



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

**DETECCIÓN AUTOMÁTICA DE GLACIARES ROCOSOS MEDIANTE
IMÁGENES SATELITALES Y DEEP LEARNING**

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN MATEMÁTICAS
APLICADAS

CATHERINE JOSEFA LEMUN LEMUN

PROFESOR GUÍA:
JAIME ORTEGA PALMA

PROFESOR CO-GUÍA:
FABIÁN MEDEL GARCÍA

MIEMBROS DE LA COMISIÓN:
VANESSA AROS ESPINOZA
JAMES MCPHEE TORRES
FELIPE TOBAR HENRÍQUEZ

Este trabajo ha sido parcialmente financiado por: CMM ANID BASAL FB210005,
FONDECYT 1201125, FONDEF ID22I10199 e IDEA I+D ID21I10305

SANTIAGO DE CHILE
2024

DETECCIÓN AUTOMÁTICA DE GLACIARES ROCOSOS MEDIANTE IMÁGENES SATELITALES Y DEEP LEARNING

En respuesta a los escenarios proyectados por el cambio climático en los últimos años, se ha vuelto fundamental cuantificar las fuentes de almacenamiento hídrico en los territorios [29, 34, 43]. El propósito es reducir la incertidumbre en cuanto a la disponibilidad de agua y gestionar de manera sostenible su uso, tanto para las comunidades locales como para las actividades vinculadas al desarrollo humano. Específicamente, cobra relevancia la necesidad de realizar mapas y cuantificar las reservas de agua en áreas de alta montaña. En este contexto, la presente tesis propone una metodología automatizada para respaldar la identificación y cartografía de glaciares rocosos. Con la ayuda del uso de imágenes satelitales de libre distribución y técnicas de aprendizaje profundo (*deep learning*) se programa un modelo de segmentación semántica de imágenes. Su arquitectura se compone de una red neuronal convolucional InceptionResNetV2-UNet, donde en la etapa de codificación se utiliza aprendizaje por transferencia (*transfer learning*). Se evalúa la sensibilidad del modelo respecto al tipo de datos de entrada mediante tres experimentos: El primer experimento (E1) emplea únicamente imágenes ópticas, el segundo experimento (E2) se basa exclusivamente en datos de velocidad de desplazamiento obtenidos de imágenes SAR, y el tercer experimento (E3) utiliza datos ópticos y de velocidad. Todos estos experimentos se aplican en un área de estudio ubicada en la zona de montaña de la Región Metropolitana.

Se logran resultados comparables con la literatura en la identificación de glaciares rocosos para áreas superiores a $0.1 [km^2]$. E1 detectó 71.3% de los glaciares totales de la zona de validación. E2 identificó un 53.9% y E2 un 79.1%. Si bien E3, identifica una mayor cantidad de GR, para el propósito de ejecución y escalabilidad el modelo, E1 entrega mejores condiciones debido a que obtiene resultados comparables en la identificación y delimitación, pero con un costo computacional mucho más bajo. Como propuestas de mejora, se sugiere ampliar la base de datos y el número de clases a identificar para potenciar la capacidad de generalización del modelo presentado en E1.

A mi familia de origen y elegida.

Agradecimientos

Quiero dedicar las primeras palabras de esta página a mi familia *Lemun*. Gracias por inculcar en mí el valor de la alegría y la resiliencia; su ejemplo ha sido mi inspiración para poder sobrellevar difíciles momentos siempre con esperanza. Gracias a mi madre *María Jacqueline* y mi padre *Sebastián Rojito* por todos estos años compartidos; los quiero montones. Agradezco a mi familia no humana, quienes me han enseñado la ternura y delicadeza de la vida: Jerby, Susy, Marti, Lucas, Bubito, D'arty, Doña, Mati y, sobre todo, a Joaquito y Vita.

Una especial mención a Nataly por ser una hermana desde el primer día que nos conocimos y a Estela por encaminarse a lo mismo conmigo. A Benjamín, por todos los años que crecimos juntos y seguiremos compartiendo nuestra amistad. También a mis amigas del colegio/universidad ♡: Kathy, Feña, Flo y Javi. A las personas que me acogieron en su hogar y se convirtieron en una segunda madre/padre/hermanita: Tía Susy, Xime, Marcelín y Cata, gracias por enseñarme a ser mejor persona. Siempre están presentes sus enseñanzas y afecto.

