



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA DE POSTGRADO Y EDUCACIÓN CONTINUA
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

**MEJORA DE LA RESOLUCIÓN ESPACIAL DE DATOS SATELITALES
SOBRE ALDEHÍDOS UTILIZANDO UNA RED NEURONAL
CONVOLUCIONAL**

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIA DE DATOS

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN COMPUTACIÓN

NICOLÁS FABIÁN CABELLO RÍOS

PROFESORA GUÍA:
NANCY HITSCHFELD KAHLER

PROFESORA CO-GUÍA:
LAURA GALLARDO KLENNER

COMISIÓN:
AIDAN HOGAN

Este trabajo ha sido parcialmente financiado por Centro de Ciencia del Clima y la Resiliencia (CR)², Chile, y parcialmente apoyado por la infraestructura de supercómputo del NLHPC (CCSS210001).

SANTIAGO DE CHILE

2024

Resumen

Las observaciones satelitales de variables atmosféricas son fundamentales para el monitoreo de la calidad del aire y para abordar la influencia humana sobre el clima. Es por esto que la precisión y resolución espacial son de gran relevancia ya que aportan a la toma de decisiones informada de políticas públicas. Dadas las mejoras en las tecnologías de percepción remota durante los últimos años, se ha generado una discontinuidad entre los datos obtenidos con las tecnologías nuevas y las antiguas, la cual es necesario homogeneizar.

Este trabajo busca aumentar la resolución espacial de datos de formaldehído (HCHO) obtenidos por el espectrómetro del instrumento satelital Ozone Monitoring Instrument (OMI) utilizando Redes Neuronales Convolucionales entrenadas con datos de Tropospheric Monitoring Instrument (TROPOMI). Las dos arquitecturas de Redes Neuronales utilizadas corresponden a la *Super Resolution Convolutional Neural Network* (SRCNN) y la *Super Resolution Generative Adversarial Network* (SRGAN). La metodología utilizada se basa en la propiedad de las Redes Convolucionales de ser entrenadas con pequeñas sub-imágenes de los datos utilizados, pero tener la capacidad de ser aplicadas a datos sin restricción de tamaño mientras se mantenga la consistencia de las resoluciones utilizadas durante el proceso de entrenamiento. Adicionalmente se obtuvo dos conjuntos de datos para entrenar y evaluar las redes, considerando un traslape mayor entre sub-imágenes de entrenamiento y otro conjunto con un solapamiento menor, donde mientras más solapamiento se dispone de una mayor cantidad de datos.

Los resultados sugieren las redes SRCNN y SRGAN entregan un mejor rendimiento que el caso de comparación base de interpolación bicúbica, siendo SRGAN la que mejor desempeño obtuvo de forma consistente. En cuanto a los conjuntos de datos, se obtuvo que mientras mayor sea la disponibilidad de datos el rendimiento de la red entrenada aumenta pese a que se espera que disponga de una menor heterogeneidad. La aplicación de Redes Adversarias logra obtener buenos resultados en este problema, por lo cual invita a realizar mejoras a las arquitecturas. Este trabajo logra contribuir a la mejora del monitoreo para Chile y América del Sur.

Para mi abuelo.

Agradecimientos

Primero agradezco a mis padres Marcela y Rodrigo por haberme dado la oportunidad de estudiar y fomentar una buena ética de trabajo. A mis hermanos Vicente y Matías por estar presentes en el día a día por más conflictos que podamos tener. A mis perros Lukas y Nerón que son los pilares de la alegría en la casa. También a mi abuelo Segundo por la paciencia y cariño que siempre tuvo, el apoyo que me brindó desde niño apoyándome en actividades que me gustaban y por enseñarme lo que es la verdadera amabilidad. A mi difunta abuela Amanda, que con todo su esfuerzo fue el pilar del surgimiento de la familia, y a su hermana Berta, por apoyarnos en cada parte del proceso.

Agradezco a mis profesoras guías Nancy y Laura por el apoyo durante el trabajo de tesis y por brindarme la oportunidad de desarrollar un tema de mi interés en cambio climático. A Santiago por el tiempo dedicado a explicarme detalles técnicos para este trabajo, sin los cuales no habría podido obtener los resultados actuales.

A mis amigos del colegio: David, Bastián, Cristóbal, Gabriel y Christian que han estado conmigo durante tantos años. Por todos los cafés, las juntas y los salidas que hemos compartido, valoro cada recuerdo con ustedes. Su apoyo y compañía ha sido fundamental durante todos este tiempo y espero poder mantenerlo por muchos años más. A los muchachos de ATR por las reuniones y risas que me han alegrado una y otra vez, han sido una gran compañía. A los amigos que hice a lo largo de la carrera en la universidad, fueron un pilar para poder llevar el día a día, desde los almuerzos, las conversaciones y las risas.

Agradezco también a la rama de natación por brindar un espacio de compañerismo y distensión que fue de gran ayuda para sobrellevar los momentos de mayor estrés. Los buenos momentos vividos con el equipo me han dejado grandes recuerdos.

Quiero agradecer de forma abierta a cada persona que me brindó ayuda o un saludo en la universidad, haya sido funcionaria, de la cafetería, de la biblioteca, del departamento de computación o del área de deportes. Todos aportaron un grano de arena en este largo camino. Una pequeña mención a mi psicóloga Nathalia, quien definitivamente me ha ayudado a llevar de forma sana todos los desafíos que se fueron presentando.

Finalmente reconocerme a mí por mi trabajo realizado, por perseverar y salir adelante aunque se presentaran dificultades. Si bien muchas veces estuve perdido, con esfuerzo, paciencia y confiando en mi mismo logré resolver los problemas. Ahora puedo decir que todo el esfuerzo valió la pena.

Tabla de Contenido

1. Introducción	1
1.1. Definición del Problema	2
1.2. Objetivos	2
1.2.1. Objetivo General	2
1.2.2. Objetivos Específicos	3
1.3. Metodología	3
1.4. Estructura de Tesis	4
2. Marco Teórico	5
2.1. Columna de Densidad Troposférica de HCHO	5
2.2. Transformada de Coseno Discreta para llenado de datos	7
2.3. Super-Resolución de imagen única	8
2.4. Redes Neuronales Convolucionales	8
2.5. Redes Generativas Adversarias	9
2.6. Redes Neuronales Convolucionales para Super Resolution	9
2.7. Métricas de evaluación	10
2.7.1. Proporción Máxima de Señal a Ruido	10
2.7.2. Similaridad estructural	10
2.7.3. Error porcentual absoluto medio	11
3. Análisis y Diseño de Datos	13
3.1. Datos satelitales	13

3.2. Datos de HCHO	14
3.3. Preprocesamiento	16
3.3.1. Regrillado	17
3.3.2. Llenado de datos	18
3.3.3. Generación de ventanas	19
3.3.4. Reemplazo de valores negativos	20
4. Implementación	22
4.1. Arquitecturas	22
4.2. Entrenamiento	24
5. Resultados y análisis	27
5.1. Conjuntos con valores negativos	28
5.1.1. Primer conjunto de datos	28
5.1.2. Segundo conjunto de datos	31
5.2. Conjuntos sin valores negativos	34
5.2.1. Primer conjunto de datos sin valores negativos	34
5.2.2. Segundo conjunto de datos sin valores negativos	36
5.3. Análisis de resultados	38
5.4. Reconstrucción de datos OMI	40
6. Conclusión y trabajo futuro	44
Bibliografía	49
Anexos	50
Anexo A	50

Índice de Ilustraciones

2.1.	Descripción de la principal producción y remoción de formaldehído atmosférico. Fuente Dias et al., 2020.	6
3.1.	Sección eficaz de absorción de HCHO a 298 °K. Fuente de datos: Meller y Moortgat, 2000.	14
3.2.	Rango espectral para OMI y TROPOMI. Fuente: Veeffkind, 2012.	15
3.3.	(a) Concentración de HCHO en grilla irregular en escala logarítmica. (b) Valor de calidad de datos. Datos del mismo archivo de origen ejemplificando el recorrido orbital.	16
3.4.	Delimitación de zona de estudio.	16
3.5.	Distribución de datos.	17
3.6.	Diagrama de regridado de dato a grilla regular. En verde se encuentran las celdas que cumplen la condición del llenado al 50 %, y en rojo la que no cumplen la condición.	18
3.7.	Diagrama de obtención de error RMSE para distintas combinaciones de size y s	19
3.8.	Selección de parámetros para función de llenado. (a) Error de reconstrucción para valores de s . (b) Error de reconstrucción para valores de size	19
3.9.	Diagrama explicativo de la generación de ventanas de datos.	20
3.10.	Diagrama explicativo de reemplazo de datos negativos.	21
4.1.	Arquitectura de red SRCNN junto al proceso de interpolación	23
4.2.	Arquitectura de la red adversaria. a) Red Generadora. b) Red Discriminadora.	23
4.3.	Diagrama de entrenamiento de redes. a) Entrenamiento de red convolucional. b) Entrenamiento de red adversaria.	25

5.1. Distribución acumulada de modelos para el conjunto de prueba del primer conjunto de datos. Mayor PSNR y SSIM es mejor, mientras que menor MAPE es mejor.	29
5.2. Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados para el primer conjunto de datos.	30
5.3. Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados con negativos filtrados para el primer conjunto de datos.	30
5.4. Gráfico de predicciones para modelos para el primer conjunto de datos considerando valores negativos. a) Dato a predecir en baja resolución. b) Reconstrucción de modelo bicúbico. c) Reconstrucción de modelo SRCNN. d) Reconstrucción de modelo SRGAN. e) Dato real de alta resolución TROPOMI.	31
5.5. Gráfico de medición de error para el primer conjunto de datos considerando valores negativos. a) Dato real de alta resolución TROPOMI. b) Reconstrucción de modelo SRCNN, c) Reconstrucción de modelo SRGAN. d) Error de predicción de SRCNN. e) Error de predicción de SRGAN.	31
5.6. Distribución acumulada de modelos para el conjunto de prueba del segundo conjunto de datos. Mayor PSNR y SSIM es mejor, mientras que menor MAPE es mejor.	32
5.7. Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI para el segundo conjunto de datos.	32
5.8. Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados con negativos filtrados para el segundo conjunto de datos.	33
5.9. Gráfico de predicciones para modelos para el segundo conjunto de datos considerando valores negativos. a) Dato a predecir en baja resolución. b) Reconstrucción de modelo bicúbico. c) Reconstrucción de modelo SRCNN. d) Reconstrucción de modelo SRGAN. e) Dato real de alta resolución TROPOMI.	33
5.10. Gráfico de medición de error para el segundo conjunto de datos considerando valores negativos. a) Dato real de alta resolución TROPOMI. b) Reconstrucción de modelo SRCNN. c) Reconstrucción de modelo SRGAN. d) Error de predicción de SRCNN. e) Error de predicción de SRGAN.	34
5.11. Distribución acumulada de modelos para el conjunto de prueba del primer conjunto de datos con valores negativos preprocesados. Mayor PSNR y SSIM es mejor, mientras que menor MAPE es mejor.	35
5.12. Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados para el primer conjunto de datos con valores negativos preprocesados.	36

5.13. Gráfico de predicciones para modelos para el primer conjunto de datos con valores negativos preprocesados. a) Dato a predecir en baja resolución. b) Reconstrucción de modelo bicúbico. c) Reconstrucción de modelo SRCNN. d) Reconstrucción de modelo SRGAN. e) Dato real de alta resolución TROPOMI.	36
5.14. Gráfico de medición de error para el primer conjunto de datos con valores negativos preprocesados. a) Dato real de alta resolución TROPOMI. b) Reconstrucción de modelo SRCNN, c) Reconstrucción de modelo SRGAN. d) Error de predicción de SRCNN. e) Error de predicción de SRGAN.	36
5.15. Distribución acumulada de modelos para el conjunto de prueba del segundo conjunto de datos con valores negativos preprocesados. Mayor PSNR y SSIM es mejor, mientras que menor MAPE es mejor.	37
5.16. Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados para el segundo conjunto de datos con valores negativos preprocesados.	38
5.17. Gráfico de predicciones para modelos para el primer conjunto de datos con valores negativos preprocesados. a) Dato a predecir en baja resolución. b) Reconstrucción de modelo bicúbico. c) Reconstrucción de modelo SRCNN. d) Reconstrucción de modelo SRGAN. e) Dato real de alta resolución TROPOMI.	38
5.18. Gráfico de medición de error para el primer conjunto de datos con valores negativos preprocesados. a) Dato real de alta resolución TROPOMI. b) Reconstrucción de modelo SRCNN, c) Reconstrucción de modelo SRGAN. d) Error de predicción de SRCNN. e) Error de predicción de SRGAN.	38
5.19. Reconstrucción de promedio anual de datos de OMI de América del Sur para al año 2007 utilizando redes entrenadas con el primer conjunto de datos sin valores negativos en escala logarítmica. a) Dato de OMI de baja resolución. b) Reconstrucción de SRCNN. c) Reconstrucción de SRGAN.	41
5.20. Reconstrucción de promedio anual de datos de OMI de la zona central y sur de Chile para al año 2007 utilizando redes entrenadas con el primer conjunto de datos sin negativos valores en escala logarítmica. a) Dato de OMI de baja resolución. b) Reconstrucción de SRCNN c) Reconstrucción de SRGAN. . . .	42
5.21. Reconstrucción de promedio anual de datos de OMI de América del Sur para al año 2007 utilizando redes entrenadas con el segundo conjunto de datos sin valores negativos en escala logarítmica. a) Dato de OMI de baja resolución. b) Reconstrucción de SRCNN c) Reconstrucción de SRGAN.	42
5.22. Reconstrucción de promedio anual de datos de OMI de la zona central y sur de Chile para al año 2007 utilizando redes entrenadas con el segundo conjunto de datos sin valores negativos en escala logarítmica. a) Dato de OMI de baja resolución. b) Reconstrucción de SRCNN. c) Reconstrucción de SRGAN. . .	43

Capítulo 1

Introducción

Actualmente, se dispone de técnicas de percepción remota implementadas en satélites que permiten detectar múltiples trazas químicas. Dependiendo de la calidad de los sensores usados y de la órbita satelital, la resolución espacial y temporal de la información puede variar. Los avances tecnológicos han permitido mejorar la resolución espacial a lo largo del tiempo, generando una diferencia significativa entre las tecnologías previas (2004) y las nuevas (2017).

Uno de estos gases que se monitorea es el formaldehído (HCHO), un gas intermedio presente en la cadena de oxidación de muchos compuestos orgánicos volátiles no metánicos (NMVOCs), llevando a la producción de compuestos como el dióxido de carbono (CO_2) [1]. El HCHO es formado por emisiones de fuente biogénicas (principalmente isopreno), antropogénicas (transporte, energía, fuentes industriales) y pirogénicas (incendios), presentando grandes variaciones estacionales e interanuales. Su tiempo de duración es pocas horas debido a la diversidad de procesos capaces de removerlo de la atmósfera, entre los que destacan principalmente la fotólisis, la oxidación y la deposición húmeda. Como resultado, se producen sustancias como el monóxido de carbono (CO), el ozono (O_3) y el ácido fórmico (HCOOH), entre otros [2]. En particular para la formación de O_3 es necesaria la interacción entre HCHO y NO_x ($\text{NO} + \text{NO}_2$), por lo cual es de interés calcular la relación HCHO/ NO_2 como indicador de si la producción de O_3 en la zona es sensible a NO_x o a compuestos orgánicos volátiles (VOCs) [3, 4].

En este trabajo se busca utilizar imágenes del espectrómetro Ozone Monitoring Instrument (OMI) a bordo del satélite AURA, de relativa baja resolución ($13 \times 24 \text{ km}^2$), que recolectó mediciones desde julio de 2004 hasta la llegada del nuevo instrumento Tropospheric Monitoring Instrument (TROPOMI) de relativa alta resolución ($3,5 \times 5,5 \text{ km}^2$), un espectrómetro lanzado en octubre de 2017. El algoritmo de obtención de datos utilizado en ambos casos obtiene la densidad de columna inclinada (SCD) y la pondera por un factor de masa de aire (AMF) para obtener los valores de densidad vertical de la columna (VCD) a utilizar [1]. Como es de esperar, existe una diferencia de calidad entre las imágenes obtenidas por ambos espectrómetros, dificultando la continuidad de los datos. Por ello, el problema a resolver consiste en aumentar la resolución de una imagen de OMI a través de inteligencia artificial, usando como datos de entrenamiento los datos de TROPOMI. Este problema es

conocido como Single Image Super Resolution (SISR), en el cual se reconstruye una imagen a una resolución mayor [5].

1.1. Definición del Problema

Previamente se desarrolló una tesis que resolvía este problema para el gas dióxido de nitrógeno (NO_2) [6], en la cual se construyeron distintas arquitecturas de redes neuronales convolucionales estocásticas. El enfoque evaluó la factibilidad de reconstruir imágenes a una mayor resolución y estimar su error. La conclusión obtenida de este trabajo fue que las *Stochastic Super Resolution Convolutional Neural Networks* (SSRCNNs) basadas en la propuesta de Dong et al. (2016) [7] pueden ser usadas, demostrando que las redes convolucionales obtuvieron un mejor rendimiento que métodos clásicos como la interpolación bicúbica y los métodos basados en codificación dispersa, gracias a la flexibilidad de su estructura. Posteriormente, se incluyó el uso de redes adversarias basado en Stengel et al. (2020) [8]. En otros trabajos de super resolución se han utilizado arquitecturas más recientes que unen redes convolucionales con mecanismos de atención de *transformers* para lidiar con los altos costos en memoria [9]. Si bien los modelos *transformers* han sido bien explorados en estudios para imágenes naturales, su efectividad en percepción remota aún necesita mayor investigación, ya que hay casos en los que no se asegura que superen el rendimiento de las redes convolucionales [10].

Para el presente proyecto de tesis, el problema principal consiste en adaptar los resultados del trabajo realizado con NO_2 y aplicarlos a aldehídos medidos por los mismos espectrómetros. Por lo tanto, es necesario continuar con el trabajo y adaptarlo a una nueva traza, utilizando metodologías que en el trabajo previo solo fueron mencionadas como trabajo futuro. De esta manera, se planea dar continuidad a las series de datos de cierta resolución para que puedan ser trabajados y luego caracterizar la evolución de la fotoquímica urbana de ciudades de América del Sur según Duncan et al. (2010) [11]. Para esto es necesario disponer tanto de HCHO como de NO_2 , considerando que ambos tiempos de residencia son cortos, del orden de horas [4]. Las emisiones de NO_2 son principalmente antropogénicas impactadas principalmente por la quema de combustibles fósiles, mientras que el origen del HCHO tiene tanto fuentes antropogénicas, como biogénicas. Debido a que sus fuentes son distintas, la percepción remota de estas trazas es diferente, por lo que es necesario estudiar ambas.

1.2. Objetivos

1.2.1. Objetivo General

El objetivo general de esta tesis es realizar una mejora de resolución espacial imágenes satelitales de aldehídos por un factor de 2, obtenidas por el espectrómetro Ozone Monitoring Instrument (OMI) a bordo del satélite AURA con resultados de mejor calidad que los obtenidos por la interpolación bicúbica.

1.2.2. Objetivos Específicos

Con el fin de alcanzar el objetivo general anteriormente mencionado, se definieron los siguientes objetivos específicos:

- Formalizar la introducción del uso de redes adversarias sobre el método SRCNN y generar la red correspondiente.
- Evaluar los resultados y la factibilidad de uso de la red generada para el propósito de la reconstrucción para el período 2004-presente.
- Explorar mejoras a las técnicas SRCNN y redes adversarias.

1.3. Metodología

La metodología a utilizar en el transcurso del trabajo de tesis consta de varias fases, comenzando con la exploración de datos. En esta fase, se llevarán a cabo análisis estadísticos y visualizaciones para describir los datos y obtener una mejor comprensión de ellos. Esta etapa es crucial para identificar patrones, anomalías y características importantes de los datos que influirán en las etapas posteriores del proceso.

