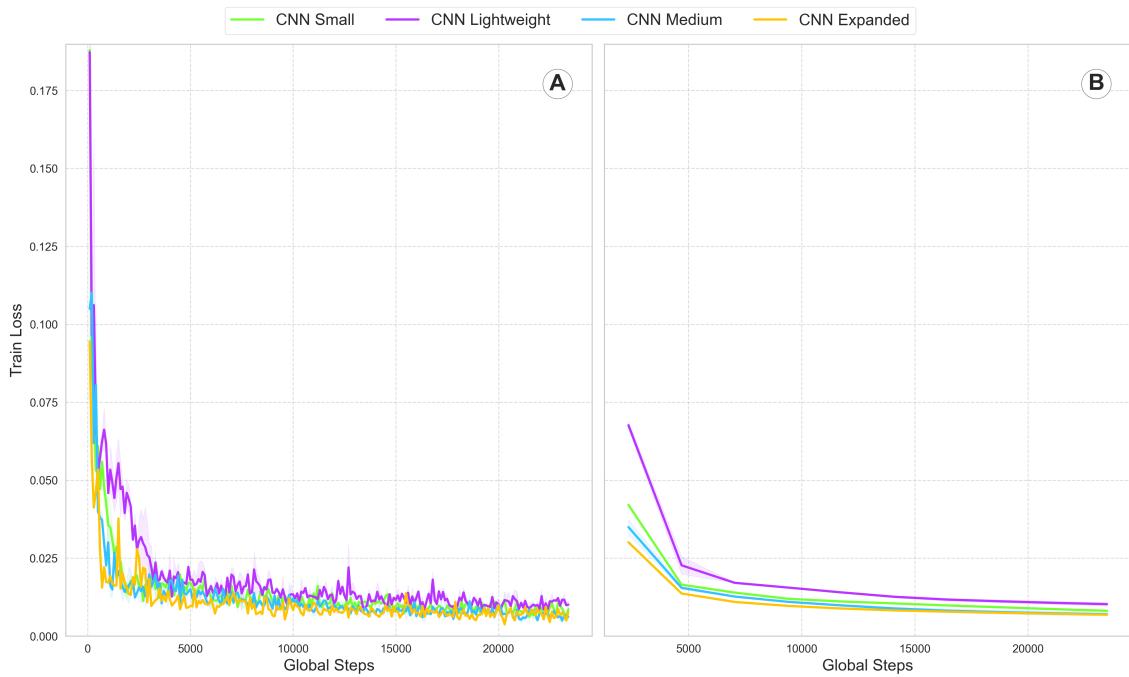


Figura 12. Evolución de la pérdida de entrenamiento durante el entrenamiento del modelo de Fusión x4 para diferentes configuraciones de CNN. En el panel A (izquierda), la pérdida se representa en función de los pasos globales, mostrando la estabilidad y comportamiento de cada configuración a medida que avanza el entrenamiento. En el panel B (derecha), se observa la pérdida suavizada, destacando la convergencia de cada arquitectura hacia menores valores de pérdida. Las arquitecturas más complejas, como CNN Expanded, logran menores pérdidas finales, indicando una mayor capacidad para capturar detalles en la super-resolución.

3.3.4.8 Análisis de la pérdida de testeo. Además de evaluar la pérdida durante el entrenamiento, es fundamental analizar el desempeño del modelo en el conjunto de testeo para validar su capacidad de generalización y asegurar que no se esté sobreajustando a los datos de entrenamiento. La Figura ?? presenta una comparación de la pérdida de testeo para las diferentes configuraciones del modelo CNN en las etapas de Fusión x2 y Fusión x4.