Sumado a esto, quiero expresar mi gratitud a las personas del DIM y el equipo Copernicus, con quienes viví este proceso. En particular, agradezco a mi director de tesis, Jaime Ortega, por siempre considerar el lado humano de las cosas; sin su comprensión y apoyo esta tesis no hubiese sido posible. Agradezco a Fabián Medel por ayudarme en la realización de este trabajo y por todas las historias de viajes que siempre sacaron carcajadas. A mis primeros amigos del DIM: Vale, Francisco, Juan Pablo, Tomás y Sebastián; gracias por acogirme y hacerme sentir tan parte del grupo, los quiero mucho. Extiendo esta gratitud a mis compañeros/amigos de LES OTRES C.S, son todos luz. Agradezco a la gente del actual Matrak Edition, en especial a Gonzalo y Pedro por las risas y amistad; Fabián y Leo por su motivación para compartir; Jota y Manu por valorar mis consejos y; sobre todo, a Benja por querer ser y convertirse en mi lugar de paz. Toda mi estadía en el DIM no hubiese sido igual, sin contar con el cariño y ayuda de Karen, Natacha, María Cecilia, Silvia, Georgina, Linda, Don Luis y Don Oscar. En la seguidilla de agradecimientos, como no nombrar a mis primeros amigos del laboratorio, Javier y Fran, son mi lugar seguro. A mis nuevos compañeros/amigos: Karina, Cristóbal, Ian, Carla, Lucho y César, gracias por su trato amable y preocupación.

También, quiero extender mis agradecimientos a los académicos y profesionales: Felipe Tobar, James McPhee y Vanessa Aros. Gracias por aceptar formar parte de mi comisión de tesis y ser parte de este proceso. De igual forma, extiendo reconocimiento al financiamiento otorgado por los proyectos CMM ANID BASAL FB210005, FONDECYT 1201125, FONDEF ID22I10199 y FONDEF IDEA I+D ID21I10305.

Finalmente, quiero dar las gracias a mi abuelito Ramiro (Tata♡), su amor incondicional siempre ha sido mi hogar.

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Hipótesis	3
1.3. Objetivo General	3
1.4. Objetivos Específicos	3
1.5. Estructura de la Tesis	4
2. Marco Teórico	5
2.1. Glaciares Rocosos y Cubiertos	5
2.1.1. Tipos de Glaciares	6
2.2. Percepción Remota e Imágenes Satelitales	8
2.2.1. Percepción Remota	8
2.2.2. Imágenes Ópticas	9
2.2.3. Imágenes SAR e Interferometría	10
2.3. Deep Learning y Visión Computacional	16
2.3.1. Preliminares	16
2.3.2. Deep Learning	18
2.3.3. Visión Computacional	20
2.3.4. Aprendizaje por Transferencia	21
2.3.5. Arquitecturas	22
2.3.6. Métricas	24

2.4. Deep Learning y Detección de Glaciares Rocosos	26
3. Metodología	29
3.1. Área de Estudio	29
3.2. Recopilación de datos	30
3.2.1. Ópticos	30
3.2.2. Polígonos	31
3.2.3. SAR	32
3.3. Muestreo	34
3.3.1. Máscara	34
3.3.2. Imágenes	34
3.4. Dataset	35
3.5. Modelo, Entrenamiento y Postprocesamiento	37
4. Resultados y Discusión	39
4.1. Desempeño de la cartografía automatizada	39
4.2. Área, Pendiente y Orientación	40
5. Conclusiones	45
Bibliografía	50

Índice de Tablas

2.1. Ejemplo de forma para la matriz de confusión para un problema de clase binaria	24
2.2. Resumen de los diversos modelos de aprendizaje profundo desarrollados para la cartografía de glaciares rocosos. Estos modelos han sido aplicados en diferentes contextos geográficos y haciendo uso de diversas fuentes de datos, teniendo buenos resultados en la identificación y cartografía de glaciares rocosos en entornos montañosos.	28
3.1. Clasificación primaria de glaciares y forma considerada en el Inventario Público de Glaciares de Chile elaborado por la DGA 2022.	31
3.2. Descripción de etiquetas y colores para las dos clases implementadas en el modelo.	34
3.3. Esquema general de la arquitectura InceptionResnetV2 - UNet utilizada para la segmentación semántica de imágenes.	37
4.1. Resultados del entrenamiento para los experimentos E1 (OPT), E2 (SAR) y E3 (OPT+SAR)	39
4.2. Número de glaciares rocosos distintos detectados por el primer experimento (OPT) y el tercer experimento (OPT+SAR), junto con la cantidad idéntica que identifican.	44