La siguiente fase es el pre-procesamiento de los datos, que implica la limpieza, transformación y preparación de los datos para su posterior manejo. Este paso asegura que los datos estén en el formato adecuado y que cualquier error, falta de información o ruido presente en los datos sea tratado adecuadamente. El objetivo es obtener un conjunto de datos refinado y listo para ser utilizado en la construcción de modelos.

Luego, se procederá con la generación de modelos de redes neuronales convolucionales con el fin de mejorar la resolución espacial de las imágenes satelitales. Esta fase implica diseñar y seleccionar arquitecturas de redes neuronales que sean capaces de realizar la tarea de superresolución de manera precisa medida según métricas de calidad.

El entrenamiento es la fase siguiente, en la cual se entrenarán los modelos y se ajustarán los hiper parámetros necesarios para resolver el problema. Este proceso consiste en comparar imágenes de baja resolución con imágenes de alta resolución, permitiendo que el modelo aprenda a generar imágenes de mayor calidad a partir de datos de menor resolución.

Finalmente, se realizará la evaluación del modelo. Esta fase implica evaluar el rendimiento del modelo mediante el uso de métricas específicas como el PSNR (*Peak Signal-to-Noise Ratio*), el SSIM (*Structural Similarity Index Measure*) y el MAPE (*Mean Absolute Percentage Error*). Estas métricas permitirán determinar la efectividad y precisión del modelo en la mejora de la resolución de imágenes satelitales.

1.4. Estructura de Tesis

La siguiente tesis está organizada de la siguiente manera: el Capítulo 2 ofrece una breve explicación teórica necesaria para entender el problema, abarcando los fundamentos del gas aldehído, la superresolución, redes neuronales convolucionales y generativas adversarias, las métricas de evaluación a utilizar, y finaliza con las operaciones aplicadas a grillas de datos. En el Capítulo 3 se encuentra la descripción y análisis de los datos satelitales de formalehído utilizados en este trabajo. La descripción de los métodos a implementar, asociados al estado del arte en técnicas de Super Resolution, puede encontrarse en el Capítulo 4. A continuación, en el Capítulo 5, se presentan los resultados obtenidos mediante la aplicación de los distintos modelos. El Capítulo 6 resume las conclusiones obtenidas en este trabajo y presenta la descripción del plan de trabajo a seguir.

Capítulo 2

Marco Teórico

En este capítulo se explicará en detalle la teoría necesaria para entender los conceptos más generales asociados al trabajo a realizar. Inicialmente, se proporcionará un contexto del gas aldehído junto a la información satelital relevante. Posteriormente, se abordarán las definiciones del problema a resolver y se describirán las herramientas utilizadas.

2.1. Columna de Densidad Troposférica de HCHO

El gas formaldehído (HCHO) tiene una gran relevancia en la atmósfera. Influye en el proceso de formación de productos como el monóxido de carbono (CO), el ozono troposférico (O₃) y otros contaminantes secundarios, es decir, no emitidos sino que formados en la atmósfera. Su principal fuente de producción global en la atmósfera corresponde a la oxidación del metano (CH₄), mientras que en áreas continentales se ve influenciado principalmente por la oxidación de compuestos orgánicos volátiles no metánicos (nonmethane volatile organic compounds o NMVOCs) [2]. En la figura 2.1 se muestran otros gases precursores y productos asociados al formaldehído. El HCHO tiene un tiempo de vida corto en la atmósfera (del orden de horas), lo cual permite asociarlo directamente a sus fuentes de emisión. Los procesos encargados de su remoción son principalmente la fotólisis, la oxidación y la deposición húmeda. El HCHO afecta a la salud de las personas, provocando tos, irritación en los ojos, alergias en la piel, problemas respiratorios como el asma, y en casos de exposición crónica genera problemas como mareos, desordenes del sueño y pérdida de memoria [12].

Debido a todo esto, monitorear los niveles de HCHO es crucial para estudiar la calidad del aire. Una forma de medición de HCHO es a través de instrumentos como los espectrómetros unidos a satélites que orbitan la Tierra. Estos dispositivos son capaces de medir distintos componentes tanto en la tropósfera como en estratósfera, proporcionando información sobre la columna de gases que analizan. Los espectrómetros OMI y TROPOMI tienen la capacidad de medir las masas de columnas de formaldehídos, además de otras sustancias, posibilitando la detección diaria de diversos contaminantes [1]. El uso de datos de formaldehído (HCHO) obtenidos por el instrumento TROPOMI ha sido validado en varias ocasiones, incluyendo estudios en América del Sur [2, 13], lo cual justifica su aplicación y confiabilidad en estudios

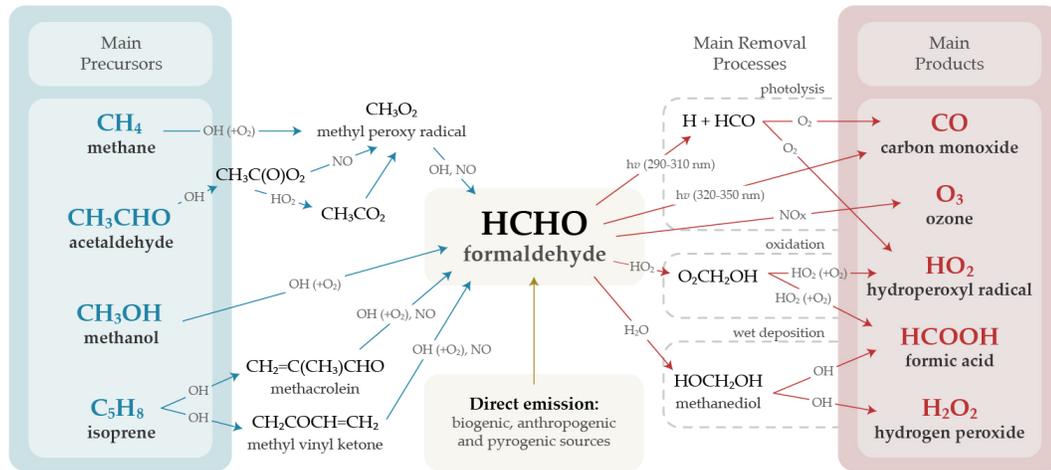


Figura 2.1: Descripción de la principal producción y remoción de formaldehído atmosférico. Fuente Dias et al., 2020 [2].

atmosféricos. Además, es crucial tener en cuenta las diferencias en resolución entre los espectrómetros OMI ($13 \times 24\text{km}^2$) y TROPOMI ($3, 5 \times 5, 5\text{km}^2$), siendo este último capaz de proporcionar mediciones con una mayor resolución espacial y temporal, lo cual representa un avance significativo para la monitorización y estudio de la calidad del aire a escala global.

La obtención de información de los instrumentos se basa en el algoritmo DOAS (Differential Optical Absorption Spectroscopy); [14] es una técnica utilizada para analizar espectros ópticos y determinar la concentración de gases atmosféricos, como HCHO, SO_2 y NO_2 [1], a partir de mediciones espectrales. Este método se fundamenta en la Ley de Beer-Lambert, que establece que la cantidad de luz absorbida por una sustancia es proporcional a la concentración de la sustancia y a la distancia que recorre la luz a través de ella. En el contexto del DOAS, se emplean espectrómetros que captan la luz solar reflejada o transmitida a través de la atmósfera. El proceso de análisis comienza separando el espectro medido en bandas de longitud de onda, generalmente en el ultravioleta al visible (UV-VIS). El objetivo es identificar las líneas espectrales características del gas de interés, como el formaldehído (HCHO), así como de otros gases atmosféricos como el dióxido de azufre (SO_2) y dióxido de nitrógeno (NO_2). A partir de estas líneas espectrales, el algoritmo DOAS calcula cuánta luz ha sido absorbida por el gas durante su trayectoria a través de la atmósfera. Una vez identificadas las características espectrales del gas, el DOAS utiliza modelos matemáticos para ajustar y cuantificar la absorción específica del gas en cada banda espectral. Este ajuste se realiza comparando el espectro observado con un espectro de referencia calculado, teniendo en cuenta factores como la dispersión y la absorción por otros componentes atmosféricos (ver figura 3.2). La salida del algoritmo proporciona mediciones de columnas verticales de concentración de gases, fundamentales para estudios atmosféricos y el monitoreo de la calidad del aire a escala global.

2.2. Transformada de Coseno Discreta para llenado de datos

Las grillas de datos son una herramienta fundamental para estructurar la información de manera ordenada y accesible. En este contexto, cada elemento de la grilla representa una celda que contiene un dato específico con el cual se pueden realizar diversas operaciones. Esta estructura facilita la manipulación, análisis y visualización de datos en diversos campos, permitiendo trabajar eficientemente con grandes volúmenes de información. Específicamente, en conjuntos de datos geofísicos es común utilizar grillas regulares de varias dimensiones para estructurar la información de manera sistemática. Sin embargo, debido a la frecuente presencia de datos faltantes en estos conjuntos de datos, es esperable la necesidad de rellenar estos vacíos de manera efectiva. Para abordar esta necesidad, se ha utilizado la Transformada de Coseno Discreta de forma extensiva en la recuperación de datos geofísicos debido a su uso estudiado en el llenado de este tipo de datos [15, 16].

La Transformada de Coseno Discreta (Discrete Cosine Transform o DCT) es una técnica ampliamente utilizada para el robusto llenado de datos faltantes en una o más dimensiones. Este método emplea un enfoque de penalización de mínimos cuadrados, mostrando una mayor eficiencia especialmente cuando los datos están distribuidos en una grilla equiespaciada [17].

El proceso queda descrito por la ecuación 2.1

$$\hat{X} = IDCT(\Gamma \circ DCT(W \circ (X - \hat{X}) + \hat{X})) \quad (2.1)$$

Donde X corresponde a los datos con valores faltantes, W es un arreglo binario del mismo tamaño de X indicando si tiene o no un valor faltante, \hat{X} es el valor a predecir y el operador \circ hace referencia al producto de Schur (producto elemento a elemento). Adicionalmente, se considera que Γ es un tensor tridimensional de filtrado, definido según la ecuación 2.2.

$$\Gamma_{i_1, i_2, i_3} = \left(1 + s \left(\sum_{j=1}^3 \left(2 - \cos \frac{(i_j - 1)\pi}{n_j} \right) \right)^2 \right)^{-1} \quad (2.2)$$

donde i_j denota el i -ésimo elemento en el eje j , y n_j corresponde al tamaño de X a lo largo de dicha dimensión. También se utiliza el parámetro s , que controla el nivel de suavizado aplicado a los datos. Para el llenado de datos, es crucial considerar un valor infinitesimal (≈ 0), ya que un valor alto de s puede llevar a la pérdida de componentes de alta frecuencia. Sin embargo, un valor demasiado pequeño de s puede provocar sobreajuste en los datos, lo que representa un problema para la capacidad de generación al predecir valores faltantes en el conjunto de datos [16].

2.3. Super-Resolución de imagen única

La Super-resolución de imagen única o single image Super-Resolution (SR) es un problema de visión computacional que consiste en inferir imágenes de alta resolución (HR) a partir de imágenes de menor resolución (LR). Dependiendo de la cantidad de imágenes LR utilizadas, el problema se clasifica como super resolución de imagen única (SISR) o super resolución de múltiples imágenes (MISR). Esta técnica tiene diversas aplicaciones, incluyendo imágenes médicas, seguridad y vigilancia [18, 19]. La super resolución es un problema mal planteado (ill-posed), ya que una imagen LR puede corresponder a múltiples imágenes HR posibles, lo que implica que no existe una única solución de reconstrucción. Además, la relación entre la imagen de alta resolución y la de baja resolución puede variar significativamente según el contexto, por lo cual es esencial considerar el origen y las características específicas de las imágenes al abordar este problema [20].

El procedimiento se modela en base a la ecuación 2.3, cuyo objetivo es obtener, a partir de la imagen de baja resolución I_x , una aproximación de alta resolución \hat{I}_y de la imagen original en resolución alta I_y .

$$\hat{I}_y = \mathcal{F}(I_x; \theta), \quad (2.3)$$

donde \mathcal{F} es el modelo de Super Resolución y θ corresponde a los parámetros de \mathcal{F} .

Para obtener buenos resultados, es fundamental utilizar un modelo cuyos parámetros se ajusten adecuadamente. Por ello, es de suma importancia que los parámetros se adapten correctamente al modelo, como se plantea en el objetivo de la ecuación 2.4:

$$\hat{\theta} = \arg \min_{\theta} \mathcal{L}(\hat{I}_y, I_y) + \lambda \Phi(\theta) \quad (2.4)$$

donde $\mathcal{L}(\hat{I}_y, I_y)$ representa la función de pérdida entre la imagen generada \hat{I}_y y la imagen original I_y , $\Phi(\theta)$ es el término de regularización y λ es parámetro de *trade-off*.

2.4. Redes Neuronales Convolucionales

Las Redes Neuronales Convolucionales (CNNs, por sus siglas en inglés) son una herramienta poderosa para procesar datos organizados en grillas de distintas dimensiones. Estas redes se utilizan comúnmente para trabajar con series de tiempo (1D), imágenes (2D) o videos (3D). Una de sus principales características es la capacidad de establecer conexiones que permiten la comunicación entre datos y sus vecindades. Esto facilita la extracción de características locales importantes en el análisis de datos complejos. Además, las CNNs destacan por su implementación simple, rapidez y alta adaptabilidad a distintos tipos de datos [7].

La convolución es una operación matemática fundamental en el funcionamiento de las Redes Neuronales Convolucionales (CNNs). En el contexto de las CNNs, la convolución se utiliza para extraer características locales de los datos mediante el deslizamiento de un filtro o kernel sobre la entrada. Este filtro realiza una serie de multiplicaciones y sumas entre sus valores y los valores correspondientes en la entrada, generando una nueva matriz de características denominada mapa de características [21].

En dos dimensiones, la convolución se define matemáticamente como:

$$G[i, j] = (H * F)(i, j) = \sum_m \sum_n H(m, n) \cdot F(i - m, j - n) \quad (2.5)$$

donde F es la matriz de entrada, por ejemplo, una imagen, H es el kernel o filtro de tamaño (m, n) , y (i, j) son las coordenadas del elemento en el mapa de características resultante. La operación es una transformación lineal asociativa y conmutativa, que consiste en desplazar el kernel H sobre la matriz de entrada F , multiplicando los elementos superpuestos y sumando los resultados para obtener cada valor en el mapa de características. Este proceso permite a las CNNs detectar patrones y características específicas en los datos, como bordes, texturas, y otras estructuras visuales, que son esenciales para tareas como el reconocimiento de imágenes y la detección de objetos.

2.5. Redes Generativas Adversarias

Las Redes Generativas Adversarias (GANs, por sus siglas en inglés) son redes en las que se entrenan simultáneamente un modelo generador y un modelo discriminador [22]. El modelo generador G busca capturar la distribución de los datos, mientras que el modelo discriminador D estima la probabilidad de que una muestra provenga de la distribución del modelo generador o de la distribución de los datos reales. El buen funcionamiento de estas redes se basa en el entrenamiento iterativo y adversarial de ambos modelos, lo cual lleva a que el generador produzca resultados cada vez más realistas y el discriminador mejore en distinguir entre datos reales y generados. Este proceso de entrenamiento puede plantearse como un problema de optimización de la ecuación 2.6 [8, 18].

$$\min_G \max_D \mathbb{E}[\log(D(I_y))] + \mathbb{E}[\log(1 - D(G(I_x)))] \quad (2.6)$$

donde I_y es una imagen del conjunto de datos de alta resolución mientras que I_x es una imagen del conjunto de imágenes de baja resolución.

2.6. Redes Neuronales Convolucionales para Super Resolution

En los últimos años, los métodos basados en Inteligencia Artificial han predominado en la resolución de problemas de Super Resolución (SR). La ventaja sobre los métodos clásicos, como la interpolación bicúbica, que se utiliza actualmente como referencia, ha establecido el uso de redes neuronales como el estado del arte desde el modelo propuesto por Dong et al. (2016) [7]. Los esfuerzos se han centrado en adaptar las redes al problema de SR, lo que ha llevado al uso de modelos de Redes Adversarias Generativas (GANs) [22] y Redes Residuales (ResNet) [23]. Aplicaciones recientes de estas arquitecturas han demostrado un rendimiento superior en comparación con modelos genéricos de SR [20, 24, 25].

2.7. Métricas de evaluación

Las métricas a utilizar son de gran relevancia, ya que establecen un criterio objetivo para evaluar la calidad de la reconstrucción de una imagen. Estas mediciones se realizan comparando la imagen original a la cual se desea llegar I_y con la imagen reconstruida \hat{I}_y .

2.7.1. Proporción Máxima de Señal a Ruido

La proporción Máxima de Señal a Ruido (*Peak Signal-to-Noise Ratio*, PSNR) es una métrica que se utiliza para detectar la pérdida de información en una imagen, evaluando la calidad de la reconstrucción en comparación con la imagen original [26]. PSNR se expresa en decibelio (dB) y es inversamente proporcional al error cuadrático medio (MSE) entre las dos imágenes. Se define en base a la ecuación 2.7

$$PSNR = 10 \cdot \log_{10} \left(\frac{L^2}{MSE} \right) \quad (2.7)$$

Donde L es el valor máximo posible de un píxel en la imagen (por ejemplo, 255 para una imagen de 8 bits), y MSE es el error cuadrático medio entre la imagen original I_y y la imagen reconstruida \hat{I}_y , calculado como:

$$MSE = \frac{1}{N} \sum_{i=1}^N (I(i) - \hat{I}(i))^2 \quad (2.8)$$

con N la cantidad de píxeles en la reconstrucción. De esta forma se mide la diferencia entre valores de píxeles, aportando información para procesos de reconstrucción.

2.7.2. Similaridad estructural

El índice de similitud estructural (*Structural Similarity*, SSIM) es una métrica para medir la similitud estructural entre dos imágenes [27]. A diferencia de las métricas tradicionales como el PSNR, que se centran en las diferencias absolutas entre los valores de los píxeles, SSIM se basa en comparaciones independientes de luminiscencia, contraste y estructura. Este enfoque permite evaluar mejor la calidad percibida de la imagen, con valores en el rango $[0,1]$, siendo 1 el máximo valor posible. La fórmula general para calcular SSIM entre dos imágenes I e \hat{I} es:

$$SSIM(I, \hat{I}) = [l(I, \hat{I})]^\alpha \cdot [c(I, \hat{I})]^\beta \cdot [s(I, \hat{I})]^\gamma \quad (2.9)$$

donde $l(I, \hat{I})$, $c(I, \hat{I})$ y $s(I, \hat{I})$ son las funciones de comparación de luminiscencia, contraste y estructura, respectivamente, y α , β y γ son parámetros que ajustan la importancia relativa de cada uno de estos componentes. Dichas funciones específicas se definen de la siguiente manera:

$$l(I, \hat{I}) = \frac{2\mu_I\mu_{\hat{I}} + C_1}{\mu_I^2 + \mu_{\hat{I}}^2 + C_1} \quad (2.10)$$

$$c(I, \hat{I}) = \frac{2\sigma_I\sigma_{\hat{I}} + C_2}{\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2} \quad (2.11)$$

$$s(I, \hat{I}) = \frac{\sigma_{I\hat{I}} + C_3}{\sigma_I\sigma_{\hat{I}} + C_3} \quad (2.12)$$

Aquí, μ_I y $\mu_{\hat{I}}$ son las medias de las imágenes I e \hat{I} , σ_I y $\sigma_{\hat{I}}$ son las desviaciones estándar de I e \hat{I} , y $\sigma_{I\hat{I}}$ es la covarianza entre I e \hat{I} . Los términos C_1 , C_2 y C_3 son constantes para estabilizar las divisiones con denominadores pequeños, típicamente definidos como funciones del rango dinámico de los valores de píxeles.