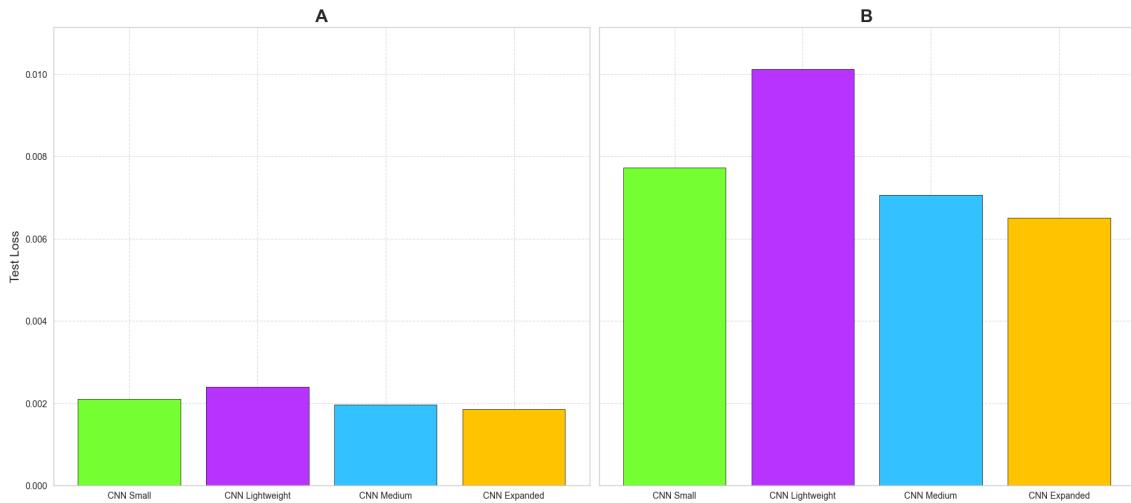


Figura 13. Comparación de la pérdida de testeo para distintas configuraciones de CNN en las etapas de Fusión x2 y Fusión x4. En el panel A (izquierda), se observa la pérdida de testeo para la Fusión x2, con resultados relativamente homogéneos entre arquitecturas, pero con una ligera ventaja en las configuraciones de mayor capacidad. En el panel B (derecha), correspondiente a la Fusión x4, se evidencia una mayor variabilidad, donde las arquitecturas más complejas, como CNN Expanded, logran menores pérdidas, indicando un rendimiento superior en la resolución final de 10 metros.

En la Figura 13, se aprecia que, especialmente en la etapa de Fusión x4, las arquitecturas con mayor capacidad tienden a lograr menores valores de pérdida en el conjunto de testeo, sugiriendo una mejor representación y captura de detalles. Sin embargo, los modelos de mayor tamaño también demandan más recursos computacionales y tiempo de entrenamiento, por lo que es necesario equilibrar el rendimiento y la eficiencia según las necesidades y limitaciones específicas de la aplicación.

3.3.4.9 Optimización de la inferencia mediante combinación de convoluciones. Durante la etapa de inferencia en Fusión x2, es esencial que el modelo CNNSR sea eficiente, especialmente en entornos con limitaciones de recursos. Para lograr una inferencia rápida sin comprometer la calidad, se ha implementado una optimización que reduce la complejidad computacional al combinar múltiples convoluciones en una sola operación.

Combinación de convoluciones en Conv3XC. El modelo utiliza bloques convolucionales Conv3XC, que durante el entrenamiento aplican tres convoluciones para aprender representaciones complejas de las imágenes:

1. **Primera convolución** (conv[0]): Convierte las características de entrada a un espacio intermedio.
2. **Segunda convolución** (conv[1]): Aplica una convolución 3x3 para capturar patrones locales.
3. **Tercera convolución** (conv[2]): Reduce la dimensionalidad a los canales de salida deseados.

Además, una conexión residual (sk) conecta la entrada con la salida del bloque mediante una convolución 1x1.

Optimización en inferencia. Para acelerar la inferencia, se combinan las tres convoluciones y la conexión residual en una sola operación mediante:

- **Precomputación de pesos y sesgos combinados:** Antes de la inferencia, se combinan los pesos y sesgos de las capas convolucionales en una única convolución.
- **Uso de una sola convolución en inferencia:** La operación combinada permite reducir el tiempo de procesamiento sin modificar la salida.