Índice de Ilustraciones

2.1. Estructura de un Glaciar Rocosó [4]	7
2.2. Comportamientos de sensores utilizados para la percepción remota. A la izquierda el sensor pasivo y a la derecha el sensor activo. Las líneas punteadas son las recibidas por el sensor y las continuas corresponden a radiación emitida ya sea por una fuente externa al sensor o el mismo sensor.	9
2.3. Representación de las bandas consideradas en imágenes espectrales. La figura abarcando diferentes rangos de longitud de onda. A la vez señala el Porcentaje de Transmisión Atmosférica el cual indica cómo la atmósfera interactúa con diferentes longitudes de onda.	9
2.4. Proceso de separación de bandas en imágenes multiespectrales e hiperespectrales. Las imágenes multiespectrales generalmente capturan la radiación en varias bandas discretas y amplias. Como se muestra en la figura, estas bandas pueden incluir segmentos específicos del espectro electromagnético, como el visible y el infrarrojo cercano. Por otro lado, las imágenes hiperespectrales tienen la capacidad de capturar información en una mayor cantidad de bandas, con intervalos de longitud de onda más estrechos y numerosos	10
2.5. Ejemplo de imágenes ópticas e imágenes SAR ante un evento de erupción volcánica. A la izquierda, ejemplo de imagen óptica y a la derecha ejemplo de imagen SAR para el volcán Kamchatka ubicado en Rusia [31].	11
2.6. Estructura de los datos binarios para el almacenamiento de imágenes SLC (Single Look Complex Image) [19].	12
2.7. Gráfico de línea base de ejemplo para el método PSI y SBAS.	15
2.8. Esquema de relación entre aprendizaje de máquinas (<i>ML</i>), aprendizaje profundo (<i>DL</i>) y otras áreas de la inteligencia artificial (<i>IA</i>).	16
2.9. Esquema resumen para los tipos de tarea, métricas y experiencias en aprendizaje mediante algoritmo de <i>machine learning</i>	17
2.10. Ejemplo de capas que puede contener una CNN. Aprendizaje profundo para la extracción de aspectos en opiniones textuales [27].	19

2.11. Ejemplo para las tareas de clasificación, detección y segmentación de imagen en visión computacional.	20
2.12. Ejemplo de los diferentes tipos de segmentación de imagen [39].	21
2.13. Ejemplo de arquitectura U-Net para 32x32 píxeles en la resolución más baja. Cada caja azul corresponde a un mapa de características multicanal. El número de canales se indica en la parte superior de la caja. El tamaño x-y se proporciona en la esquina inferior izquierda de la caja. Las cajas blancas representan mapas de características copiados.[25].	22
2.14. Ejemplo de arquitectura InceptionResNetV2-UNet. La estructura de InceptionResNet-UNet utiliza la estructura U-Net tradicional, que consta de partes de codificador (muestreo reducido) y decodificador (muestreo ascendente) con bloques residuales [1].	24
2.15. Ejemplo de predicción según el valor de IoU, donde en la izquierda, el índice es 0, en el centro es 1/3 y en la derecha es 1.	26
3.1. Extensión geográfica de la zona de estudio seleccionada. En esta área, se identificaron y catalogaron un total de 392 glaciares clasificados como glaciares rocosos. La región abarca aproximadamente 2750 [km ²] y se ubica entre las longitudes -70.11 a -69.78 al oeste, así como las latitudes -34.04 a -33.24 al sur.	30
3.2. En la Figura se presentan las cuatro imágenes ópticas correspondientes a las bandas BGR (Azul - banda 1, Verde- banda 2, Rojo- banda 3) y NIR (Infra-rojo Cercano- banda 1) para el área de entrenamiento. Estas imágenes fueron obtenidas a partir de la misión Sentinel 2 y se utilizaron en el proceso de análisis.	31
3.3. Mapa de velocidad promedio generados para el área de estudio. Donde a) mapa para el área de validación y b) mapa para el área de entrenamiento.	33
3.4. PCA para imágenes óptica del área de entrenamiento a utilizar en el primer experimento.	35
3.5. Parches para la creación del dataset. En la figura se muestra la grilla creada para obtener los parches de las imágenes provenientes de los datos Ópticos, Ópticos+SAR y SAR en el área de estudio ocupada como entrenamiento y la ocupada como validación. De estos, solo se tomaron los parches marcados en azul, para equilibrar el conjunto de datos de entrenamiento. La imagen a) corresponde a la zona utilizada como validación y la imagen b) la utilizada como entrenamiento.	36
3.6. Ejemplo de parches (representados en falso color) generados para cada uno de los experimentos con su respectiva máscara.	36