En la práctica, la fórmula SSIM simplificada se usa con $\alpha = \beta = \gamma = 1$, $C_3 = C_2/2$, $C_1 = (k_1L)^2$ y $C_2 = (k_2L)^2$ que son constantes que evitan inestabilidad, donde además $k_1 \ll 1$ y $k_2 \ll 1$:

$$SSIM(I, \hat{I}) = \frac{(2\mu_I\mu_{\hat{I}} + C_1)(2\sigma_I\sigma_{\hat{I}} + C_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + C_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2)} \quad (2.13)$$

SSIM es ampliamente utilizado en aplicaciones de procesamiento de imágenes y visión computacional, ya que proporciona una evaluación perceptualmente más relevante de la similitud entre dos imágenes en comparación con las métricas tradicionales basadas únicamente en los errores de píxeles [28].

2.7.3. Error porcentual absoluto medio

El error porcentual absoluto medio (Mean Absolute Percentage Error, MAPE en inglés) es una métrica que mide la magnitud media del error entre dos imágenes en términos porcentuales [29]. Es ampliamente utilizado para la validación de modelos predictivos debido a su interpretación intuitiva como el porcentaje de error.

La fórmula para calcular MAPE entre una imagen original I_y y una imagen reconstruida \hat{I}_y es:

$$MAPE = \frac{100}{N} \sum_{i=1}^N \frac{|I_y(i) - \hat{I}_y(i)|}{\max(|I_y(i)|, th)} \quad (2.14)$$

donde N es el número total de píxeles en la imagen, $I_y(i)$ es el valor del píxel i en la imagen original, $\hat{I}_y(i)$ es el valor del píxel i en la imagen reconstruida y th es el valor umbral (*threshold*, abreviado *th*); en este trabajo es considerado igual a 2,36 [Pmolec/cm²]. Este valor se obtuvo donde la columna de HCHO es cercana a 0 y calculando promedios sobre regiones sobre el Océano Pacífico, donde es esperable que los valores de concentración sean cercanos a cero.

MAPE tiene varias ventajas. En primer lugar, expresa el error como un porcentaje, lo que facilita la interpretación y comparación entre diferentes conjuntos de datos o modelos. En segundo lugar, MAPE es indiferente a la escala de los valores de los píxeles, lo que hace adecuado para comparar errores en imágenes con diferentes rangos de valores de píxeles.

Sin embargo, el MAPE tiene algunas limitaciones. Es particularmente sensible a los valores cercanos a cero en la imagen original I_y , ya que dividir por valores pequeños puede resultar en valores extremadamente grandes de MAPE. Para mitigar este efecto, se pueden aplicar técnicas de suavizado o excluir los píxeles con valores cercanos a cero. En este trabajo, se optó por este último método, excluyendo los valores por debajo del umbral definido para el valor del píxel máximo.

Capítulo 3

Análisis y Diseño de Datos

3.1. Datos satelitales

El gas formaldehído (HCHO) tiene un gran impacto en la calidad del aire, por lo que es crucial disponer de mediciones confiables de su concentración. Durante años, se han utilizado observaciones satelitales de las columnas troposféricas de HCHO para el monitoreo de la calidad del aire y estudios climáticos, tanto a escala regional como global [2, 3, 30], . Existe un gran interés en determinar las concentraciones en lugares donde las emisiones son mayores, tales como áreas con vegetación densa, zonas afectadas por incendios y regiones con alto tráfico vehicular [1]. La determinación de la distribución geográfica del HCHO es, por lo tanto, de gran relevancia. Debido al corto tiempo de permanencia del formaldehído en la atmósfera, sus concentraciones presentan una alta heterogeneidad espacial [13]. En consecuencia, los instrumentos satelitales han mejorado su resolución espacial a lo largo de los años para proporcionar datos más precisos sobre la cantidad y la ubicación de este contaminante.

Uno de los métodos ampliamente utilizado para obtener datos de la columna de densidad troposférica de formaldehído (HCHO) es mediante el uso de instrumentos satelitales, que emplean el algoritmo Differential Optical Absorption Spectroscopy (DOAS). Como se dijo antes, este algoritmo se fundamenta en la aplicación de la ley de Beer-Lambert, que permite separar las estructuras espectrales de absorción en el espectro para aislar gases traza [31]. Una característica del formaldehído es que comparte bandas espectrales con gases como el O₃, el SO₂ y BrO, las cuales es necesario aislar (ver figura 3.2).

Para obtener la columna de densidad vertical de HCHO, el método transforma una columna de densidad inclinada obtenida por el espectrómetro mediante la operación con un factor de masa de aire (Air Mass Factor o AMF). La columna inclinada se deriva de la absorción de la sección eficaz (cross section¹) del HCHO (ver figura 3.1).

El HCHO es de absorción débil y más difícil de detectar que otros gases debido a que tiene una densidad óptica baja², llegando a ser 1 orden de magnitud menor a la del NO₂ y 3

¹Cantidad de flujo óptico reflejado hacia la fuente emisora.

²Habilidad de un material de transmitir luz a través de él.

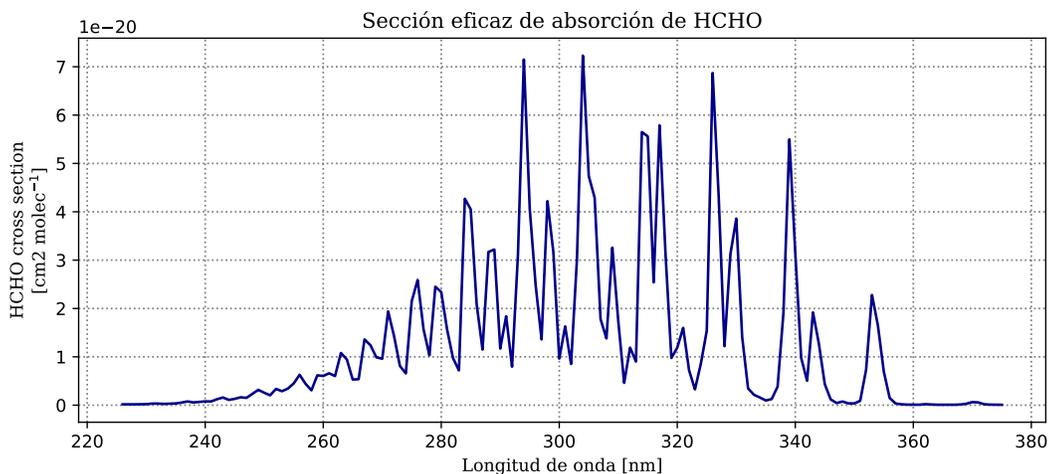


Figura 3.1: Sección eficaz de absorción de HCHO a 298 °K. Fuente de datos: Meller y Moortgat, 2000 [32].

ordenes de magnitud más bajo que el del O₃ [31].

3.2. Datos de HCHO

Las fuentes de datos utilizadas en este trabajo provienen de los instrumentos Ozone Monitoring Instrument (OMI) y el Tropospheric Monitoring Instrument (TROPOMI). OMI es un espectrómetro a bordo del satélite AURA, lanzado en Julio de 2004, diseñado para medir la radiación solar retrodispersada por la atmósfera y la superficie de la Tierra en líneas de absorción de longitudes de onda que van desde 270 a 500 nm [1, 34]. La resolución de los píxeles varía desde $13 \times 24 \text{ km}^2$ en el punto nadir hasta $28 \times 150 \text{ km}^2$ a los bordes del área de barrido. Por otra parte, TROPOMI es un espectrómetro a bordo del satélite Sentinel-5 Precursor (S5P), lanzado en Octubre de 2017, diseñado para medir en las regiones espectrales ultravioleta, visible, infrarrojo cercano e infrarrojo de onda corta. TROPOMI ofrece una alta resolución espacial en nadir de $3,5 \times 7 \text{ km}^2$, mejorada a $3,5 \times 5,5 \text{ km}^2$ en Agosto de 2019.

Cada archivo de datos corresponde a las mediciones realizadas a lo largo de la órbita del satélite, presentando los datos en un formato de grilla irregular donde cada píxel de concentración está georreferenciando por su latitud y longitud. Además de la concentración de formaldehído, cada valor en la grilla incluye un valor de calidad (quality assurance value o qa_value) [35] (ver figura 3.3). Este valor proporciona una indicación de la confianza en la medición asociada, considerando variables como la presencia de nubes, nieve, hielo y otros factores ambientales que pueden influir en la precisión de las mediciones.

Los datos seleccionados para este trabajo están limitados a América del Sur, en el rango de longitud (-86° , -31°) y en latitud (-58° , 14°). Esta región presenta una diversidad notable entre áreas urbanas y zonas de vegetación, lo cual es crucial para comprender cómo se distribuye el formaldehído a una mayor resolución. Este enfoque se espera que contribuya significativamente a estudios diversos en áreas relacionadas.

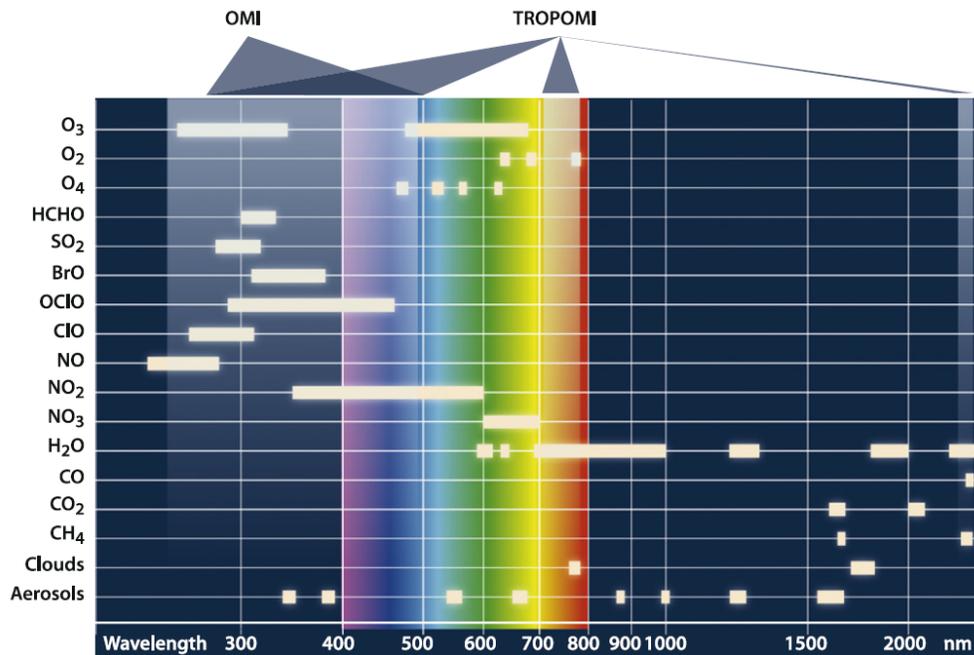


Figura 3.2: Rango espectral para OMI y TROPOMI. Fuente: Veefkind, 2012 [33].

Debido a la naturaleza irregular de las grillas de datos utilizadas, no es posible realizar operaciones directas entre grillas irregulares debido a la falta de coincidencia en sus celdas en términos de latitud y longitud. Por lo tanto, surge la necesidad de transformar estas grillas irregulares a una grilla regular común, y de esta manera poder operar grillas asociadas a mediciones de distintos tramos orbitales; este proceso es conocido como regrillado (regridding). Esta transformación permite alinear cada píxel de la grilla irregular con las celdas de la nueva grilla regular, asignando a cada píxel una contribución proporcional basada en el área que ocupa en la nueva celda.

Para manejar correctamente los datos, es esencial comprender su distribución, ya que esta influye en la toma de decisiones y los procedimientos a realizar, moldeando finalmente los resultados. La distribución de los datos originales muestra que la mayoría de las concentraciones se encuentran en un rango acotado de valores. En el histograma presente en la figura 3.5, se puede observar que los valores bajos (cerca de 0) tienen una alta frecuencia, mientras que los más alejados de 0, tanto positivos como negativos, tienen una frecuencia muy baja. Esto se confirma con la función de distribución acumulada, que indica que la mayoría de los valores se concentran alrededor de 0, mientras que las concentraciones muy negativas o muy altas tienen poca probabilidad de ocurrencia.

Un detalle importante a resaltar es la presencia de valores negativos. Las concentraciones de HCHO corresponden a cantidades del contaminante, por lo que un valor negativo carece de sentido físico y debe considerarse cuidadosamente en el análisis. Además, se identifican valores fuera de rango tanto positivos como negativos, que pueden afectar el análisis. Estos valores extremos son observaciones numéricamente distantes del resto de los datos y pueden distorsionar las conclusiones si no se maneja adecuadamente.

Otra característica notable de los datos es la presencia de datos faltantes. Aproximada-

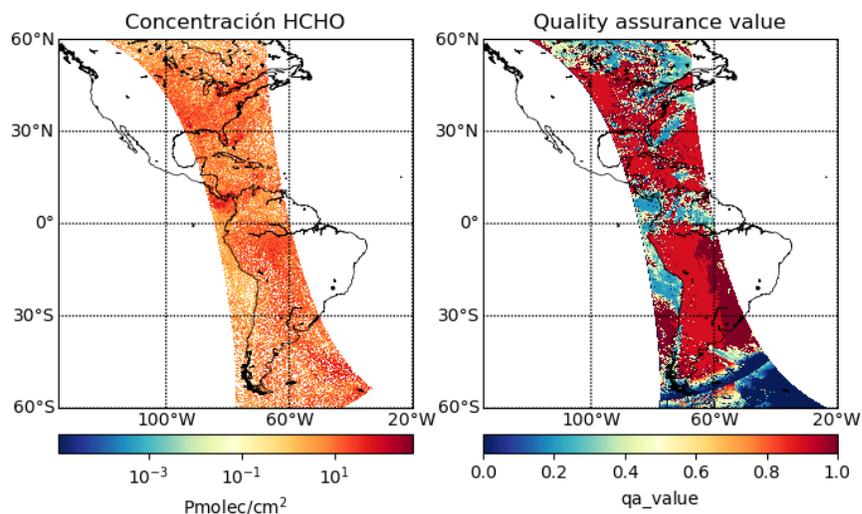


Figura 3.3: (a) Concentración de HCHO en grilla irregular en escala logarítmica. (b) Valor de calidad de datos. Datos del mismo archivo de origen ejemplificando el recorrido orbital.

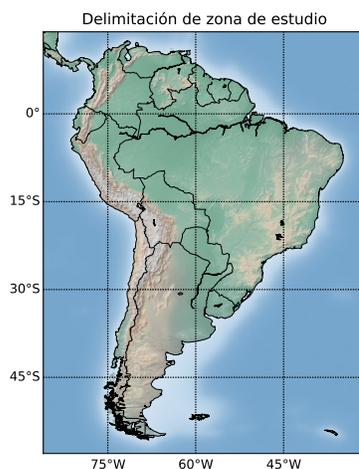


Figura 3.4: Delimitación de zona de estudio.

mente un 18 % de los datos originales carecen de valor. Esta falta de datos presenta problemas para tareas posteriores, ya que estos valores ausentes están distribuidos por todo el archivo, generando discontinuidades en las vecindades dentro de la misma grilla. Estas discontinuidades pueden complicar el análisis y procesamiento de los datos, afectando la precisión y consistencia de los resultados obtenidos. Por lo tanto, es crucial abordar de manera efectiva la gestión de estos datos faltantes para garantizar la integridad del análisis posterior.

3.3. Preprocesamiento

Dado que el objetivo es mejorar la resolución espacial de datos satelitales, es necesario adaptar los datos para cumplir con este propósito. Para resolver los problemas de falta de datos y el uso irregular de grillas, se realiza un preprocesamiento exhaustivo. Todos los

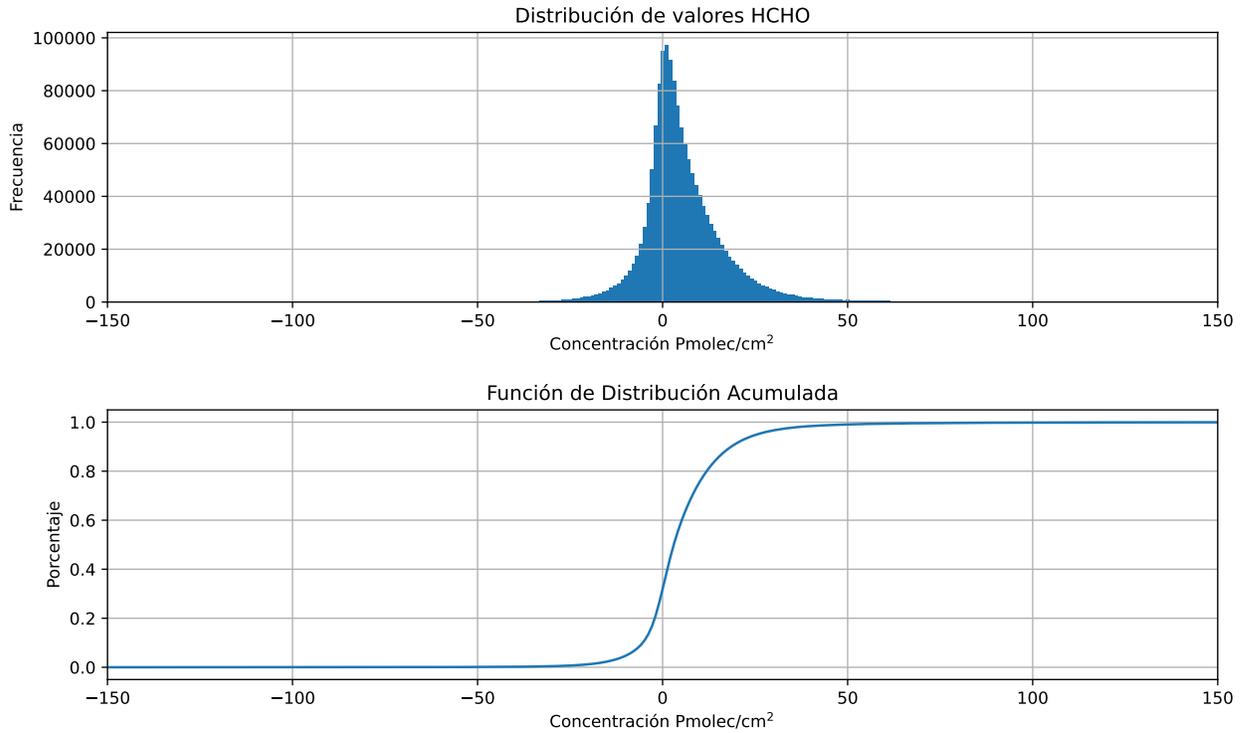


Figura 3.5: Distribución de datos.

procedimientos se aplican a los datos de TROPOMI, ya que estos datos serán fundamentales para realizar entrenamientos y predicciones sobre datos de OMI. El primer paso es considerar solo los valores útiles según el valor de calidad (figura 3.3.b). Para ello, se verifica que los valores de concentraciones de la grilla original tengan asociado a un `qa_value` mayor o igual a 0,5; de lo contrario, esos valores son filtrados [1, 2, 35]. Este proceso asegura que solo se utilicen datos de alta calidad, mejorando así la precisión y fiabilidad de los análisis posteriores.

3.3.1. Regrillado

El siguiente procedimiento que se realiza es el de regrillado, que como ya se mencionó consiste en transformar los datos de una grilla irregular (celdas de $3,5 \times 5,5 \text{ km}^2$) de origen a una grilla regular común. Este proceso se realiza pasando a una grilla regular que cubre todo el espacio de coordenadas, es decir, $(-180^\circ, 180^\circ)$ en longitud y $(-90^\circ, 90^\circ)$ en latitud. De esta forma cualquier grilla irregular puede ser transformada a esta nueva grilla regular. El tamaño de cada celda de la nueva grilla es de $(0,05^\circ, 0,05^\circ)$, puesto que cada dato final a predecir de OMI se encuentra en una grilla de $(0,1^\circ, 0,1^\circ)$, esperando realizar una mejora de reconstrucción en un factor de 2.

Al realizar el proceso de regrillado, cada celda de la nueva grilla es llenada con el aporte de datos de la grilla original, ocupando una fracción del área de esta nueva celda. El valor final que tiene una celda de la nueva grilla es representativo del área que aborda en latitud y longitud, por lo que los datos con los que se obtuvo desde la grilla original deben aportar la suficiente información para indicar que el valor representado en esa celda es representativo.