Esta optimización se realiza con la función `update_params()` dentro de la clase Conv3XC:

```
1      class Conv3XC(nn.Module):  
2          def __init__(...):  
3              # Inicialización de capas y variables  
4          def update_params(self):  
5              # Combina las convoluciones en una operación ú  
nica  
6          def forward(self, x: torch.Tensor) -> torch.Tensor:  
7              if self.train_mode:  
8                  # Durante entrenamiento, usa las  
convoluciones originales  
9              out = self.conv(x) + self.sk(x)
```

```
10         else:
11             # En inferencia, usa la convolución
12             combinada
13             self.update_params()
14             out = self.eval_conv(x)
15
16
17     return out
```

Para activar esta optimización:

```
1     # Configurar el modelo para inferencia
2
3     model.train_mode = False
4
5     model.eval()
6
7     with torch.no_grad():
8
9         output = model(input_data)
```

Beneficios de la optimización.

- **Eficiencia:** Reduce el tiempo de inferencia y uso de recursos.
- **Escalabilidad:** Permite procesar imágenes de mayor resolución sin aumentar significativamente el tiempo.
- **Calidad mantenida:** La salida es idéntica a la obtenida sin la optimización.

Esta técnica es ideal para aplicaciones que requieren procesamiento rápido de grandes volúmenes de datos en tiempo real o cercano al tiempo real.

Implementación práctica. Durante la inferencia en Fusión x2, el modelo utiliza la convolución combinada:

```
1     # Preparar el modelo para la inferencia
2
3     fusionx2_model.train_mode = False
4
5     fusionx2_model.eval()
6
7     # Procesar las imágenes de entrada
```

```
5         with torch.no_grad():
6             output = fusionx2_model(input_images)
7
```

Esta optimización reduce el tiempo de procesamiento sin afectar la calidad de las imágenes generadas.

4. IDENTIFICACIÓN DE REQUISITOS

4.1 Trabajo Previo Realizado

En este capítulo se describe el trabajo previo realizado para guiar el desarrollo del software, considerando la identificación del problema y el contexto habitual de uso o funcionamiento de la aplicación.

4.1.1 Identificación del Problema

Uno de los principales retos en el análisis de imágenes satelitales es la falta de sensores de alta resolución de acceso abierto. Esto limita la capacidad de realizar estudios detallados y precisos en diversas aplicaciones geoespaciales.

4.1.2 Contexto de Uso

El software desarrollado está diseñado para funcionar en entornos de inferencia, donde se espera que procese imágenes satelitales y aplique técnicas de superresolución en tiempo real o en procesos de análisis. La identificación de requisitos ha sido realizada en conjunto con expertos en la materia para asegurar su aplicabilidad y eficiencia en escenarios reales de teledetección.

5. DESCRIPCIÓN DE LA HERRAMIENTA SOFTWARE DESARROLLADA

5.1 Detalles del Proceso de Desarrollo

Este capítulo detalla el proceso de desarrollo del software, incluyendo las fases, los hitos alcanzados y los diagramas explicativos de la arquitectura y funcionamiento del modelo de superresolución.

5.1.1 *Fases del Desarrollo*

El desarrollo de la herramienta se dividió en fases clave, cada una con objetivos específicos:

- **Fase de Planificación:** Identificación de requerimientos y objetivos.
- **Fase de Diseño:** Estructuración del modelo SR (Superresolución) y diseño de arquitectura.
- **Fase de Implementación:** Desarrollo de los módulos y entrenamiento del modelo de superresolución.
- **Fase de Pruebas:** Validación y ajuste de parámetros para garantizar precisión y rendimiento.

5.1.2 *Diagramas de Flujo y Arquitectura*

Se presentan diagramas de flujo y arquitecturas que ilustran el funcionamiento del programa, explicando cada etapa del procesamiento y superresolución de imágenes. Estos diagramas proporcionan una visión clara de la estructura interna y la secuencia de operaciones.

5.1.3 *Capturas de Pantalla*

Para facilitar la comprensión del programa, se incluyen capturas de pantalla que muestran la interfaz de usuario y el proceso de superresolución, permitiendo al lector vi-

sualizar el flujo y resultado del programa.