4.1. Comparación de la cartografía de glaciación rocosos (RG) generada por los modelos entrenados (E1: Azul, E2: Verde y E3: Celeste) y la proveniente del inventario de la DGA (verde lima).	40
4.2. Relación entre área, pendiente y orientación de los glaciares rocosos detectados en cada experimento: a) Resultados del experimento 1 utilizando imágenes ópticas. b) Resultados del experimento 2 basados en imágenes SAR. c) Resultados del experimento 3 utilizando imágenes mixtas.	41
4.3. Porcentaje de glaciares de roca detectados respecto al primer experimento basado en imágenes ópticas, el segundo basado en datos procesados de imágenes SAR y el tercero, correspondiente datos ópticos y SAR	42
4.4. Comparación entre los glaciares rocosos detectado por cada experimento según su área.	42
4.5. Comparación entre los glaciares rocosos detectado por cada experimento según su pendiente.	43
4.6. Comparación entre los glaciares rocosos detectado por cada experimento según su orientación.	43
4.7. Comparación de las características de los glaciares detectados por el primer y tercer experimento.	44

Capítulo 1

Introducción

Este capítulo de introducción sienta las bases para comprender la relevancia y los objetivos específicos que orientan este trabajo hacia el desarrollo de una metodología automatizada en el ámbito de la cartografía de glaciares rocosos.

1.1. Motivación

Durante últimos años, y en respuesta a los escenarios proyectados en el contexto del cambio climático, los estudios orientados a cuantificar las reservas hídricas, mejorar su gestión y mitigar los efectos de las transformaciones climáticas sobre estos, se vuelven cada vez más relevantes a nivel internacional. Los cambios climáticos actuales y provistos ejercen impacto sobre los sistemas de aguas subterráneas y superficiales, manifestándose tanto de forma directa, a través de procesos como la reposición por recarga y evaporación, como de manera indirecta, mediante modificaciones en el uso del agua [29, 34, 43]. Simultáneamente, las variaciones en el ciclo hídrico afectan los ecosistemas que dependen de este recurso, así como las actividades humanas que se apoyan en él. En esa línea, la seguridad del suministro de agua para el consumo humano y diversas actividades, como la agricultura y la minería, se ve amenazada por estas variaciones [3].

Con respecto a Chile, este cuenta con un total de 922 [km^3] de aguas renovables al año, colocándolo en el puesto 14 a nivel mundial. No obstante, se observa un notorio desequilibrio geográfico entre la distribución de estos recursos y la densidad poblacional en la zona norte y central del país. El aumento de las temperaturas en los últimos años ha ejercido un impacto significativo en los sistemas acuáticos; generando un déficit de precipitaciones [15]. Junto a esto, ambas zonas concentran el 65% de la población del país [14]. En este contexto, es esencial cuantificar las fuentes de almacenamiento hídricos presentes en el territorio con el fin de disminuir la incertidumbre en términos de disponibilidad de agua y manejar de forma sostenible el uso de este recurso tanto para las comunidades locales como para las actividades asociadas al desarrollo humano. En particular, destaca la necesidad de cartografiar y cuantificar las reservas de agua en áreas de alta montaña. Siendo los glaciares un foco principal, en parte, porque Chile concentra el 80% de los glaciares de Sudamérica (26180 glaciares

en el país) y por otro lado, debido a que éstos son reservas estratégicas de agua en estado sólido y actúan como indicadores del Cambio Climático [11, 44]. Ante lo expuesto y dada su proximidad a importantes centros urbanos e instalaciones industriales, tanto los glaciares descubiertos como los glaciares rocosos son recursos de gran relevancia a nivel nacional.