Para ello, solo se consideran las celdas de la nueva grilla que tienen su área ocupada en al menos un 50% (ver figura 3.6). Con esto las celdas que tienen un porcentaje menor ocupado son filtradas ya que no se considera que su valor sea representativo. Todo el proceso de regrillado y filtrado mencionados anteriormente con el toolbox Harp³ gracias a su uso estandarizado en otros trabajos, simplicidad de uso y soporte de manejo de datos de los productos requeridos [1].

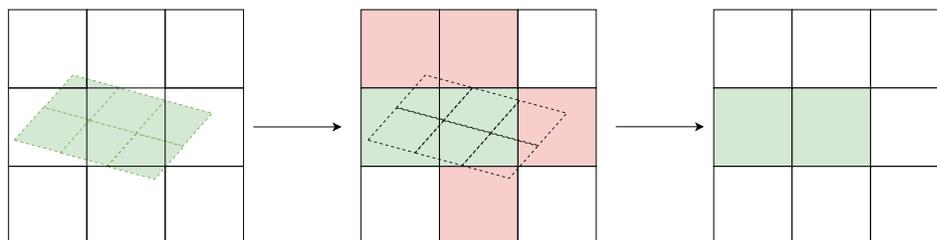


Figura 3.6: Diagrama de regrillado de dato a grilla regular. En verde se encuentran las celdas que cumplen la condición del llenado al 50%, y en rojo la que no cumplen la condición.

3.3.2. Llenado de datos

Una forma de mitigar el problema de la falta de datos es utilizando una función de llenado de datos faltantes. Para su utilización es necesario establecer dos parámetros de relevancia, el tamaño máximo de datos espacialmente cercanos a rellenar (**size**) y el parámetro del nivel de suavizado (**s**). El parámetro de tamaño indica la cantidad máxima de celdas de grilla que puede tener un agujero sin datos para ser llenado, por lo que si el agujero tiene un número mayor a dicha cantidad de celdas no será llenado. Es importante considerar que esta función puede llenar múltiples agujeros de una grilla mientras cumplan con la condición de tamaño. Por otra parte el parámetro de suavizado indica cuan suave será el resultado.

Para la obtención de estos parámetros se generó un agujero de tamaño arbitrario **size** en una posición aleatoria dentro del archivo y se generó de forma aleatoria agujeros de una celda en un 18% del archivo (fracción de datos faltantes promedio por archivo de TROPOMI HCHO), y se rellenó el con los datos disponibles para un parámetro de suavizado **s** y se calculó el error RMSE entre los datos originales y los datos rellenados (ver figura 3.7). Este proceso fue repetido para distintas combinaciones de tuplas (**s**, **size**), con lo cual se obtuvo un error asociado a cada tupla. Una vez obtenido el error para distintas combinaciones de (**s**, **size**) para múltiples archivos, se promediaron las curvas obtenidas generando una curva representativa para el valor de **s** y otra para el valor de **size**, presentes en la figura 3.8. En la curva de **s** es directo evidenciar que el valor con menor error asociado es $s = 0,01$. En cuanto al valor de **size**, para valores de tamaño mayores a 14 tiende a ser constante. El valor escogido fue $size = 18$ puesto que para tamaños con un error considerablemente menor la cantidad de celdas a llenar era muy baja, mientras que es posible experimentar con valores mayores puesto que como se puede notar en el gráfico, el error no es tan sensible al aumento de tamaño del agujero a llenar.

³<https://atmospherictoolbox.org/harp/>

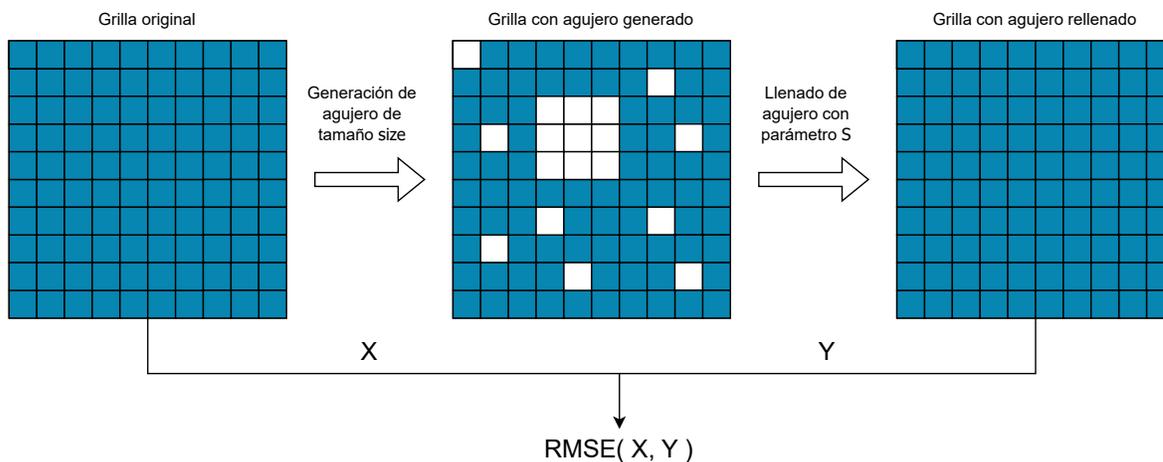


Figura 3.7: Diagrama de obtención de error RMSE para distintas combinaciones de **size** y **s**.

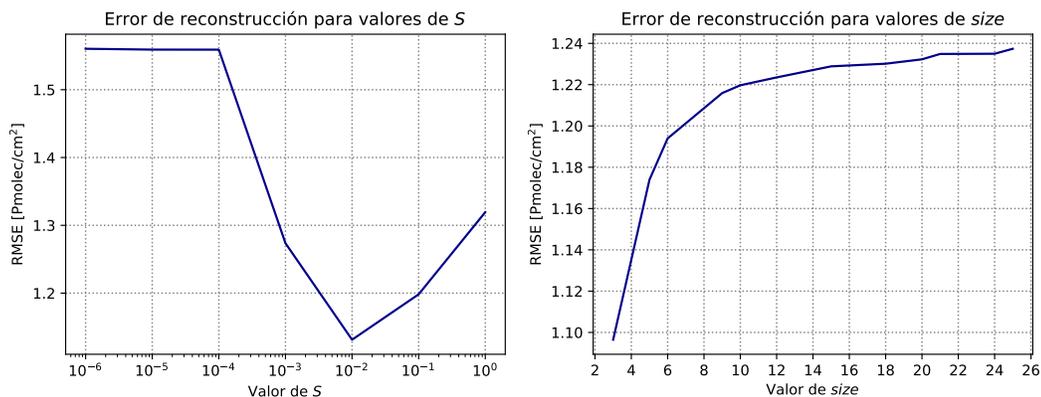


Figura 3.8: Selección de parámetros para función de llenado. (a) Error de reconstrucción para valores de **s**. (b) Error de reconstrucción para valores de **size**.

Una vez realizado el filtrado por calidad, el regrillado y el llenado de datos se procede a filtrar los datos a la zona de interés en la latitud y longitud especificadas anteriormente.

3.3.3. Generación de ventanas

Tras el regrillado y filtrado de los datos según condiciones de calidad, queda una grilla regular con una mayor cantidad de datos faltantes. Como el objetivo final es entrenar redes neuronales para realizar una mejora espacial de los datos, es necesario disponer de información sin datos faltantes. Para resolver este problema y disponibilizar de información para posteriormente entrenar una red se obtienen ventanas de datos. Estas ventanas consisten en secciones de 128×128 celdas de las grillas regulares ya procesadas en las cuales todos los datos de las celdas seleccionadas están disponibles. De esta forma se realiza una ventana deslizante que obtiene las ventanas del tamaño especificado, de lo contrario sigue buscando por la siguiente. Al encontrar una ventana se busca la siguiente saltando N celdas lo cual

genera un traslape de $M = 128 - N$ celdas, similar al caso especificado en la figura 3.9 . Esto asegura que al poder controlar la distancia de muestreo N , el traslape de ventanas M no sea tan grande y aporte a la heterogeneidad de los datos en caso de que exista una zona con muchas celdas sin datos faltantes.

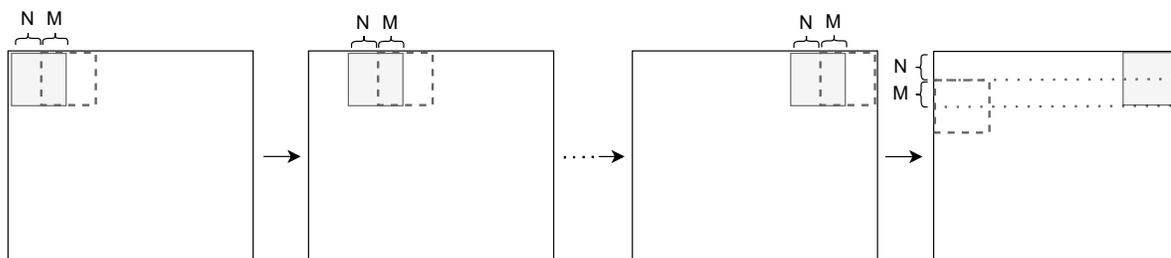


Figura 3.9: Diagrama explicativo de la generación de ventanas de datos.

Esta metodología permite obtener información de forma flexible ya que recorre la grilla buscando solamente secciones del tamaño especificado que tengan la completitud de sus datos. Luego, de cada grilla regular procesada es posible obtener múltiples ventanas que posteriormente conformarán el conjunto de datos con el cual se entrenarán y evaluarán las redes a utilizar. Cada ventana de 128×128 corresponde a una imagen en alta resolución (HR) que serán los datos a predecir. Una vez obtenidas las ventanas de HR se procede a obtener las de baja resolución (LR), donde por cada ventana HR se realiza un promedio de vecindad de píxeles para obtener una ventana de tamaño 64×64 . Con esto, cada celda de LR equivale a 4 celdas de HR. Debido a la similitud entre las ventanas generadas con imágenes, se usarán los conceptos píxel y celda de forma equivalente, haciendo una analogía con los trabajos de SR que utilizan imágenes.

Un potencial problema al entrenar una red es que las ventanas ingresadas no aporten información útil para el objetivo final de reconstruir datos de OMI. Una forma de abordar esta problemática es filtrar los datos que aportan poca información, y el criterio para su selección fue el filtrado por varianza de las ventanas. A cada ventana HR se le calculó su varianza, y la mediana de dichas varianzas fue utilizada como valor umbral para determinar el mínimo valor de varianza permitido para que una ventana sea considerada dentro del conjunto de datos.

3.3.4. Reemplazo de valores negativos

Como se pudo ver en la figura 3.5, los datos tienen presencia de valores negativos. Al trabajar con datos de columnas verticales de densidad de HCHO, los valores menores a 0 no tienen un sentido físico, por lo que se pueden atribuir directamente a ruido en las mediciones del espectrómetro. Una opción para evitar su presencia es realizar un filtrado directo de los datos negativos. Considerando que los datos negativos corresponden a un 5% de los datos totales, esta opción limitaría la cantidad de ventanas disponibles para el entrenamiento, ya que aumentaría la cantidad de datos faltantes (debido al filtrado), por lo que bastaría que falte un píxel de la ventana para que esta ya no sea considerada.

Para mitigar este problema sin afectar la disponibilidad de datos, se reemplazaron los datos negativos en base a los valores de su vecindad. Cada valor negativo presente en los datos fue ignorado y en su lugar se utilizó el valor promedio de la vecindad (8 celdas más cercanas) considerando solo los valores válidos, es decir, dejando fuera los datos negativos y faltantes (ver figura 3.10).

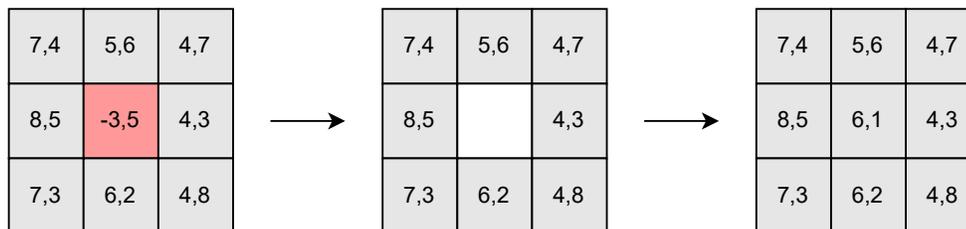


Figura 3.10: Diagrama explicativo de reemplazo de datos negativos.

Para los casos en los que los datos sean muy pequeños (bajo un valor umbral) o no fue posible obtener un valor promedio debido a que la vecindad solo tuvo datos faltantes, el valor del píxel es fijado como un valor base. Este valor base o umbral corresponde a $2,36$ [$\text{Pmolec}/\text{cm}^2$], y fue obtenido al calcular el promedio sobre valores en el Océano Pacífico, sector donde se espera que la columna vertical de densidad alcance su valor más bajo. De esta forma se logra reemplazar los datos negativos sin afectar la disponibilidad de información para la obtención de ventanas.

Capítulo 4

Implementación

En este capítulo se detalla la implementación de las redes neuronales utilizadas, las características específicas de cada una y las razones que justifican su elección. Los datos empleados corresponden a ventanas de HCHO de TROPOMI, obtenidos siguiendo el procedimiento detallado en el capítulo anterior.

4.1. Arquitecturas

El tipo de redes consideradas son totalmente convolucionales, es decir, se basan en convoluciones como operación principal en sus capas. Estas redes tienen la ventaja de ser agnósticas al tamaño de los datos de entrada que deben predecir. Gracias a esta característica, es posible entrenar una red utilizando datos de ventanas locales obtenidas y luego realizar una predicción a lo largo de una grilla de datos de mayor tamaño [8]. De esta manera, se resuelve el problema de los datos faltantes distribuidos a través de los archivos, ya que el conjunto de entrenamiento se obtiene solo de los sectores con datos disponibles, sin limitar el tamaño de los datos a predecir.

La primera red es una Super Resolution Convolutional Neural Network (SRCNN) planteada por Dong et al. [7]. Esta red marcó el comienzo de la mejora de las redes sobre métodos tradicionales de reconstrucción, como el algoritmo bicúbico de interpolación, que no utiliza machine learning. La SRCNN se destaca por su simplicidad y rapidez de entrenamiento, y su objetivo principal es validar la posibilidad de mejorar la resolución espacial de los datos, sin enfocarse excesivamente en una gran capacidad predictiva. Por esta razón, la red consta solo de tres capas convolucionales con tamaños variantes de canales y filtros, análoga a la red original en la que se basa. El preprocesamiento de esta red incluye una interpolación bicúbica de cada imagen LR, aumentando su tamaño de 64×64 a 128×128 (Up Sampling). De esta manera, la imagen se amplía al tamaño deseado antes que ingrese a la red, y luego, al pasar por esta se mejora su resolución (ver figura 4.1).

La segunda red a entrenar es una red generativa adversaria (GAN) ideada por Ledig et al. [18]. Esta consta de dos redes neuronales: una generadora G , que produce los resul-

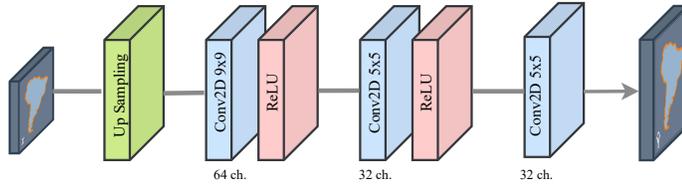


Figura 4.1: Arquitectura de red SRCNN junto al proceso de interpolación

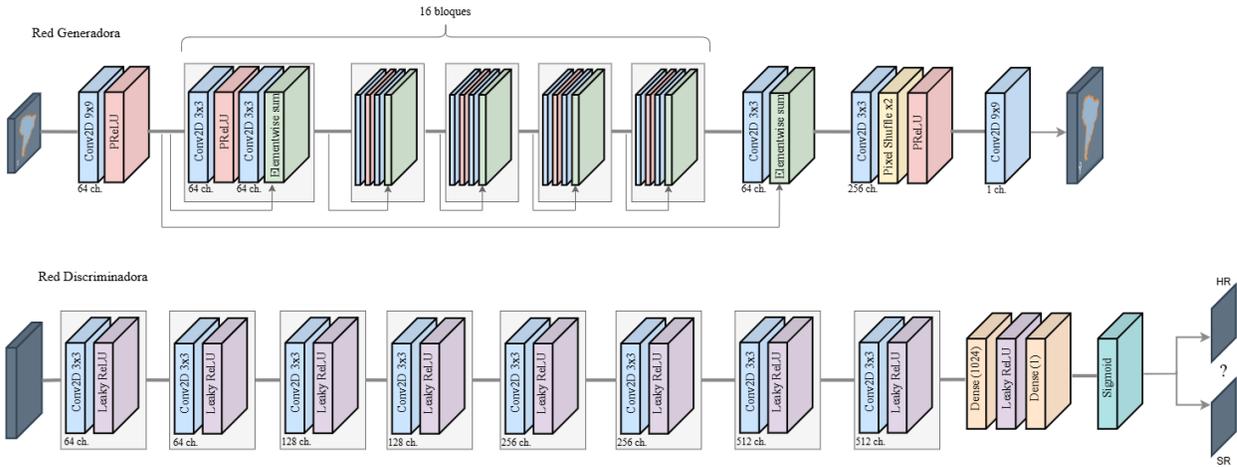


Figura 4.2: Arquitectura de la red adversaria. a) Red Generadora. b) Red Discriminadora.

tados de la reconstrucción espacial, y una discriminadora D , que distingue si las ventanas de tamaño 128×128 ingresadas son producidas por G o provienen del conjunto de datos de entrenamiento. Anteriormente, se han obtenido buenos resultados con GANs para problemas de súper resolución [18, 8, 25], proporcionando un método para insertar detalles físicamente realistas. Las dos redes se entrenan iterativamente en contra, de manera que el generador crea soluciones muy similares al conjunto de imágenes de entrenamiento, dificultando que el discriminador distinga entre las imágenes generadas y las reales.

La arquitectura de ambas redes se basa en una estructura de bloques. Para el generador, se utilizan 16 bloques en los que cada uno realiza dos convoluciones y luego suma la salida con los elementos de la entrada del bloque, en un proceso denominado *skip-connections*. Las *skip-connections* liberan a la red de la necesidad de modelar la función identidad, dado que su representación como resultado de convoluciones no es trivial. Además, en sus últimas capas, la red generadora incluye una convolución con $64 \times 2^2 = 256$ canales y una capa de *pixel shuffle* (reordena los datos según un factor) con un factor de aumento de 2, permitiendo así el aumento de tamaño de 64×64 a 128×128 . De este modo, la red generadora realiza el aumento de tamaño de la imagen de entrada a lo largo de la misma red, sin necesidad de un preprocesamiento, a diferencia de la SRCNN.

La red discriminadora, por su parte, recibe una imagen aumentada a 128×128 y determina si esta es real (pertenece al conjunto de entrenamiento) o falsa (fue generada por la red generadora). Su arquitectura consta de 8 bloques convolucionales con una función de activación Leaky ReLU con $\alpha = 0,2$ [18], donde la cantidad de canales en la convolución aumenta según la profundidad de la red. Posteriormente, la red dispone capas lineales de

conexión que culminan en una función sigmoide, asignando un valor cercano a 1 si considera que la entrada pertenece al conjunto de entrenamiento y cercano a 0 si fue producida por el generador.

La principal modificación de la red propuesta por Ledig et al. [18] es la eliminación de las capas de *Batch Normalization*, una operación que normaliza la salida de la capa anterior restando la media y dividiendo por la desviación estándar. Estas capas pueden restringir el rango de flexibilidad de la red; por ello, se ha demostrado experimentalmente que eliminarlas mejora el rendimiento del modelo [20, 8].