6. EVALUACIÓN

6.1 Usabilidad de la Herramienta

La evaluación de la herramienta incluye una valoración de su usabilidad, centrada en la accesibilidad y eficiencia de la interfaz y su adaptabilidad para distintos usuarios en el ámbito de la teledetección.

6.2 Aplicabilidad al Problema Propuesto

Se analiza la aplicabilidad de la herramienta al problema de mejorar la resolución de imágenes satelitales. Este análisis se realiza con la colaboración de usuarios expertos, quienes evalúan la capacidad del modelo para ofrecer soluciones viables y de alta resolución en el contexto de aplicaciones geoespaciales avanzadas.

REFERENCES

- Ajmar, A., Boccardo, P., Broglia, M., Kucera, J., Giulio-Tonolo, F., & Wania, A. (2017). Response to flood events: The role of satellite-based emergency mapping and the experience of the Copernicus emergency management service. *Flood damage survey and assessment: New insights from research and practice*, 211-228.
- Alparone, L., Aiazzi, B., Baronti, S., & Garzelli, A. (2015). *Remote sensing image fusion*. Crc Press.
- Amani, M., Mohseni, F., Layegh, N. F., Nazari, M. E., Fatolazadeh, F., Salehi, A., Ahmadi, S. A., Ebrahimi, H., Ghorbanian, A., Jin, S., et al. (2022). Remote sensing systems for ocean: A review (Part 2: Active systems). *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 1421-1453.
- Anwar, S., Khan, S., & Barnes, N. (2020). A deep journey into super-resolution: A survey. *ACM Computing Surveys (CSUR)*, 53(3), 1-34.
- Aybar, C., Bautista, L., Montero, D., Contreras, J., Ayala, D., Prudencio, F., Loja, J., Ysuhuaylas, L., Herrera, F., Gonzales, K., et al. (2024). CloudSEN12+: The largest dataset of expert-labeled pixels for cloud and cloud shadow detection in Sentinel-2. *Data in Brief*, 56, 110852.
- Caballero, I., Ruiz, J., & Navarro, G. (2019). Sentinel-2 satellites provide near-real time evaluation of catastrophic floods in the west mediterranean. *Water*, 11(12), 2499.
- Cao, J., Wang, Q., Xian, Y., Li, Y., Ni, B., Pi, Z., Zhang, K., Zhang, Y., Timofte, R., & Gool, L. V. (2023). Ciaosr: Continuous implicit attention-in-attention network for arbitrary-scale image super-resolution. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1796-1807.
- Choe, J., & Shim, H. (2019). Attention-based dropout layer for weakly supervised object localization. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2219-2228.

- Choi, H., Lee, J., & Yang, J. (2023). N-gram in swin transformers for efficient lightweight image super-resolution. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2071-2081.
- Cooley, S. W., Smith, L. C., Stepan, L., & Mascaro, J. (2017). Tracking dynamic northern surface water changes with high-frequency planet CubeSat imagery. *Remote Sensing*, 9(12), 1306.
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE signal processing magazine*, 35(1), 53-65.
- Denis, G., Claverie, A., Pasco, X., Darnis, J.-P., de Maupeou, B., Lafaye, M., & Morel, E. (2017). Towards disruptions in Earth observation? New Earth Observation systems and markets evolution: Possible scenarios and impacts. *Acta Astronautica*, 137, 415-433.
- Denis, G., de Boissezon, H., Hosford, S., Pasco, X., Montfort, B., & Ranera, F. (2016). The evolution of Earth Observation satellites in Europe and its impact on the performance of emergency response services. *Acta Astronautica*, 127, 619-633.
- Deshmukh, R., Roros, C. J., Kashyap, A., & Kak, A. C. (2023). An aligned multi-temporal multi-resolution satellite image dataset for change detection research. *arXiv preprint arXiv:2302.12301*.
- Deudon, M., Kalaitzis, A., Goytom, I., Arefin, M. R., Lin, Z., Sankaran, K., Michalski, V., Kahou, S. E., Cornebise, J., & Bengio, Y. (2020). Highres-net: Recursive fusion for multi-frame super-resolution of satellite imagery. *arXiv preprint arXiv:2002.06460*.
- Dong, C., Loy, C. C., He, K., & Tang, X. (2014). Learning a deep convolutional network for image super-resolution. *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part IV* 13, 184-199.

- Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2), 295-307.
- Dong, C., Loy, C. C., & Tang, X. (2016). Accelerating the super-resolution convolutional neural network. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part II* 14, 391-407.
- Donike, S. (2022). Deep Super-Resolution of Sentinel-2 Time Series. *eplus.uni-salzburg.at*.
<https://eplus.uni-salzburg.at/obvusbhs/content/titleinfo/8498881/full.pdf>
- ED Chaves, M., CA Picoli, M., & D. Sanches, I. (2020). Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review. *Remote Sensing*, 12(18), 3062.
- ESA. (2019). Sentinel-2 User Handbook, ESA Standard Document, Issue I, Rev. 2.
- European Space Agency. (2024, septiembre). *Sentinel-2C se une a la familia Copernicus en órbita* [Consulted on 2024-11-09]. https://www.esa.int/Newsroom/Press_Releases/Sentinel-2C_se_une_a_la_familia_Copernicus_en_orbita
- Gargiulo, M., Mazza, A., Gaetano, R., Ruello, G., & Scarpa, G. (2019). Fast super-resolution of 20 m Sentinel-2 bands using convolutional neural networks. *Remote Sensing*, 11(22), 2635.
- Ginting, D., Setiawan, K., Anggraini, N., Suardana, A., Nandika, M., Ulfa, A., Aziz, K., & Dewanti, R. (2024). Comparison between top and bottom of atmosphere Sentinel-2 image for mangrove mapping in Balikpapan Bay, East Kalimantan. *BIO Web of Conferences*, 89, 07003.
- Griffiths, P., Nendel, C., & Hostert, P. (2019). Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote sensing of environment*, 220, 135-151.
- Haase, D., & Amthor, M. (2020). Rethinking depthwise separable convolutions: How intra-kernel correlations lead to improved mobilenets. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14600-14609.

- He, J., Yuan, Q., Li, J., Xiao, Y., Liu, D., Shen, H., & Zhang, L. (2023). Spectral super-resolution meets deep learning: Achievements and challenges. *Information Fusion*, 97, 101812.
- Hubble, A. L. (2021). Single-Image Super-Resolution: Towards the Enhancement of Sentinel-2 Satellite Imagery. *adamleonardhubble.github.io*. <https://adamleonardhubble.github.io/portfolio/assets/document/Thesis.pdf>
- Hui, Z., Gao, X., Yang, Y., & Wang, X. (2019). Lightweight image super-resolution with information multi-distillation network. *Proceedings of the 27th ACM International Conference on Multimedia*, 2024-2032.
- Jiao, B., Su, Y., Li, Q., Manara, V., & Wild, M. (2023). An integrated and homogenized global surface solar radiation dataset and its reconstruction based on a convolutional neural network approach. *Earth System Science Data*, 15(10), 4519-4535.
- Karagiannopoulou, A., Tsertou, A., Tsimiklis, G., & Amditis, A. (2022). Data fusion in earth observation and the role of citizen as a sensor: A scoping review of applications, methods and future trends. *Remote Sensing*, 14(5), 1263.
- Kaushik, S. K., Mishra, V. N., Punia, M., Diwate, P., Sivasankar, T., & Soni, A. K. (2021). Crop health assessment using Sentinel-1 SAR time series data in a part of central India. *Remote Sensing in Earth Systems Sciences*, 4(4), 217-234.
- Kawulok, M., Tarasiewicz, T., Nalepa, J., Tyrna, D., & Kostrzewska, D. (2021). Deep learning for multiple-image super-resolution of Sentinel-2 data. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 3885-3888.
- Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1646-1654.
- Kong, F., Li, M., Liu, S., Liu, D., He, J., Bai, Y., Chen, F., & Fu, L. (2022). Residual local feature network for efficient super-resolution. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 766-776.