En general, el reconocimiento y cartografía de glaciares descubiertos en nuestro territorio se ha llevado a cabo de forma exitosa, debido a sus características superficiales distintivas y visibles de estas crioformas. Los glaciares descubiertos son masas de hielo que se encuentran sin cobertura de roca o sedimentos; y presentan una superficie clara de nieve y hielo, lo que facilita su identificación y cartografía mediante el uso de imágenes satelitales [11, 20]. Por otro lado, los glaciares cubiertos se encuentran parcial o totalmente cubiertos bajo una capa de detritos tales como escombros, rocas, gravilla o tierra [23]. Similarmente, los glaciares rocosos (GR) también se encuentran cubiertos por una capa de detritos y poseen una estructura interna compuesta por una combinación de rocas y sedimentos finos. No obstante, su porcentaje de hielo es menor a los de los glaciares cubiertos, el cual va entre 40% a 60% del volumen total [11, 5].

La tarea de localizar y delimitar los glaciares rocosos mediante métodos *in situ* se vuelve costosa debido a las complicadas características geográficas en las que se pueden encontrar estas formaciones geológicas. Si bien, el empleo de imágenes satelitales ha permitido sobrellevar las barreras geográficas como avanzar en la capacidad de cartografiar; dicho proceso todavía requiere interpretación manual, lo que conlleva un trabajo laborioso y subjetivo que a menudo conduce a inconsistencias en su categorización [4, 36]. Con el fin de mejorar el reconocimiento y cartografía de estos glaciares, es que en los últimos años se han estado generando nuevas metodologías que involucran el uso de imágenes satelitales y aprendizaje de máquinas (*machine learning*) [12, 18, 36] para disminuir los sesgos a causa de la interpretación manual y con esto poder obtener un criterio más unificado de clasificación visual. Si bien se han registrado avances en la detección, los resultados aún carecen de precisión y presentan diversos desafíos. Entre ellos se encuentra la dificultad de distinguir entre la superficie del glaciar y los elementos rocosos circundantes, lo que requiere la búsqueda continua de métodos más avanzados y adaptativos. A su vez, muchos de estos desarrollos de detección automática precisan el uso de bases de datos provenientes de imágenes de alta resolución, lo que conlleva altos costos económicos. Se suma la generación de datos de mayor exigencia computacional, como el procesamiento de interferogramas. Estas problemáticas dificultan la escalabilidad del método a grandes regiones, así como la implementación y uso de estas técnicas por parte de expertos no pertenecientes al área de la percepción remota y el aprendizaje de máquinas.

Con lo anterior, esta tesis tiene como finalidad desarrollar una metodología automatizada para apoyar la identificación de glaciares rocosos bajo un enfoque basado en el uso de imágenes satelitales de libre distribución y técnicas de aprendizaje profundo (*deep learning*). Se busca superar las limitaciones de modelos previamente diseñados mediante la aplicación de diferentes estrategias que permitan una identificación más eficiente de estos tipos de glaciares, disminuyendo tiempos de procesamiento y costos económicos con el fin de llegar a un producto que sea aplicable a extensas regiones y accesible para usuarios no expertos.

1.2. Hipótesis

La hipótesis de este estudio postula que es factible diseñar una metodología automatizada basado en imágenes satelitales que identifique glaciares rocosos y cartografie estas crioformas. La metodología toma imágenes satelitales de libre distribución y técnicas de aprendizaje profundo. Se espera superar las limitaciones de los modelos previamente diseñados mediante la aplicación de nuevas estrategias que disminuyan el tiempo de procesamiento y simplifiquen los requisitos de los modelos previos con el fin de que sea escalable a extensas regiones y más accesible para usuarios no expertos.

1.3. Objetivo General

El objetivo general de este trabajo es desarrollar una metodología basada en técnicas de inteligencia artificial, en particular con *deep learning*, junto con imágenes satelitales provenientes de sensores remotos de libre distribución (sensores ópticos y radar), que permitan detectar y cartografiar glaciares rocosos presentes en la Región Metropolitana.