4.2. Entrenamiento

Como se mencionó anteriormente, el conjunto de datos fue formado por ventanas obtenidas de TROPOMI, donde las de baja resolución (LR, denominadas también x) se obtuvieron de las de alta resolución (HR, denominadas también y). Del conjunto de datos, se reservó una fracción del 20 % para validación durante el entrenamiento y otro 20 % para disponer de un conjunto de pruebas. La red SRCNN (ahora denominada también F) sigue un proceso de entrenamiento en el cual recibe datos de entrada x y los procesa de forma iterativa. El procesamiento genera predicciones $F(x)$, que se comparan con el valor real y del conjunto de entrenamiento. La función de pérdida computa el error de predicción y busca ser minimizada durante el entrenamiento para que los resultados generados se acerquen cada vez más a los datos reales de entrenamiento. Utilizando el algoritmo de *backpropagation* [36] y el optimizador *Adam* [37], que computa *learning rates* adaptativos individuales basados en estimaciones adaptativas de momentos de primer y segundo orden y propaga el error a través de la red.

El objetivo del entrenamiento es minimizar la función de pérdida, que determina el error de cada predicción y guía el aprendizaje de la red; por esta misma razón es de gran relevancia determinar una función que se adapte al problema planteado. Para la SRCNN, se usa la función de pérdida L2 normalizada, también conocida como error cuadrático medio (MSE), que mide la diferencia cuadrática normalizada entre las celdas de $F(x)$ y y .

En cuanto al entrenamiento de la red adversaria (SRGAN), se requiere tener un enfoque diferente, ya que implica entrenar dos redes en conjunto. Inicialmente, el dato de entrada x se procesa a través de la red generadora G , obteniendo su versión aumentada $G(x)$. Luego, la red discriminadora D evalúa tanto el valor generado por la red generadora $D(G(x))$ como el valor real $D(y)$ del conjunto de datos de entrenamiento. Se espera que la red D asigne un valor bajo a $D(G(x))$ porque reconoce que proviene de la red generadora, mientras que debe asignar un valor alto a $D(y)$ debido a que proviene del conjunto de datos de reales.

Cada una de estas redes tiene su propia función de pérdida: \mathcal{L}_G para la red generadora y \mathcal{L}_D para la red discriminadora. La función de pérdida de la red de generadora \mathcal{L}_G tiene dos componentes. La primera, denominada de contenido $\mathcal{L}_{\text{contenido}}$ (ecuación 4.1), realiza una comparación directa entre $G(x)$ y el dato real y . La segunda, la componente adversaria $\mathcal{L}_{\text{adversaria}}$ (ecuación 4.2), mide qué tan bien la red generadora puede engañar a la red discriminadora D . Así, la función de pérdida del generador se define según la ecuación 4.3, con un valor de $\alpha = 10^{-3}$ [8, 18].

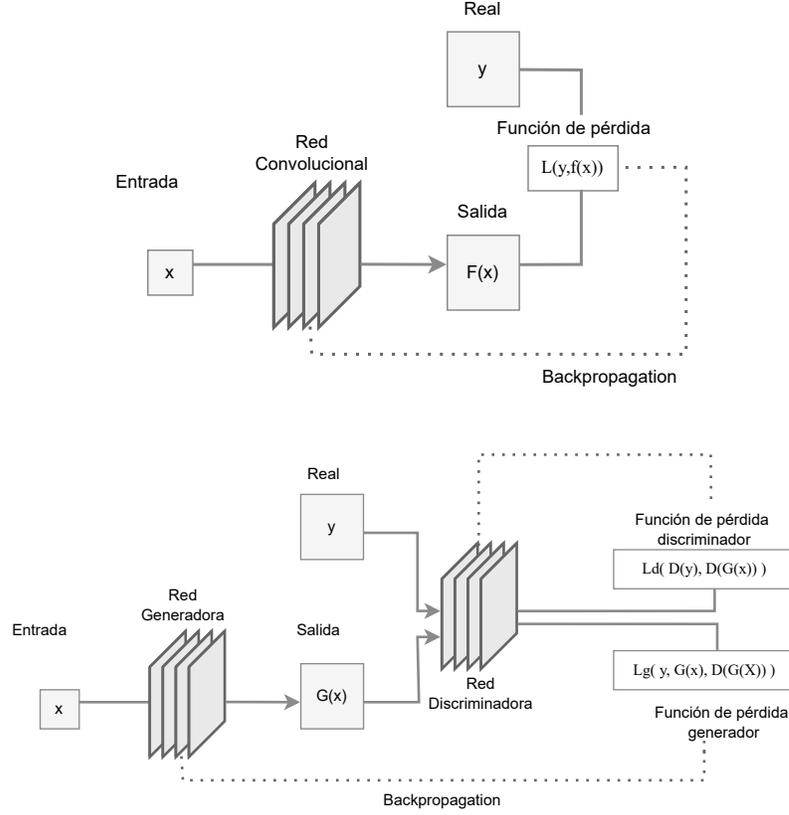


Figura 4.3: Diagrama de entrenamiento de redes. a) Entrenamiento de red convolucional. b) Entrenamiento de red adversaria.

$$\mathcal{L}_{\text{contenido}}(x, y) = \frac{1}{N} \sum (y - G(x))^2 \quad (4.1)$$

$$\mathcal{L}_{\text{adversaria}} = -\log(D(G(x))) \quad (4.2)$$

$$\mathcal{L}_G = \mathcal{L}_{\text{contenido}} + \alpha \mathcal{L}_{\text{adversaria}} \quad (4.3)$$

Con respecto a la función de pérdida \mathcal{L}_D , en 4.4, el objetivo es evaluar la capacidad de la red discriminadora para distinguir entre datos reales del conjunto de entrenamiento y , asignándoles alto puntaje, y datos generados por el generador $G(x)$, a los cuales se les asigna un puntaje bajo. Esto permite a la red discriminadora aprender a discernir eficazmente entre los datos auténticos y generados.

$$\mathcal{L}_D = -\log(D(y)) - \log(1 - D(G(x))) \quad (4.4)$$

El entrenamiento de la red adversaria se compone de dos fases: una de pre-entrenamiento y una de entrenamiento adversario. En la fase de pre-entrenamiento, solo la red generadora aprende de manera similar a la SRCNN, sin la participación del discriminador. Este mecanismo permite que la red generadora mejore sus predicciones antes de enfrentarse al

discriminador. Durante esta etapa, la función de pérdida asociada para el generador es la misma que la ecuación 4.3, pero sin la componente adversaria, es decir, con $\alpha = 0$. En la segunda fase, se lleva a cabo el entrenamiento adversario, donde ambas redes (generadora y discriminadora) interactúan. Aquí se utiliza un valor de $\alpha = 10^{-3}$ en la función de pérdida del generador \mathcal{L}_G basado en la literatura [8, 18], lo que incluye la componente adversaria. Este proceso iterativo permite que la red generadora envíe predicciones más precisas al discriminador, y a su vez, que el discriminador refine su capacidad de distinguir entre datos reales y generados. El proceso de entrenamiento de las redes mencionadas se ilustra en la figura 4.3 y se llevó a cabo utilizando el framework Pytorch, una librería de aprendizaje de máquinas que facilita la creación y entrenamiento de modelos [38].

Capítulo 5

Resultados y análisis

En este capítulo se presentan los resultados de los estudios realizados utilizando las arquitecturas de redes especificadas en los capítulos anteriores. Para ello, se obtuvieron dos conjuntos de datos con la metodología de ventanas variando el parámetro de distancia de muestreo entre ventanas (parámetro N de la figura 3.9) con el objetivo de mostrar los cambios que esto produce en los resultados de las redes utilizadas. Cabe destacar que ambos conjuntos provienen de los mismos datos de TROPOMI¹, y solo varía el parámetro de distancia de muestreo en la obtención de ventanas. Para cada conjunto se obtuvo un subconjunto de prueba, correspondiente al 20 % de las ventanas generadas para la realización de pruebas. En cada caso de estudio, se realizó una comparación de los modelos entrenados utilizando como referencia el método de interpolación bicúbica, el cual realiza reconstrucciones sin el uso de aprendizaje de máquinas.

Los conjuntos de datos a evaluar se diferencian por tener un traslape de distinta cantidad de celdas al momento de la extracción de ventanas. Esto puede significar una diferencia en la calidad de los datos con los que la red aprende. Se espera que mientras menor sea el traslape las ventanas estarán espacialmente más distantes, por ende los datos tendrán menor posibilidad de repetirse y existirá una mayor heterogeneidad. Las redes convolucionales especificadas basan su aprendizaje en la operación de convolución, que actúa sobre vecindades de valores y no en píxeles individuales, por lo que evitar la repetición de áreas de datos espacialmente cercanos puede aumentar la calidad de la reconstrucción de la red. Por otra parte, se espera que mientras mayor sea la disponibilidad de datos, mejor sean los resultados obtenidos. Por ende formar un conjunto de datos con un espaciado excesivo entre las ventanas extraídas puede limitar la capacidad predictiva de la red. El primer conjunto de datos a evaluar considera una distancia entre ventanas contiguas de 32 celdas, correspondiente a la mitad de la dimensión de una ventana de baja resolución. El segundo conjunto de datos fue obtenido con una distancia entre ventanas de 14 celdas, similar al utilizado en el trabajo de Dong et al. [7].

Para tratar los valores negativos, se utilizó la metodología descrita en la sección 3.3.4 la cual fue aplicada para ambos conjuntos de datos mencionados. De esta forma se obtuvo

¹Los datos utilizados se encuentran disponibles públicamente en el repositorio NASA GES DISC para OMI y TROPOMI.

resultados para ambos conjuntos de datos (muestreo cada 32 y 14 celdas) considerando para cada uno el caso en el que presenta datos negativos, y también el caso en el que los datos negativos fueron removidos (reemplazados por el promedio de la vecindad). Esto permite evaluar el efecto que tienen los valores no físicos en la red y cuantificar el impacto de su presencia en el entrenamiento en base a las métricas de evaluación, y verificar si el preprocesamiento de los valores negativos es un aporte al aprendizaje de las redes.

Este trabajo fue acelerado con el uso de una Unidad de Procesamiento Gráfico (*Graphical Processing Unit*, GPU en inglés), la cual ofrece varias ventajas. Las GPUs están diseñadas para realizar cálculos en paralelo, lo que acelera significativamente el procesamiento de grandes volúmenes de datos y el entrenamiento de modelos. Esto permite reducir considerablemente los tiempos de entrenamiento en comparación con el uso de CPUs. Además, las GPUs son especialmente eficaces para manejar las operaciones matriciales intensivas que son comunes en los algoritmos de aprendizaje profundo, mejorando así la eficiencia y el rendimiento general del proceso de entrenamiento.

Cada conjunto de datos dispone de su propio subconjunto de entrenamiento con el cual la red aprende y de un subconjunto de prueba para obtener las métricas que evalúan la calidad de la reconstrucción. La evaluación sobre cada conjunto fue realizada con los modelos entrenados específicamente con dichos datos, es decir, las redes fueron entrenadas de forma separada para cada conjunto de datos. Con ello, se dispone de cuatro redes SRCNN y cuatro redes SRGAN, cada una entrenada con un conjunto de datos distinto y evaluada sobre el mismo conjunto (es decir, para el conjunto de muestreo 14 con y sin valores negativos, y para el conjunto de muestreo 32 con y sin valores negativos). Cabe mencionar que las redes mantienen la arquitectura antes especificada, y solo varía el conjunto de datos utilizado en cada una.

5.1. Conjuntos con valores negativos

Inicialmente se realiza la evaluación de los datos considerando los datos negativos, es decir, sin el preprocesamiento de valores de vecindad. En la tabla 5.1, se presentan los tiempos de entrenamiento obtenidos utilizando una GPU NVidia Tesla V100 para las arquitecturas de redes implementadas, para cada conjunto de datos descrito anteriormente.

Tabla 5.1: Tiempos de entrenamiento de los modelos SRCNN y SRGAN para casos con datos negativos.

Tiempo de entrenamiento [hr]	SRCNN	SRGAN
Primer conjunto de datos	3,1	19,25
Segundo conjunto de datos	14,13	94,18

5.1.1. Primer conjunto de datos

Como se mencionó, el primer conjunto de datos a evaluar considera una distancia de muestreo de 32 celdas entre ventanas ($N = 32$ en figura 3.9). Se evaluaron los modelos sobre

el conjunto de prueba, con lo cual se obtuvieron las métricas especificadas en la tabla 5.2. Para PSNR ambas redes logran un mejor rendimiento que el caso base de comparación, indicando que logran tener un menor error MSE entre píxeles. En cuanto a SSIM la diferencia entre modelos es aún mayor, indicio de que la similaridad estructural de las imágenes reconstruidas con respecto a las originales es mayor para los modelos entrenados. Finalmente MAPE indica un error porcentual de las predicciones, donde nuevamente las redes entregaron un rendimiento superior a la interpolación bicúbica, pero se puede considerar que 68 % o 71 % de error porcentual es relativamente alto. Para todas estas métricas el rendimiento de la red adversaria fue superior, destacando con un mejor rendimiento en todos los aspectos de precisión utilizados en este trabajo.

Tabla 5.2: Métricas promedio obtenidas de los modelos evaluados para el primer conjunto de datos.

Métrica / Modelo	Bicúbico	SRCNN	SRGAN
PSNR [dB]	28,92	29,23	29,57
SSIM	0,74	0,76	0,79
MAPE [%]	72,61	71,0	68,86

Es importante comprender cómo se distribuyen estas métricas a lo largo del conjunto de datos sobre el cual fueron evaluadas. Es por eso que se presentan en la figura 5.1 los gráficos de distribución acumulada para las redes entrenadas y el método de interpolación bicúbica. En las tres comparaciones se puede notar que el rendimiento más bajo corresponde a la interpolación bicúbica, mejorada por la red SRCNN para finalmente ser ambas superadas por la red SRGAN. Esta tendencia se mantiene a lo largo de todo el conjunto de prueba, donde para cada imagen evaluada se puede esperar que los modelos entrenados obtengan mejor rendimiento. En donde la diferencia se acrecienta es en la curva de SSIM, en la cual ambas redes se desenvuelven notoriamente mejor que la interpolación.

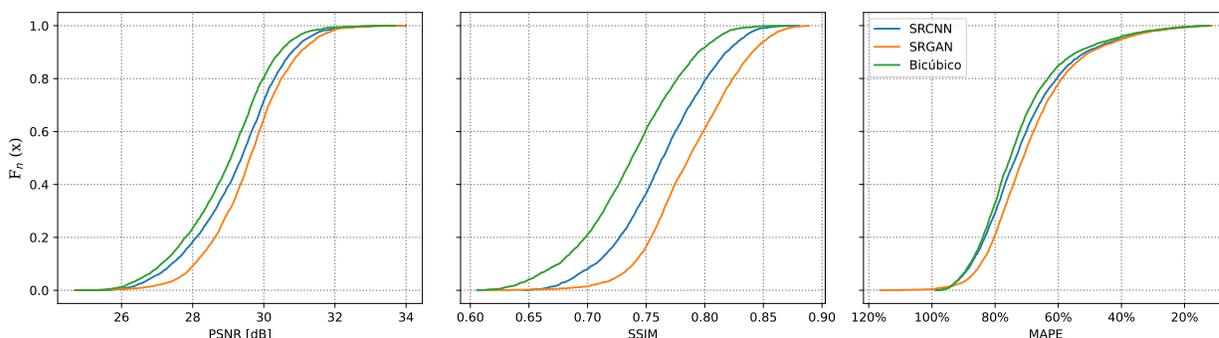


Figura 5.1: Distribución acumulada de modelos para el conjunto de prueba del primer conjunto de datos. Mayor PSNR y SSIM es mejor, mientras que menor MAPE es mejor.

Otra evaluación realizada fue comparar la correlación obtenida en los datos esperados con los obtenidos. Para ello se presentan gráficos de dispersión en la figura 5.2, donde cada píxel de una ventana predicha corresponde a un punto en el gráfico. El comportamiento deseado es que los puntos se encuentren lo más cercanos a la línea diagonal, indicando que la reconstrucción de su valor fue precisa. Como se puede esperar, la red adversaria obtuvo

mejor rendimiento que los otros casos al tener mayor coeficiente R^2 , aunque la mejora es leve con respecto al modelo SRCNN. Finalmente, ambas redes logran tener un ajuste superior al caso bicúbico.

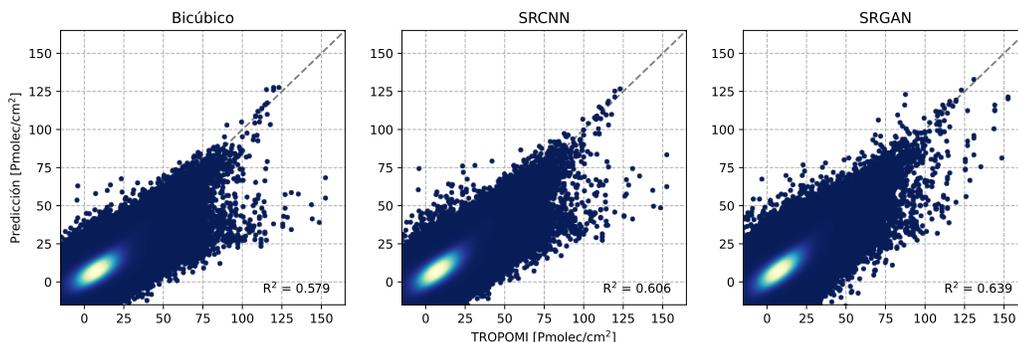


Figura 5.2: Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados para el primer conjunto de datos.

Para evaluar el impacto de los datos negativos en la predicción se realizó un gráfico de dispersión filtrando los datos menores que 0 en la figura 5.3. El primer resultado a destacar es que la correlación para los tres casos disminuye considerablemente según el coeficiente R^2 con respecto al caso sin filtrado. Esto sugiere que parte de la capacidad del ajuste de los modelos se enfoca en la predicción de valores negativos, lo cual como ya se mencionó, no presenta utilidad para este trabajo.

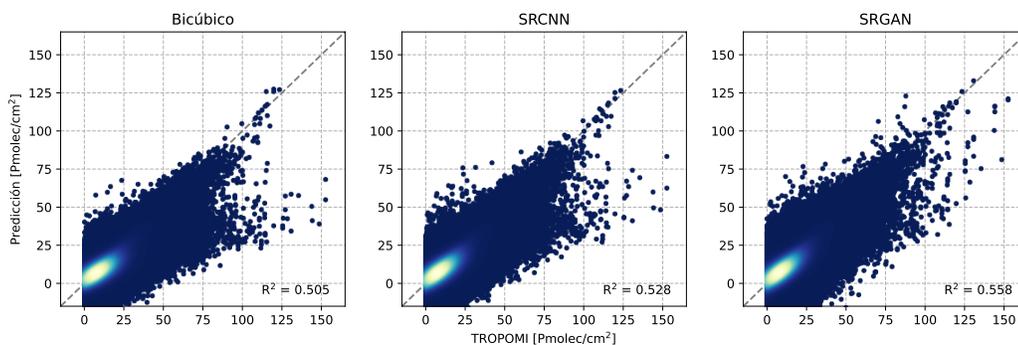


Figura 5.3: Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados con negativos filtrados para el primer conjunto de datos.

Con el fin de obtener una perspectiva visual de la calidad de reconstrucción se grafican las mejoras en resolución de los modelos en la figura 5.4. La primera ventana corresponde a píxeles en baja resolución, la cual se busca aumentar a través de los distintos modelos. Luego, se muestra la reconstrucción bicúbica, la de SRCNN y la de SRGAN, presentando ventanas en alta resolución. Finalmente se encuentra la ventana con los datos provenientes del subconjunto de prueba la cual se busca replicar a través de los modelos.