- Lanaras, C., Bioucas-Dias, J., Galliani, S., Baltsavias, E., & Schindler, K. (2018). Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 305-319.
- Landsat, U. (2016). Landsat 8 captured the BlueCutFire in California on Aug. 17, 2016. View the image: <https://t.co/5i8gk4q8bX>. <https://x.com/usgslandsat/status/773939936755982336>
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4681-4690.
- Li, C., Liu, X., & Li, S. (2023). Transformer meets GAN: Cloud-free multispectral image reconstruction via multi-sensor data fusion in satellite images. *IEEE Transactions on Geoscience and Remote Sensing*.
- Liang, J., Cao, J., Sun, G., Zhang, K., Gool, L. V., & Timofte, R. (2021). Swinir: Image restoration using swin transformer. *IEEE/CVF International Conference on Computer Vision Workshops, ICCVW 2021, Montreal, BC, Canada, October 11-17, 2021*, 1833-1844.
- Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). *Remote sensing and image interpretation*. John Wiley & Sons.
- Lim, B., Son, S., Kim, H., Nah, S., & Mu Lee, K. (2017). Enhanced deep residual networks for single image super-resolution. *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 136-144.
- Lisboa, F., Brotas, V., & Santos, F. D. (2024). Earth Observation—An Essential Tool towards Effective Aquatic Ecosystems' Management under a Climate in Change. *Remote Sensing*, 16(14), 2597.
- Liu, J., Tang, J., & Wu, G. (2020). Residual feature distillation network for lightweight image super-resolution. *Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16*, 41-55.

- Malczewska, A., Malczewski, J., & Hejmanowska, B. (2023). Challenges in Preparing Datasets for Super-Resolution on the Example of SENTINEL-2 and Planet Scope Images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, 91-98.
- Mao, X., Shen, C., & Yang, Y.-B. (2016). Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. *Advances in neural information processing systems*, 29.
- Matgen, P., Martinis, S., Wagner, W., Freeman, V., Zeil, P., McCormick, N., et al. (2020). Feasibility assessment of an automated, global, satellite-based flood-monitoring product for the Copernicus Emergency Management Service. *Luxembourg: Publications Office of the European Union*.
- Michel, J., Vinasco-Salinas, J., Inglada, J., & Hagolle, O. (2022). Sen2ven μ s, a dataset for the training of sentinel-2 super-resolution algorithms. *Data*, 7(7), 96.
- Muheet, A., Hwang, J., Yang, S., Kang, J. H., Kim, Y., & Bae, S.-H. (2020). Multi-attention based ultra lightweight image super-resolution. *Computer Vision - ECCV 2020 Workshops - Glasgow, UK, August 23-28, 2020, Proceedings, Part III*, 103-118.
- Navarro, M. A., & Sánchez, J. (2020). Sharp estimates of semistable radial solutions of k-Hessian equations. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 150(4), 2083-2115. <https://doi.org/10.1017/prm.2019.14>
- Pampanoni, V., Fascetti, F., Cenci, L., Laneve, G., Santella, C., & Boccia, V. (2024). Analysing the Relationship between Spatial Resolution, Sharpness and Signal-to-Noise Ratio of Very High Resolution Satellite Imagery Using an Automatic Edge Method. *Remote Sensing*, 16(6), 1041.
- Prendergast, S., Payne, C., Lifson, M., Haughwout, C., Tieppo, M., Figueiredo, P., Guerrra, A., Costa, A., Magalhães, H., Hormigo, T., et al. (2022). AEROS: Oceanographic Hyperspectral Imaging and Argos-Tracking CubeSat. *arXiv preprint arXiv:2211.05124*.