1.4. Objetivos Específicos

Los objetivos específicos de esta tesis son:

1. Determinar los parámetros geomorfológicos, geológicos y climáticos bajo los cuales se desarrollan los glaciares rocosos.
2. Revisar arquitecturas de *deep learning* y métodos de procesamiento datos satelitales que tengan la capacidad de resolver el problema planteado.
3. Generar diferentes bases de datos para revisar que el modelo de *deep learning* seleccionado capte la información necesaria para detectar glaciares rocosos.
4. Revisar la contribución de los datos InSAR para la detección de glaciares rocosos de forma automatizada.
5. Implementar un modelo basado en *deep learning* para detectar glaciares rocosos mediante datos satelitales.
6. Comparar modelo de *deep learning* utilizado con trabajos previos para la detección de glaciares rocoso.
7. Automatizar el procesamiento de datos satelitales y entrenamiento del modelo de *deep learning* para la detección de glaciares rocosos.
8. Analizar el funcionamiento de la metodología establecida, con el fin de identificar áreas de mejora futura.

1.5. Estructura de la Tesis

En el segundo capítulo, se lleva a cabo una revisión conceptual para comprender la problemática de detección y cartografía de glaciares rocosos mediante *deep learning* e imágenes satelitales. En una primera instancia, se detallan las características distintivas de estos glaciares, posteriormente, se abordan conceptos de percepción remota, proporcionando una explicación más detallada de los atributos presentes en las imágenes satelitales y en particular de la técnica de InSAR. En las secciones subsiguientes, se introduce de manera general el tema sobre *deep learning* y se expone en mayor profundidad las técnicas necesarias para abordar la problemática de detección de objetos. El capítulo culmina con una revisión del estado del arte sobre el empleo de técnicas de *deep learning* basadas en imágenes satelitales para la detección de glaciares rocosos.

En el tercer capítulo, se detallan las metodologías empleadas para la recopilación de datos, el preprocesamiento y el entrenamiento del modelo, con el objetivo de abordar la problemática de detección de glaciares rocosos. Se explora la importancia de la información proveniente de InSAR, mediante la elaboración de tres experimentos.

El cuarto capítulo se centra en las discusiones derivadas de los resultados obtenidos por el modelo en los diferentes experimentos, así como la relevancia de los datos provenientes de InSAR en dicho contexto.

Finalmente, en el quinto capítulo, se presentan las conclusiones obtenidas y se proponen mejoras para la metodología implementada que van orientadas en aumentar la precisión tanto para la detección automatizada de nuevos glaciares como para la cartografía automatizada de estas formas criogénicas.

Capítulo 2

Marco Teórico

El marco teórico se enfoca en tres áreas principales: primero, se aborda la temática de glaciares rocosos, explorando diversos tipos y sus características distintivas. Luego, se examina la importancia de percepción remota e imágenes satelitales, detallando conceptos fundamentales y diferencias entre imágenes ópticas y SAR, junto con lo que es interferometría.

En la segunda sección se expone sobre *deep learning* y visión computacional, proporcionando antecedentes y analizando conceptos clave como aprendizaje por transferencia, segmentación semántica de imágenes, entre otros. Además, se exploran diversas arquitecturas y métricas en el contexto del aprendizaje profundo y la visión computacional que son relevantes para este estudio.

2.1. Glaciares Rocosos y Cubiertos

Actualmente, existen múltiples definiciones que refieren a los glaciares, lo cual ha llevado a la generación de variadas tipologías para describir las formas que estas masas de hielo pueden adoptar en la naturaleza. Una definición clásica corresponde a la dada por Cogley et al. (2011), donde establece que un glaciar es “una masa perenne de hielo, y posiblemente nevada, que se origina en la superficie terrestre por la recristalización de la nieve u otras formas de precipitación sólida y que muestra evidencia de flujo pasado o presente” [9]. No obstante, esta definición tiene diversas limitaciones. Por un lado, un cuerpo de hielo puede llegar a ser considerado un glaciar; debido a que no propone un acuerdo sobre el límite superior para estas masas de hielo. Por otro lado, no existe claridad sobre el tiempo de duración de la capa de hielo, por lo cual parche de nieve no podrían ser distinguidos hidrológicamente de un glaciar si es que persiste por más de una estación [21].