De manera similar, se presenta la figura 5.5 con las reconstrucciones señaladas anteriormente junto al error absoluto píxel a píxel de las dos redes utilizadas. De esta forma se puede identificar visualmente si existe a algún patrón en las diferencias de concentración entre las predicciones y los valores esperados. En particular, se observa que la presencia de secciones

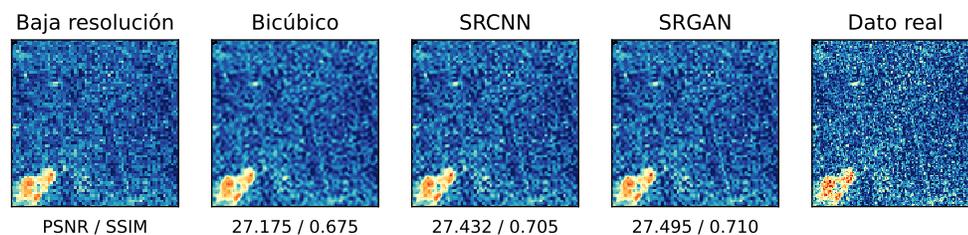


Figura 5.4: Gráfico de predicciones para modelos para el primer conjunto de datos considerando valores negativos. a) Dato a predecir en baja resolución. b) Reconstrucción de modelo bicúbico. c) Reconstrucción de modelo SRCNN. d) Reconstrucción de modelo SRGAN. e) Dato real de alta resolución TROPOMI.

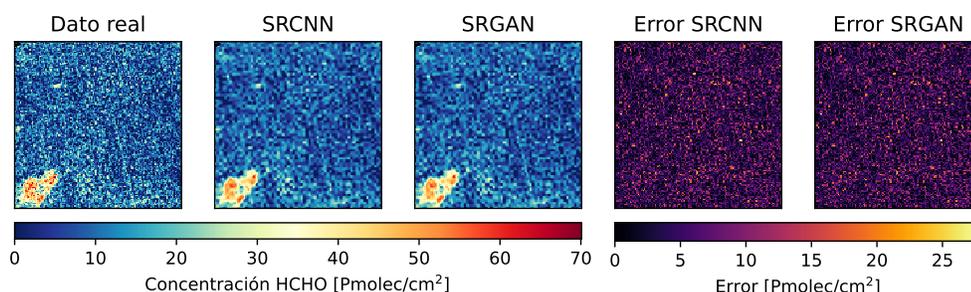


Figura 5.5: Gráfico de medición de error para el primer conjunto de datos considerando valores negativos. a) Dato real de alta resolución TROPOMI. b) Reconstrucción de modelo SRCNN, c) Reconstrucción de modelo SRGAN. d) Error de predicción de SRCNN. e) Error de predicción de SRGAN.

con valores altos a predecir no es un indicador directo de un mayor error en la reconstrucción para este caso. Las diferencias representadas no parecen tener alguna distribución evidente.

5.1.2. Segundo conjunto de datos

El segundo conjunto de datos presenta una distancia de muestreo de 14 celdas ($N = 14$ en figura 3.9), por lo cual se dispone de una cantidad mayor de ventanas para entrenar ambas redes. Las métricas obtenidas se presentan en la tabla 5.3 donde nuevamente SRGAN obtuvo un mejor rendimiento y SRCNN continuó presentando una mejora con respecto al caso bicúbico. Es directo notar que existe una mejora en comparación al conjunto de datos anterior, entregando valores de métricas superiores. Aún así, cabe destacar que la obtención de un error porcentual promedio del 66% en el mejor caso continúa siendo alto.

Para estudiar la distribución de las métricas de reconstrucción se obtuvo la figura 5.3. En ella, se observa que existe una gran diferencia de precisión obtenida entre modelos a lo largo del conjunto de datos si se compara con el conjunto de datos anterior. La red SRCNN sigue siendo levemente superior a la interpolación bicúbica en las métricas PSNR y MAPE, mientras que tiene una pequeña mejoría para SSIM, similar al obtenido para el conjunto de

Tabla 5.3: Métricas promedio obtenidas de los modelos evaluados para el segundo conjunto de datos.

Métrica / Modelo	Bicúbico	SRCNN	SRGAN
PSNR [dB]	30,7	31,01	31,68
SSIM	0,77	0,79	0,82
MAPE [%]	72,72	70,95	66,08

datos anterior. Por otra parte, la red adversaria mantiene su mejora con respecto a los otros dos modelos. Para este conjunto de datos SRGAN logró alcanzar métricas destacablemente mejores que las obtenidas por SRCNN y el caso bicúbico.

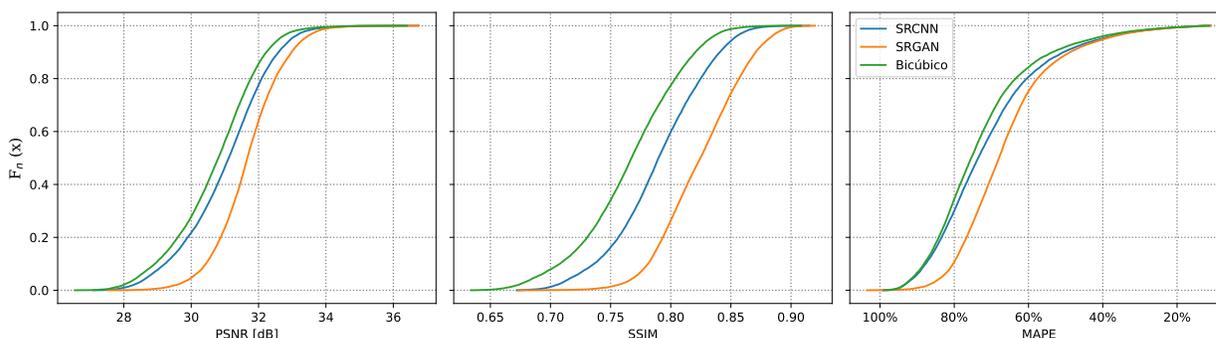


Figura 5.6: Distribución acumulada de modelos para el conjunto de prueba del segundo conjunto de datos. Mayor PSNR y SSIM es mejor, mientras que menor MAPE es mejor.

Al comparar las predicciones de los modelos con los valores esperados de TROPOMI se vuelve a notar la mejoría obtenida en las métricas. El coeficiente de determinación R^2 presenta un leve aumento con respecto en los casos bicúbico y SRCNN, mientras que para el caso adversario la mejora se acrecienta. Un aspecto interesante a notar es que los modelos tienden a subestimar los valores predichos, ya que estos se encuentran en su mayoría debajo de la diagonal de referencia. Esto es un indicador de que el valor predicho es menor al esperado.

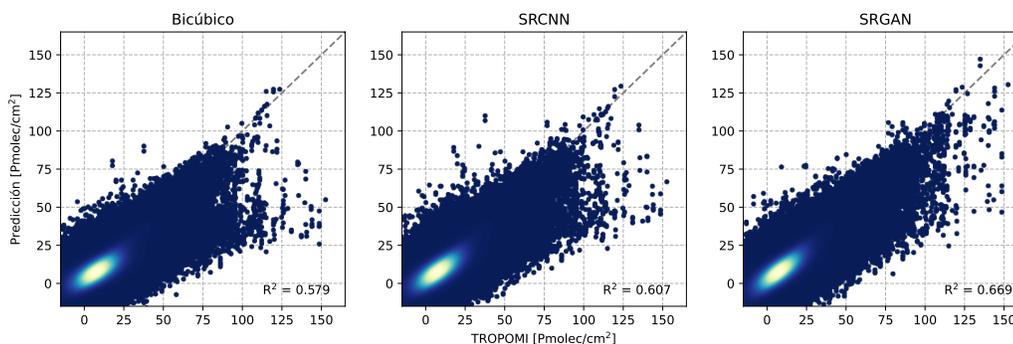


Figura 5.7: Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI para el segundo conjunto de datos.

Si se realiza el filtro de datos negativos a las predicciones individuales de los datos de TROPOMI el rendimiento de los modelos vuelve a empeorar. Como este comportamiento fue

obtenido con dos conjuntos de datos puede indicar que las redes podrían beneficiarse de un preprocesamiento en sus valores negativos, de manera que las redes mejoren su aprendizaje mejorando el error obtenido en los datos positivos a predecir y quitar énfasis en los negativos, que no son un aporte para el objetivo de la mejora de resolución.

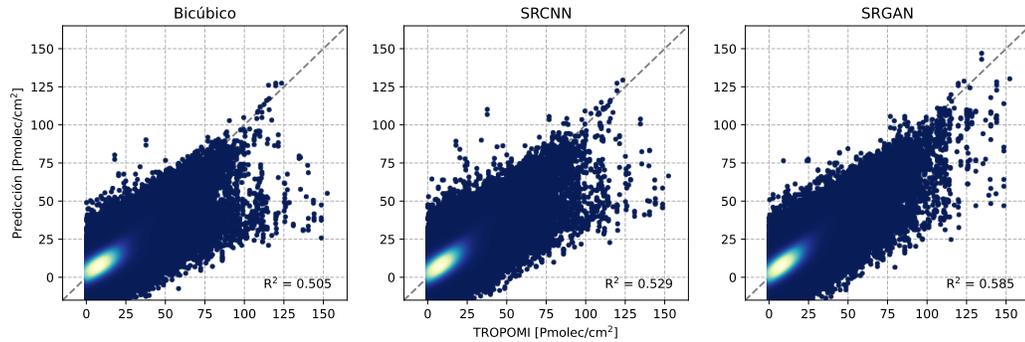


Figura 5.8: Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados con negativos filtrados para el segundo conjunto de datos.

Análogo al conjunto de datos anterior, se evalúa una ventana de concentraciones a través de los distintos modelos en la figura 5.9, entrenados esta vez con el segundo conjunto de datos. Perceptualmente, todos los modelos logran captar la alta concentración y aislarla de las valores bajos, entregando resultados visiblemente similares.

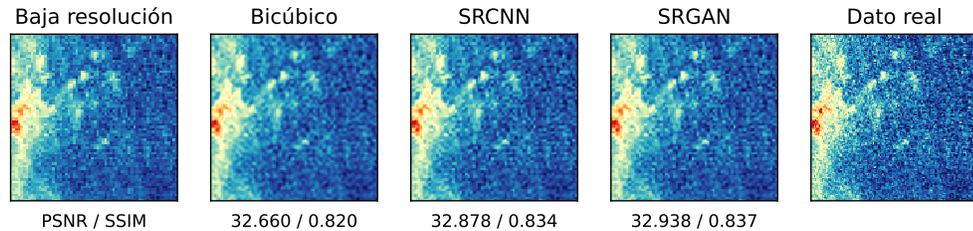


Figura 5.9: Gráfico de predicciones para modelos para el segundo conjunto de datos considerando valores negativos. a) Dato a predecir en baja resolución. b) Reconstrucción de modelo bicúbico. c) Reconstrucción de modelo SRCNN. d) Reconstrucción de modelo SRGAN. e) Dato real de alta resolución TROPOMI.

Luego, pasando al error obtenido en esta predicción, similar a lo ocurrido en el conjunto de datos anterior no se presenta un patrón de error reconocible en sectores de alta o de baja concentración (ver figura 5.10).

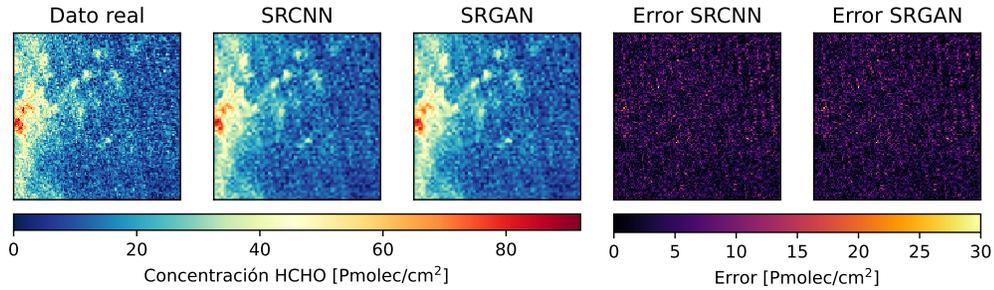


Figura 5.10: Gráfico de medición de error para el segundo conjunto de datos considerando valores negativos. a) Dato real de alta resolución TROPOMI. b) Reconstrucción de modelo SRCNN. c) Reconstrucción de modelo SRGAN. d) Error de predicción de SRCNN. e) Error de predicción de SRGAN.

5.2. Conjuntos sin valores negativos

Ahora se consideran los datos para los que los datos negativos fueron preprocesados. En la tabla 5.4, se presentan los tiempos de entrenamiento obtenidos utilizando la misma GPU NVidia Tesla V100 para las arquitecturas de redes implementadas, para cada conjunto de datos en los cuales no hay valores negativos.

Tabla 5.4: Tiempos de entrenamiento de los modelos SRCNN y SRGAN para casos con valores negativos preprocesados.

Tiempo de entrenamiento [hr]	SRCNN	SRGAN
Primer conjunto de datos	3,15	20,78
Segundo conjunto de datos	14,28	89,55

5.2.1. Primer conjunto de datos sin valores negativos

Análogo al caso anterior, este conjunto de datos considera las mismas 32 celdas como distancia de muestreo entre ventanas. Los modelos fueron evaluados sobre su conjunto de pruebas correspondiente, obteniendo las métricas de la tabla 5.5. En ella es posible notar que el rendimiento de PSNR y MAPE mejoró considerablemente con respecto a los casos que sí utilizan datos negativos. En cuanto a SSIM se obtuvo una pequeña mejora con respecto al caso del primer conjunto de datos con valores negativos. En particular el error porcentual medio bajó del orden de 68 % (tabla 5.2) en el conjunto de datos con valores negativos del mismo parámetro de muestreo a un 42 % para este conjunto que no considera valores negativos, lo cual es una mejora considerable.

La distribución de las métricas se presenta en la figura 5.11. En ella se puede ver que las curvas PSNR y SSIM se comportan de forma similar y están relativamente cercanas, similar al caso con valores negativos de la misma distancia de muestreo. La gran diferencia que presenta con dicho caso es que las curvas están desplazadas horizontalmente hacia mayores

Tabla 5.5: Métricas promedio obtenidas de los modelos evaluados para el primer conjunto de datos con valores negativos preprocesados.

Métrica / Modelo	Bicúbico	SRCNN	SRGAN
PSNR [dB]	33,52	33,84	34,05
SSIM	0,77	0,79	0,81
MAPE [%]	46,47	43,73	42,38

valores, reafirmando la noción del aumento de rendimiento. En cuanto a MAPE el comportamiento difiere respecto al caso con valores negativos, ya que las curvas SRCNN y SRGAN se comportan notablemente mejor que el caso bicúbico (3 %) mientras que para el caso de consideraba valores no físicos, el rendimiento destacado solo lo tenía SRGAN.

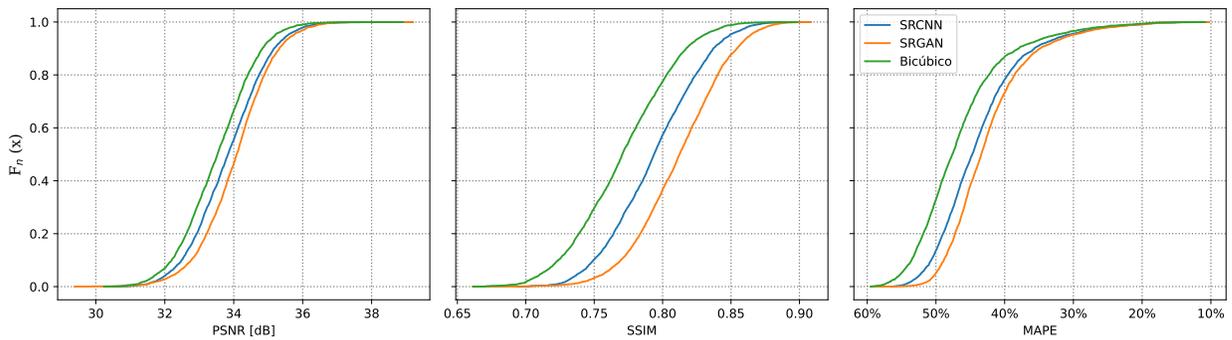


Figura 5.11: Distribución acumulada de modelos para el conjunto de prueba del primer conjunto de datos con valores negativos preprocesados. Mayor PSNR y SSIM es mejor, mientras que menor MAPE es mejor.

Las predicciones realizadas con los modelos entrenados con este conjunto de datos se presenta en la figura 5.12, donde se mantiene la tendencia en la que SRCNN obtiene mejor R^2 que la interpolación bicúbica, mientras que SRGAN es superior a ambas. Cabe destacar que si bien SRGAN tiene mayor coeficiente de determinación R^2 , visualmente presenta una mayor dispersión en sus valores predichos. Por otra parte, tanto el caso bicúbico como SRCNN tienen un sesgo negativo de los datos (tienden a subestimar el valor predicho), SRGAN presenta un leve sesgo positivo (sobrestima el valor predicho).

Como fue mencionado, en los conjuntos de datos que consideran valores no físicos se tiene que al filtrar dichos valores el rendimiento disminuye considerablemente. Por eso es importante notar que para el caso de la figura mencionada con preprocesamiento de valores negativos se tiene un rendimiento levemente mejor en R^2 con respecto al que sí considera valores negativos (ver figura 5.2), pero esta diferencia es acrecentada considerablemente al filtrar los valores negativos y dejar solo valores físicos (ver figura 5.3).

En la figura 5.13 se encuentran las predicciones de las ventanas predichas con este conjunto de datos para los distintos modelos junto a sus datos en alta y baja calidad. En la figura 5.14 se enseñan los errores píxel a píxel de las predicciones, en los que no se observan patrones reconocibles visiblemente en dichos errores.

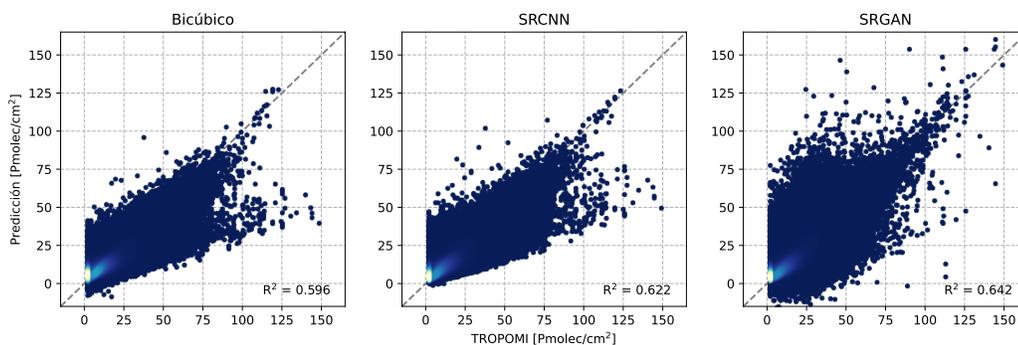


Figura 5.12: Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados para el primer conjunto de datos con valores negativos preprocesados.

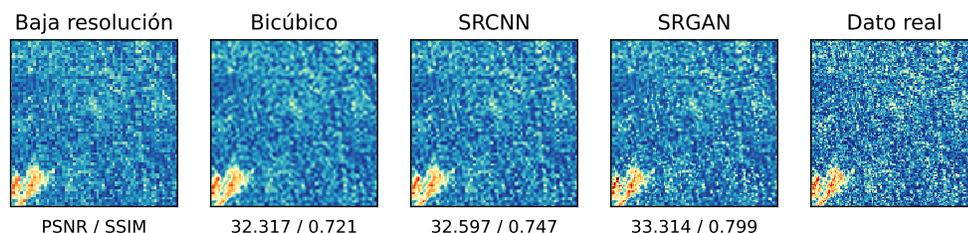


Figura 5.13: Gráfico de predicciones para modelos para el primer conjunto de datos con valores negativos preprocesados. a) Dato a predecir en baja resolución. b) Reconstrucción de modelo bicúbico. c) Reconstrucción de modelo SRCNN. d) Reconstrucción de modelo SRGAN. e) Dato real de alta resolución TROPOMI.