- Rast, M., Ananasso, C., Bach, H., Ben-Dor, E., Chabrilat, S., Colombo, R., Del Bello, U., Féret, J.-B., Giardino, C., Green, R. O., et al. (2019). Copernicus hyperspectral imaging mission for the environment: mission requirements document.
- Rohith, G., & Kumar, L. S. (2021). Paradigm shifts in super-resolution techniques for remote sensing applications. *The Visual Computer*, 37(7), 1965-2008.
- Salgueiro Romero, L., Marcello, J., & Vilaplana, V. (2020). Super-resolution of sentinel-2 imagery using generative adversarial networks. *Remote Sensing*, 12(15), 2424.
- Salvetti, F., Mazzia, V., Khaliq, A., & Chiaberge, M. (2020). Multi-image super resolution of remotely sensed images using residual attention deep neural networks. *Remote Sensing*, 12(14), 2207.
- Santilli, G., Cappelletti, C., Battistini, S., & Vendittozzi, C. (2016). Disaster Management of remote areas by constellation of CubeSats. *Proceedings of the 67th Astronautical Congress (IAC), Guadalajara, Mexico*, 26-30.
- Scheffler, D., Hollstein, A., Diedrich, H., Segl, K., & Hostert, P. (2017). AROSICS: An automated and robust open-source image co-registration software for multi-sensor satellite data. *Remote sensing*, 9(7), 676.
- Schiavon, E., Taramelli, A., Tornato, A., & Pierangeli, F. (2021). Monitoring environmental and climate goals for European agriculture: User perspectives on the optimization of the Copernicus evolution offer. *Journal of Environmental Management*, 296, 113121.
- Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., Rueckert, D., & Wang, Z. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1874-1883.
- Valenzuela, A., Reinke, K., & Jones, S. (2022). A new metric for the assessment of spatial resolution in satellite imagers. *International Journal of Applied Earth Observation and Geoinformation*, 114, 103051.

- Vasilescu, V., Datcu, M., & Faur, D. (2023). A CNN-based Sentinel-2 image super-resolution method using multiobjective training. *IEEE Transactions on Geoscience and Remote Sensing*, 61, 1-14.
- Wald, L., Ranchin, T., & Mangolini, M. (1997). Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogrammetric engineering and remote sensing*, 63(6), 691-699.
- Wan, C., Yu, H., Li, Z., Chen, Y., Zou, Y., Liu, Y., Yin, X., & Zuo, K. (2024a). Swift Parameter-free Attention Network for Efficient Super-Resolution. <https://arxiv.org/abs/2311.12770>
- Wan, C., Yu, H., Li, Z., Chen, Y., Zou, Y., Liu, Y., Yin, X., & Zuo, K. (2024b). Swift parameter-free attention network for efficient super-resolution. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6246-6256.
- Wang, H., Chen, X., Ni, B., Liu, Y., & Liu, J. (2023). Omni aggregation networks for lightweight image super-resolution. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 22378-22387.
- Wang, J., Gao, K., Zhang, Z., Ni, C., Hu, Z., Chen, D., & Wu, Q. (2021). Multisensor remote sensing imagery super-resolution with conditional GAN. *Journal of Remote Sensing*.
- Wang, Q., Shi, W., Li, Z., & Atkinson, P. M. (2016). Fusion of Sentinel-2 images. *Remote sensing of environment*, 187, 241-252.
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Change Loy, C. (2018). Esrgan: Enhanced super-resolution generative adversarial networks. *Proceedings of the European conference on computer vision (ECCV) workshops*, 0–0.
- Wulder, M. A., Roy, D. P., Radeloff, V. C., Loveland, T. R., Anderson, M. C., Johnson, D. M., Healey, S., Zhu, Z., Scambos, T. A., Pahlevan, N., et al. (2022). Fifty years of Landsat science and impacts. *Remote Sensing of Environment*, 280, 113195.
- Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., & Fu, Y. (2018). Image super-resolution using very deep residual channel attention networks. *Computer Vision - ECCV*

2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part VII, 294-310.

Zhu, J., Zhang, & Su. (2024). Comparison. *Proceedings of the International Conference on Agro-Geoinformatics*. <https://ieeexplore.ieee.org/abstract/document/10660709/>