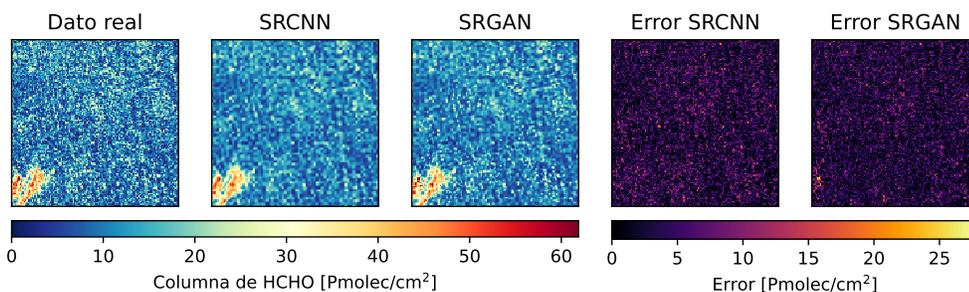


Figura 5.14: Gráfico de medición de error para el primer conjunto de datos con valores negativos preprocesados. a) Dato real de alta resolución TROPOMI. b) Reconstrucción de modelo SRCNN, c) Reconstrucción de modelo SRGAN. d) Error de predicción de SRCNN. e) Error de predicción de SRGAN.

5.2.2. Segundo conjunto de datos sin valores negativos

Finalmente, este conjunto de datos considera muestreo cada 14 celdas con preprocesamiento de valores negativos. Los modelos fueron evaluados en el conjunto de pruebas respectivo y las métricas resultantes se presentan en la tabla 5.6. Este caso obtuvo los mejores resultados para PSNR, SSIM y MAPE. En particular destaca el bajo valor de error porcentual absoluto

medio del orden de 40%, dado que para el caso que consideraba valores no físicos con el mismo muestreo esta métrica alcanzó valores del orden de 66% de error porcentual (tabla 5.3).

Tabla 5.6: Métricas promedio obtenidas de los modelos evaluados para el segundo conjunto de datos con valores negativos preprocesados.

Métrica / Modelo	Bicúbico	SRCNN	SRGAN
PSNR [dB]	33,49	33,81	34,52
SSIM	0,77	0,79	0,83
MAPE [%]	46,57	43,78	40,68

La distribución de las métricas de la figura 5.15 indica que el rendimiento de SRGAN fue consistentemente mejor que SRCNN y la interpolación bicúbica. La distribución que mantuvo un rendimiento notoriamente superior en su distribución fue las de SRGAN para la métrica SSIM.

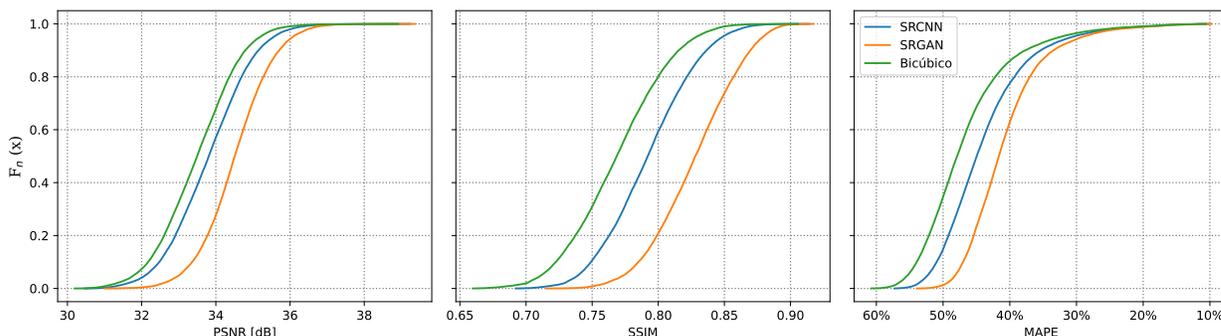


Figura 5.15: Distribución acumulada de modelos para el conjunto de prueba del segundo conjunto de datos con valores negativos preprocesados. Mayor PSNR y SSIM es mejor, mientras que menor MAPE es mejor.

Los gráficos de dispersión con su el valor de R^2 para este conjunto de datos se encuentran en la figura 5.16. Es importante notar que los resultados de SRCNN y caso bicúbico se mantuvieron similares al obtenido conjunto con un muestreo menor de datos sin valores negativos de la figura 5.12, pero destaca un mejor rendimiento en SRGAN, donde se puede notar visiblemente que los valores se encuentran menos dispersos, lo cual guarda directa relación con su valor mayor de R^2 .

En la figura 5.17 se muestran las predicciones de los modelos para este conjunto de datos, junto al puntaje de PSNR y SSIM obtenido en cada una. Por otra parte, en la figura 5.18 se incluye el error obtenido al comparar los píxeles obtenidos por los modelos con el valor esperado, donde nuevamente no existen distribuciones o patrones reconocibles de errores.

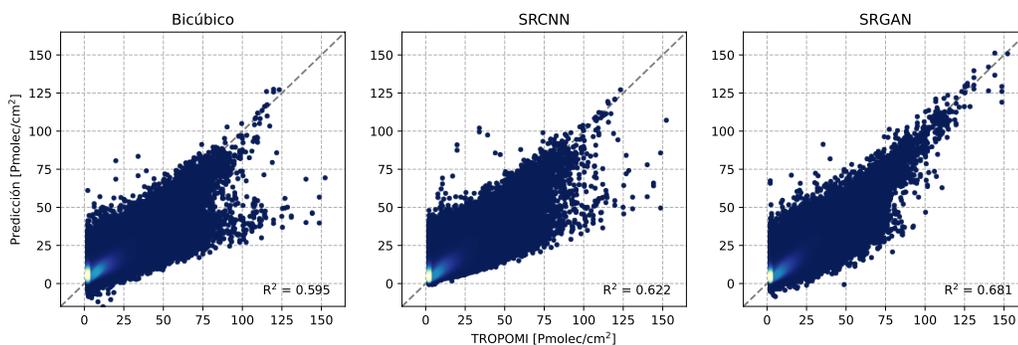


Figura 5.16: Gráfico de dispersión de valores predichos por los modelos y valores de TROPOMI observados para el segundo conjunto de datos con valores negativos preprocesados.

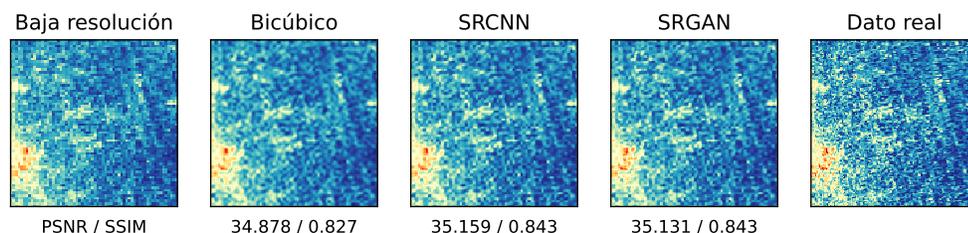


Figura 5.17: Gráfico de predicciones para modelos para el primer conjunto de datos con valores negativos preprocesados. a) Dato a predecir en baja resolución. b) Reconstrucción de modelo bicúbico. c) Reconstrucción de modelo SRCNN. d) Reconstrucción de modelo SRGAN. e) Dato real de alta resolución TROPOMI.

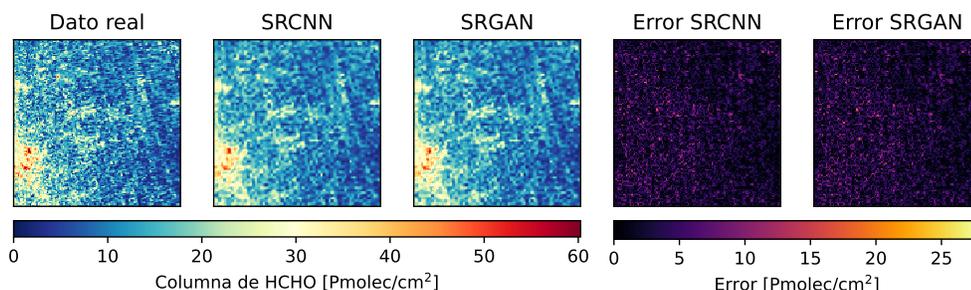


Figura 5.18: Gráfico de medición de error para el primer conjunto de datos con valores negativos preprocesados. a) Dato real de alta resolución TROPOMI. b) Reconstrucción de modelo SRCNN, c) Reconstrucción de modelo SRGAN. d) Error de predicción de SRCNN. e) Error de predicción de SRGAN.

5.3. Análisis de resultados

Para realizar una buena mejora en la resolución espacial primero es necesario entender cómo se comportan y distribuyen los datos espacial y temporalmente. Se comentó previamente que los conjuntos de datos destacan por la presencia valores negativos, los cuales no tienen

sentido físico para concentraciones de HCHO. Por ello se realizó un análisis para todos los casos entrenados, donde al filtrar los valores menores a 0 el ajuste de los datos empeoró. Esto sugiere que parte considerable de la capacidad predictiva fue utilizada en aprender la distribución de los datos negativos. Dicha característica es propia de los datos utilizados, por lo que fue necesario mitigar su efecto preprocesando los datos negativos. De esta forma se obtuvo dos nuevos conjuntos que solo consideraron datos físicos, utilizando las mismas distancias de muestreo que en los primeros conjuntos considerados (32 y 14 celdas) para realizar una comparación cuantitativa de la calidad de reconstrucción para cada caso.

Los conjuntos de datos donde los datos negativos fueron preprocesados presentaron un rendimiento superior respecto al caso en el que se dejaron los datos negativos (considerados valores sin sentido físico). Dentro de los resultados se obtuvo que todas las métricas mejoraron de forma consistente, sugiriendo que la distribución de los datos logró ser mejor representada por las redes tras quitar los valores negativos que no significaban un aporte para este trabajo. Debido a que el rendimiento de predicción de la interpolación bicúbica también aumentó cabe pensar que la distribución de los datos físicos es más fácil de predecir para un modelo, aunque no utilice aprendizaje de máquinas.

El uso de redes con más parámetros podría resultar beneficioso, ya que tienen un mayor potencial y pueden capturar de manera más efectiva la distribución de los datos con los que se entrenan. Los resultados obtenidos por SRGAN demostraron ser superiores dentro de un mismo conjunto de datos. Para los conjuntos de datos con mayor superposición (y, por lo tanto, una mayor cantidad de datos disponibles) los resultados obtenidos mejoraban, lo cual muestra una capacidad predictiva más alta. Esta red se beneficia de características como su mayor profundidad y el uso de un entrenamiento adversario diseñado específicamente para resolver problemas de súper resolución. En contraste, la red SRCNN no logró alcanzar el mismo rendimiento que SRGAN, aunque fue superior al método base bicúbico. Esta red logró obtener mejores resultados mientras más datos se encontraran disponibles, excepto para el caso de los conjuntos de datos sin valores negativos (tablas 5.5 y 5.6), donde la métrica de PSNR empeoró levemente. Es relevante mencionar que SRCNN destaca por su ligereza y la menor cantidad de parámetros, lo cual la hace una opción más eficiente en términos de costos computacionales de entrenamiento, como se ilustra en las tablas 5.1 y 5.4. Además, es importante destacar que los resultados obtenidos por Dong et al. [7] no se lograron replicar en las dos redes que utilizaron los valores no físicos debido a las diferencias significativas en los conjuntos de datos utilizados, pero se logró un rendimiento comparable en la métrica PSNR para los modelos que utilizaron los conjuntos sin valores negativos en la sección 5.2. Este trabajo se diferencia a trabajos con imágenes en que en estos últimos el dominio de valores es restringido (valores de 0 a 255), junto a que existe una disponibilidad mucho mayor de datos con los cuales trabajar, mientras que los datos satelitales son más ruidosos y menos masivos.

Como ya fue mencionado, de los datos disponibles de HCHO se obtuvo dos conjuntos de datos según su distancia de muestreo, cada uno con su variante que considera datos negativos y otra que no considera datos negativos. El primero presentó una mayor distancia de muestreo (por ende, un traslape de ventanas pequeño), con lo cual se dispone de una mayor heterogeneidad en los datos. En contraste, el segundo conjunto presentó una menor distancia de muestreo (es decir, un mayor traslape de ventanas), lo que incrementó la cantidad de datos disponibles para el entrenamiento. El objetivo de este experimento fue evaluar el rendimiento predictivo de los modelos en ambos casos. Se observó que para los casos que consideran

valores negativos, aunque el segundo conjunto de datos tiene una menor heterogeneidad, la mayor disponibilidad de ventanas para el entrenamiento benefició el rendimiento de las redes, en especial para el caso adversario. En cuanto a los casos en los que solo se consideraron los valores físicos (es decir, con datos negativos preprocesados), para SRGAN el rendimiento aumentó mientras más datos estuvieran disponibles, pero para el caso de SRCNN la mayor cantidad de datos y menor heterogeneidad empeoró las métricas de PSNR y MAPE levemente. Este resultado muestra la diferencia en la capacidad predictivas de las redes, donde en el caso de usar solo valores físicos (no negativos) SRGAN mejoró su rendimiento con más datos, pero SRCNN no logró realizar dicha mejora. Esto puede ser un indicio de la necesidad de complejizar dicha red, con tal de que logre capturar las distribuciones de mejor forma. Cabe destacar que la baja en el valor de las métricas de SRCNN ocurre también para el caso de la interpolación bicúbica, la cual forma parte del preprocesamiento de la red SRCNN (ver figura 4.1).

Los valores negativos forman un gran impacto dentro de las datos y como se mostró, no aportan al aprendizaje de las redes. Para esto se realizó el preprocesamiento de los valores negativos en los conjuntos de datos, con lo cual se obtuvo una mejora considerable de rendimiento en las métricas utilizadas. La diferencia más destacable es el alto rendimiento en PSNR que llegó a valores de 33,8[dB] y 34[dB], y para la métrica MAPE que bajó a 43,73% y 42,38% para las redes SRCNN y SRGAN respectivamente.

5.4. Reconstrucción de datos OMI

Uno de los objetivos de este trabajo de tesis es la reconstrucción de datos de OMI. Para ello se obtuvo un promedio anual de los datos del año 2007 y estos fueron reconstruidos por las redes entrenadas con ambos conjuntos de datos (muestreo cada 32 y 14 celdas) pero considerando el caso donde los valores negativos fueron preprocesados, ya que su rendimiento fue superior. Gracias a que las redes convolucionales son agnósticas al tamaño de entrada de los datos a predecir, fue posible realizar la reconstrucción de los datos de América del Sur del año especificado en las figuras 5.19 y 5.21, con las redes entrenadas con el primer conjunto de datos (distancia de muestreo de 32 celdas) y el segundo conjunto de datos (distancia de muestreo de 14 celdas) respectivamente, para el caso en el que los valores negativos fueron preprocesados.

Para realizar una mejor evaluación de los resultados predichos se realizó también una predicción sobre la zona central y sur de Chile en las figura 5.20 y 5.22 para los modelos de datos entrenados con el primer y segundo conjunto de datos respectivamente. De esta forma se puede identificar de mejor manera la reconstrucción de la red en una escala espacial más pequeña. Una de las características de la reconstrucción de estas redes es que SRCNN tiende a entregar resultados más suavizados mientras que SRGAN capta más detalles de píxeles individuales que la primera red mencionada suavizaría [8].

Dentro de los resultados obtenidos en las figuras 5.19 y 5.21, los mapas de América del Sur entregan reconstrucciones visualmente similares para ambos conjuntos de datos, que al compararlos con los datos de baja resolución de OMI mantienen los patrones de concentraciones en las localizaciones correspondientes. En cuanto a las figuras 5.20 y 5.22, las reconstruc-

ciones de la zona centro y sur de Chile muestran el aumento en la resolución de forma más visible debido a la vista más cercana de los datos, donde ambas redes muestran resultados similares. De forma general, las reconstrucciones no se diferenciaron notablemente entre el conjunto con muestreo de 32 celdas (mayor cantidad de datos) y el de muestreo de 14 celda (menor cantidad de datos).

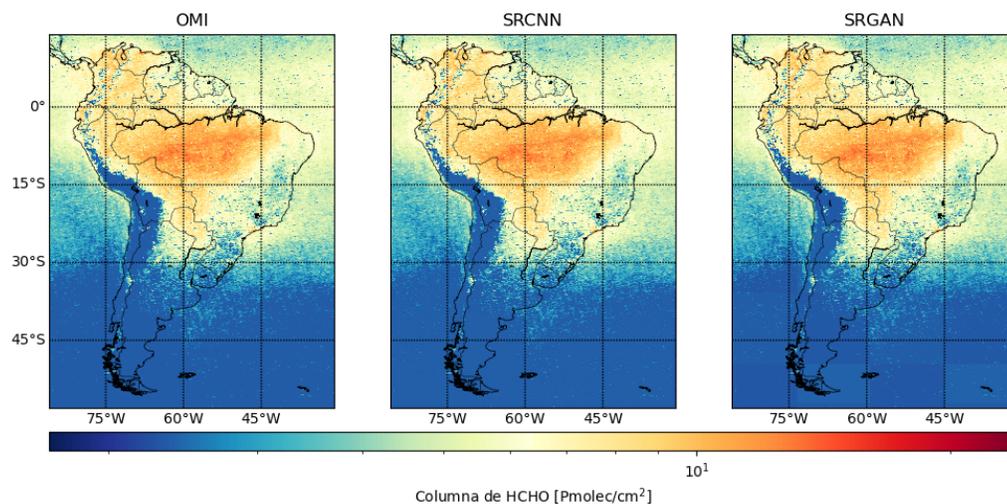


Figura 5.19: Reconstrucción de promedio anual de datos de OMI de América del Sur para al año 2007 utilizando redes entrenadas con el primer conjunto de datos sin valores negativos en escala logarítmica. a) Dato de OMI de baja resolución. b) Reconstrucción de SRCNN. c) Reconstrucción de SRGAN.

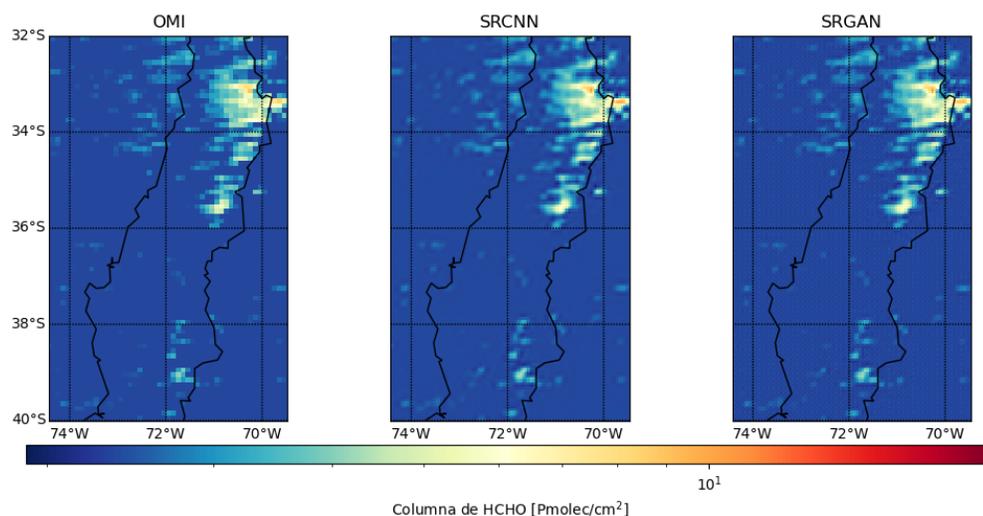


Figura 5.20: Reconstrucción de promedio anual de datos de OMI de la zona central y sur de Chile para al año 2007 utilizando redes entrenadas con el primer conjunto de datos sin negativos valores en escala logarítmica. a) Dato de OMI de baja resolución. b) Reconstrucción de SRCNN c) Reconstrucción de SRGAN.

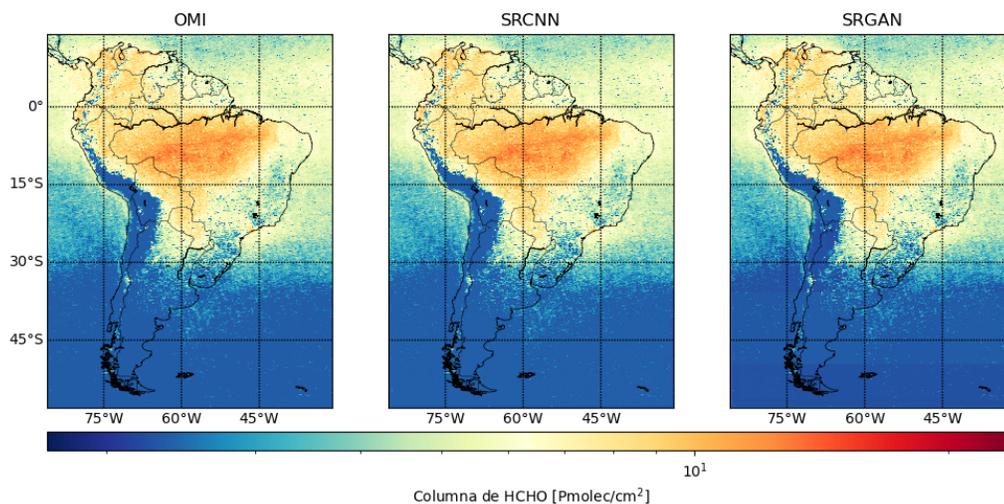


Figura 5.21: Reconstrucción de promedio anual de datos de OMI de América del Sur para al año 2007 utilizando redes entrenadas con el segundo conjunto de datos sin valores negativos en escala logarítmica. a) Dato de OMI de baja resolución. b) Reconstrucción de SRCNN c) Reconstrucción de SRGAN.

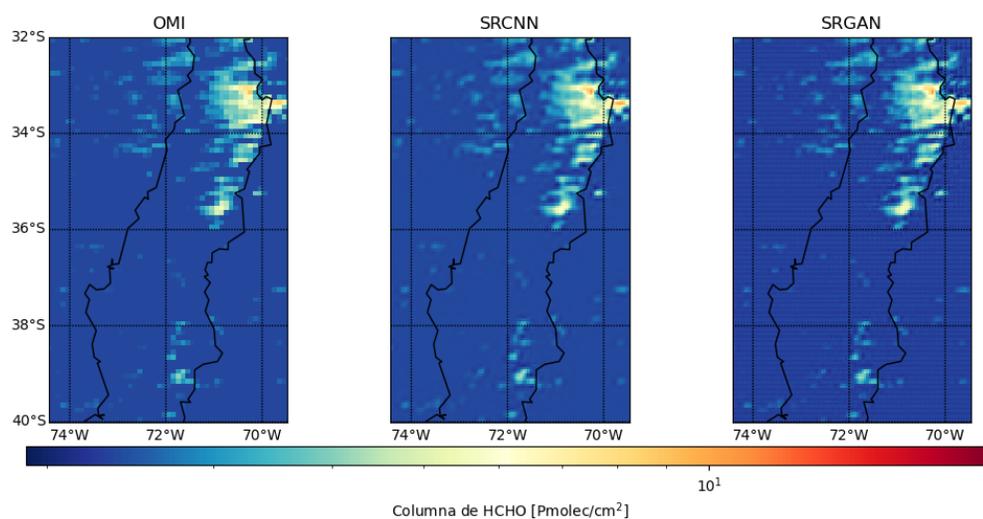


Figura 5.22: Reconstrucción de promedio anual de datos de OMI de la zona central y sur de Chile para al año 2007 utilizando redes entrenadas con el segundo conjunto de datos sin valores negativos en escala logarítmica. a) Dato de OMI de baja resolución. b) Reconstrucción de SRCNN. c) Reconstrucción de SRGAN.

Capítulo 6

Conclusión y trabajo futuro

El objetivo de este trabajo de tesis fue el de evaluar la factibilidad del uso de Redes Neuronales Convolucionales para la mejora en resolución espacial de mediciones satelitales de HCHO, reconstruyendo desde una menor a una mayor resolución. Las imágenes de baja resolución corresponden a HCHO de OMI, que se mejoran a la resolución espacial de TROPOMI que es ocupado como referencia. La metodología consistió en el uso de la red SRCNN para determinar qué tan efectivo es el modelo sobre los datos HCHO y SRGAN para introducir el uso de redes adversarias sobre este conjunto de datos.

Para evaluar la factibilidad de uso de ambas redes estas fueron entrenadas y evaluadas sobre dos conjuntos de datos obtenidos de la misma fuente de datos de TROPOMI HCHO, considerando para cada uno el caso en el que se utilizaron valores negativos y otro caso en el que los valores negativos fueron preprocesados. Estos modelos fueron comparados con la interpolación bicúbica mediante métricas de reconstrucción. De manera general los mejores resultados se obtuvieron con los conjuntos de datos que solo utilizaron valores físicos, es decir, con preprocesamiento de valores negativos. Estos buenos resultados se ven acrecentados si se toman consideraciones que beneficien la disponibilidad de los datos. En cuanto a los modelos, el de mejor rendimiento de forma consistente fue la red adversaria, logrando adaptarse a los datos debido a su mayor cantidad de parámetros y al proceso de aprendizaje adversario empleado. Otra ventaja de las redes adversarias es que realizan el aumento de resolución internamente, sin necesidad de un preprocesamiento como el requerido por SRCNN, que podría llevar a la pérdida de información y afectar el rendimiento. Por otro lado, SRCNN mostró que, aunque su capacidad predictiva es limitada, su ligereza representa una ventaja significativa al requerir menos recursos computacionales tanto para el entrenamiento como para la evaluación.

Uno de los aspectos a destacar de este trabajo fue la metodología de entrenamiento empleada. Las redes convolucionales utilizadas para este trabajo son agnósticas al tamaño de entrada a predecir. Esto resultó crucial frente a la considerable cantidad de datos faltantes y los valores posteriormente filtrados, distribuidos espacialmente en el conjunto de datos. La metodología permitió aprovechar la información disponible a través de ventanas, enfocándose en áreas donde los datos estaban presentes y evitando secciones con datos faltantes. De este modo, se logró entrenar los modelos utilizando información disponible en localidades

pequeñas para realizar una reconstrucción posterior sin importar el tamaño de los datos de entrada. Cabe destacar el positivo impacto del preprocesamiento de los datos negativos al obtener las ventanas, lo cual aportó a mucho mejores resultados.

El objetivo principal de este trabajo de tesis consistió en la reconstrucción de datos de HCHO obtenidos del espectrómetro OMI. La evaluación de dichos datos fue crucial para verificar su utilidad, realizando una reconstrucción sobre América del Sur que demostró la capacidad predictiva de ambas redes. Estos resultados prometen continuar utilizando estas técnicas en futuros estudios sobre América del Sur y posiblemente en otras áreas de interés. Además, se logró captar características de concentraciones tanto a gran escala como a baja escala, enriqueciendo la comprensión de los datos obtenidos. Es por esto que se espera continuar la línea de tiempo de datos disponibles de OMI y aplicarla a sectores de interés, tales como la zona central de Santiago de Chile.

Posibles mejoras pueden ser examinadas y analizadas a futuro para obtener resultados más confiables, que por limitaciones de tiempo tuvieron que ser excluidas. La exploración de distintos procesamientos previos de los datos para la generación de ventanas puede generar un impacto en la predicción que realicen los modelos a entrenar. También analizar cómo varían los resultados ante distintos umbrales de varianza para la elección de las ventanas que conforman el conjunto de datos.

Sería posible implementar entrenamientos adversarios adaptativos para mantener un equilibrio óptimo entre el generador y el discriminador y que ninguna de las dos redes esté sobreentrenada frente a la otra, lo cual teóricamente mejoraría el proceso de aprendizaje. Además, explorar diferentes inicializaciones de las redes podría influir significativamente en su capacidad de aprendizaje. Aunque hubo limitaciones de tiempo que impidieron la realización de experimentos, sería deseable investigar redes con mayor profundidad para mejorar la capacidad predictiva. Incrementar el número de capas y parámetros en SRCNN y probar distintos tipos de bloques y configuraciones en SRGAN podrían ofrecer resultados mejorados. Explorar nuevas arquitecturas tales como *transformers* aplicados a percepción remota también es relevante para estudiar los posibles resultados que se puedan obtener. Por último, analizar cómo las diferentes redes responden a las variaciones de ruido proporcionaría una visión más completa sobre la robustez y fiabilidad de los resultados obtenidos, y permitiría comprobar si SRGAN es más robusta que SRCNN o si es solo superior en las métricas utilizadas para este trabajo.

En este trabajo se ha destacado la importancia de comprender y analizar los datos, sin limitarse únicamente a los modelos predictivos. Una diferencia de este trabajo con el realizado para estudiar el NO₂ [6] es que en este último se estimó el error de reconstrucción, por lo que un trabajo a futuro es el desarrollar una red que pueda calcular el error de reconstrucción para los datos de HCHO. Esta estimación permitirá investigar las imprecisiones presentes en los datos satelitales, las cuales pueden influir significativamente en la precisión de los modelos predictivos. Este enfoque proporcionará una comprensión más completa del comportamiento de los datos ante condiciones tan ruidosas como las observadas en datos satelitales. .

Bibliografía

- [1] I. De Smedt, G. Pinardi, C. Vigouroux, S. Compernelle, A. Bais, N. Benavent, F. Boersma, K.-L. Chan, S. Donner, K.-U. Eichmann, P. Hedelt, F. Hendrick, H. Irie, V. Kumar, J.-C. Lambert, B. Langerock, C. Lerot, C. Liu, D. Loyola, A. Piters, A. Richter, C. Rivera Cárdenas, F. Romahn, R. G. Ryan, V. Sinha, N. Theys, J. Vlietinck, T. Wagner, T. Wang, H. Yu, and M. Van Roozendael. Comparative assessment of tropomi and omi formaldehyde observations and validation against max-doas network column measurements. *Atmospheric Chemistry and Physics*, 21(16):12561–12593, 2021.
- [2] Arthur Dias Freitas and Adalgiza Fornaro. Atmospheric formaldehyde monitored by tropomi satellite instrument throughout 2020 over são paulo state, brazil. *Remote Sensing*, 14(13), 2022.
- [3] D. Li, S. Wang, R. Xue, J. Zhu, S. Zhang, Z. Sun, and B. Zhou. Omi-observed hcho in shanghai, china, during 2010–2019 and ozone sensitivity inferred by an improved hcho / no₂ ratio. *Atmospheric Chemistry and Physics*, 21(20):15447–15460, 2021.
- [4] J. J. M. Acdan, R. B. Pierce, A. F. Dickens, Z. Adelman, and T. Nergui. Examining tropomi formaldehyde to nitrogen dioxide ratios in the lake michigan region: implications for ozone exceedances. *Atmospheric Chemistry and Physics*, 23(14):7867–7885, 2023.
- [5] Hanadi Al-Mekhlafi and Shiguang Liu. Single image super-resolution: a comprehensive review and recent insight. *Frontiers of Computer Science*, 18(1):181702, Sep 2023.
- [6] Santiago Parraguez. Improving omi-no 2 spatial resolution using a stochastic convolutional neural network over central and southern chile. Master’s thesis, Universidad de Chile - Facultad de Ciencias Físicas y Matemáticas, 2021.
- [7] C. Dong, C. Loy, K. He, and X. Tang. *Image Super-Resolution Using Deep Convolutional Networks*. Transactions on Pattern Analysis and Machine Intelligence. IEEE, 2016.
- [8] Karen Stengel, Andrew Glaws, Dylan Hettinger, and Ryan N. King. Adversarial super-resolution of climatological wind and solar data. *Proceedings of the National Academy of Sciences*, 117(29):16805–16815, 2020.
- [9] Runtai Zou, Li Wei, and Lei Guan. Super resolution of satellite-derived sea surface temperature using a transformer-based model. *Remote Sensing*, 15(22), 2023.
- [10] Junjie Li, Yizhuo Meng, Chongxin Tao, Zhen Zhang, Xining Yang, Zhe Wang, Xi Wang, Linyi Li, and Wen Zhang. Convformersr: Fusing transformers and convolutional neural

- networks for cross-sensor remote sensing imagery super-resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–15, 2024.
- [11] Bryan N. Duncan, Yasuko Yoshida, Jennifer R. Olson, Sanford Sillman, Randall V. Martin, Lok Lamsal, Yongtao Hu, Kenneth E. Pickering, Christian Retscher, Dale J. Allen, and James H. Crawford. Application of omi observations to a space-based indicator of nox and voc controls on surface ozone formation. *Atmospheric Environment*, 44(18):2213–2223, 2010.
- [12] Ki-Hyun Kim, Shamin Ara Jahan, and Jong-Tae Lee. Exposure to formaldehyde and its potential human health hazards. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev*, 29(4):277–299, October 2011.
- [13] C. Vigouroux, B. Langerock, C. A. Bauer Aquino, T. Blumenstock, Z. Cheng, M. De Mazière, I. De Smedt, M. Grutter, J. W. Hannigan, N. Jones, R. Kivi, D. Loyola, E. Lutsch, E. Mahieu, M. Makarova, J.-M. Metzger, I. Morino, I. Murata, T. Nagahama, J. Notholt, I. Ortega, M. Palm, G. Pinardi, A. Röhlings, D. Smale, W. Stremme, K. Strong, R. Sussmann, Y. Té, M. van Roozendaal, P. Wang, and H. Winkler. Tropomi–sentinel-5 precursor formaldehyde validation using an extensive network of ground-based fourier-transform infrared stations. *Atmospheric Measurement Techniques*, 13(7):3751–3767, 2020.
- [14] Ulrich Platt and Jochen Stutz. *Differential optical absorption spectroscopy*. Physics of Earth and Space Environments. Springer, Berlin, Germany, 2008 edition, June 2008.
- [15] Hengzi Liu, Ning Lu, Hou Jiang, Jun Qin, and Ling Yao. Filling gaps of monthly terra/modis daytime land surface temperature using discrete cosine transform method. *Remote Sensing*, 12(3), 2020.
- [16] Guojie Wang, Damien Garcia, Yi Liu, Richard de Jeu, and A. Johannes Dolman. A three-dimensional gap filling method for large geophysical datasets: Application to global satellite soil moisture observations. *Environmental Modelling and Software*, 30:139–142, 2012.
- [17] Damien Garcia. Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics and Data Analysis*, 54(4):1167–1178, 2010.
- [18] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network, 2017.
- [19] Jiqing Zhang, Chengjiang Long, Yuxin Wang, Xin Yang, Haiyang Mei, and Baocai Yin. Multi-context and enhanced reconstruction network for single image super resolution. In *2020 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, 2020.
- [20] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution, 2017.

- [21] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [22] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.
- [23] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.
- [24] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, and Xiaoou Tang. Esrgan: Enhanced super-resolution generative adversarial networks, 2018.
- [25] Zhongyuan Wang, Kui Jiang, Peng Yi, Zhen Han, and Zheng He. Ultra-dense gan for satellite imagery super-resolution. *Neurocomputing*, 398:328–337, 2020.
- [26] Umme Sara, Morium Akter, and Mohammad Shorif Uddin. Image quality assessment through fsim, ssim, mse and psnr—a comparative study. *Journal of Computer and Communications*, 2019.
- [27] Z. Wang, E.P. Simoncelli, and A.C. Bovik. Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems and Computers, 2003*, volume 2, pages 1398–1402 Vol.2, 2003.
- [28] Alain Horé and Djemel Ziou. Image quality metrics: Psnr vs. ssim. In *2010 20th International Conference on Pattern Recognition*, pages 2366–2369, 2010.
- [29] Rob J. Hyndman and Anne B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006.
- [30] Wenfu Sun, Lei Zhu, Isabelle De Smedt, Bin Bai, Dongchuan Pu, Yuyang Chen, Lei Shu, Dakang Wang, Tzung-May Fu, Xiaofei Wang, and Xin Yang. Global significant changes in formaldehyde (hcho) columns observed from space at the early stage of the covid-19 pandemic. *Geophysical Research Letters*, 48(4):2e020GL091265, 2021. 2e020GL091265 2020GL091265.
- [31] I. De Smedt, N. Theys, H. Yu, T. Danckaert, C. Lerot, S. Compernelle, M. Van Roozendael, A. Richter, A. Hilboll, E. Peters, M. Pedergnana, D. Loyola, S. Beirle, T. Wagner, H. Eskes, J. van Geffen, K. F. Boersma, and P. Veefkind. Algorithm theoretical baseline for formaldehyde retrievals from s5p tropomi and from the qa4ecv project. *Atmospheric Measurement Techniques*, 11(4):2395–2426, 2018.
- [32] Richard Meller and Geert K. Moortgat. Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323 k in the wavelength range 225–375 nm. *Journal of Geophysical Research: Atmospheres*, 105(D6):7089–7101, 2000.
- [33] J.P. Veefkind, I. Aben, K. McMullan, H. Förster, J. de Vries, G. Otter, J. Claas, H.J. Eskes, J.F. de Haan, Q. Kleipool, M. van Weele, O. Hasekamp, R. Hoogeveen, J. Landgraf, R. Snel, P. Tol, P. Ingmann, R. Voors, B. Kruizinga, R. Vink, H. Visser, and P.F. Levelt. Tropomi on the esa sentinel-5 precursor: A gmes mission for global observations of the

- atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 120:70–83, 2012. The Sentinel Missions - New Opportunities for Science.
- [34] P.F. Levelt, G.H.J. van den Oord, M.R. Dobber, A. Malkki, Huib Visser, Johan de Vries, P. Stammes, J.O.V. Lundell, and H. Saari. The ozone monitoring instrument. *IEEE Transactions on Geoscience and Remote Sensing*, 44(5):1093–1101, 2006.
- [35] Fabian Romahn, Mattia Pedergnana, Diego Loyola, Arnoud Apituley, Maarten Sneep, and J. P. Veefkind. *Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Formaldehyde HCHO*, 2022.
- [36] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, Octubre 1986.
- [37] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.
- [38] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Z. Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. *CoRR*, abs/1912.01703, 2019.
- [39] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015.
- [40] Chih-Yuan Yang, Chao Ma, and Ming-Hsuan Yang. Single-image super-resolution: A benchmark. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 372–386, Cham, 2014. Springer International Publishing.
- [41] Wenming Yang, Xuechen Zhang, Yapeng Tian, Wei Wang, Jing-Hao Xue, and Qingmin Liao. Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia*, 21(12):3106–3121, 2019.
- [42] M. A. A. Ghaffar, A. McKinstry, T. Maul, and T. T. Vu. Data augmentation approaches for satellite image super-resolution. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W7:47–54, 2019.

Anexos

Anexo A

En este anexo se presentan algunos términos necesarios para entender este trabajo de tesis.

Función de activación Una función de activación corresponde a una función que recibe la señal de entrada de un nodo de una red neuronal y la transforma en una señal de salida. El objetivo principal de estas funciones es introducir no linealidad dentro de la red, lo cual permite que aprendan patrones o distribuciones complejas de los datos.

Función ReLU La Unidad Lineal Rectificador (Rectified Linear Unit, ReLU) es una función de activación que filtra los valores negativos. Este comportamiento permite introducir no linealidad en las predicciones de la red.

$$ReLU(x, \alpha) = \max(0, x) \quad (\text{A.1})$$

Función Leaky ReLU La *Leaky Rectified Linear Unit* (Leaky ReLU) es una función de activación similar a la ReLU convencional en el sector positivo, pero considera una pendiente α en el dominio negativo de la función.

$$LeakyReLU(x, \alpha) = \max(0, x) + \alpha \cdot \min(0, x) \quad (\text{A.2})$$

Función sigmoide Función que procesa un valor el dominio real y lo asocia a un valor entre 0 y 1. Utilizado ampliamente en problemas de clasificación binarios en aprendizaje de máquinas, ya que permite modelar una salida como una probabilidad.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (\text{A.3})$$

Backpropagation Algoritmo de entrenamiento de redes neuronales que busca minimizar una función objetivo. El error es calculado desde la salida de la red y propagado hasta las capas iniciales (desde la salida hasta la entrada), lo cual permite el ajuste de los pesos en las capas de la red y con ello su aprendizaje.

Learning rate Fracción o porcentaje de cambio con el cual se actualizan los pesos de la red. En caso de utilizar un learning rate muy alto puede que la red no converja al quedarse en mínimos locales, mientras que un learning rate muy pequeño aumenta el tiempo de entrenamiento.