Una definición operativa y aplicable a sensores remotos es la utilizada por la Dirección General de Aguas de Chile (DGA) en el 2009, para la Estrategia Nacional de Glaciares [11]. Aquí establece que “un glaciar es toda superficie de hielo y nieve permanente generada sobre suelo, que sea visible por períodos de al menos 2 años y de un área igual o superior a $0,01[km^2]$ (una hectárea). O cualquier superficie rocosa con evidencia superficial de flujo

viscoso, producto de un alto contenido de hielo actual o pasado en el subsuelo”. Si bien esta definición ayuda a abordar las confusiones relacionadas con la extensión y la duración de la capa de hielo, puede generar contradicciones con otras caracterizaciones establecidas [35].

Con el fin de precisar una visión clara de las diferentes formas y condiciones en que se pueden encontrar estos cuerpos de hielo en Chile. Se presenta una caracterización tradicional de tres niveles para describir las formaciones glaciares en los Andes [20]. Las categorías que componen la clasificación tradicional corresponden a:

- Glaciares Descubiertos o Blancos
- Glaciares Cubiertos
- Glaciares Rocosos.

Cabe destacar que la definición de glaciar puede variar según el contexto y los criterios utilizados.

2.1.1. Tipos de Glaciares

Glaciar Descubierto o Blanco

Los glaciares descubiertos corresponden a aquellas masas de hielo que se encuentran sin cobertura de roca o sedimentos. El hielo es claramente visible y la cantidad de detritos o impurezas tanto internas como externas que contienen es muy baja [20]. Este tipo de glaciares se forman a partir de la acumulación de nieve, que se compacta con el tiempo transformándose en neviza y finalmente en hielo [35].

Glaciar Cubierto

Los glaciares cubiertos tienen como principal característica que se encuentran parcial o totalmente bajo una capa de detritos tales como escombros, rocas, gravilla o tierra, lo que les proporciona una protección adicional contra el derretimiento y la sublimación. A diferencia de los glaciares descubiertos, la superficie del glaciar no es visible directamente y por completo; debido a la cobertura.

Estas masas de hielo puede tener un origen glaciogénico, es decir, a partir de glaciares inicialmente descubiertos, los cuales, debido a procesos como adelgazamiento progresivo, retroceso o aumento de recubrimiento, empiezan a desarrollar una capa más gruesa y extensa de detritos sobre las estructuras de hielo superficial [35]. Paralelamente, la existencia de eventos como avalanchas, caída de rocas u otros movimientos, dejan las estructuras internas de hielo cubiertas con estratos que contienen impurezas y escombros [23]. En la literatura provista por Janke et al. (2015) se establecen tres subclases para los glaciares cubiertos [20]:

- Clase 1: Cobertura parcial de la superficie del glaciar.

- Clase 2: Cobertura completa de la superficie del glaciar con espesor inferior a 3 [m].
- Clase 3: Cobertura completa de la superficie del glaciar con espesor superior a 3 [m].

Adicionalmente a su posible origen, de acuerdo con Crose y Milana (2002), se plantea que estos glaciares pueden ser considerados como una etapa de transición en la transformación entre un glaciar expuesto y un glaciar rocoso [10].

Glaciar Rocosos

Los glaciares rocosos son masas de hielo que están cubiertas por una capa de detritos y que poseen una estructura interna compuesta por una combinación de rocas, sedimentos finos y un porcentaje de hielo menor a los de los glaciares cubiertos, el cual va entre 40 % a 60 % del volumen total (Figura 2.1). Tal composición hace que tengan una velocidad de desplazamiento lenta sobre laderas o valles cuesta abajo. Se ha estimado una velocidad entre los 10-150 [cm] al año [5].

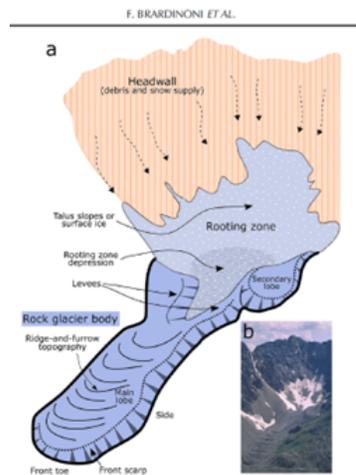


Figura 2.1: Estructura de un Glaciar Rocosos [4]

Una caracterización de la forma de los glaciares rocosos la entrega Janke (2015), estableciendo 3 clases continuas a las descritas para los glaciares cubiertos [20]:

- Clase 4: Corresponde a glaciares rocosos que tienen un porcentaje de hielo mayor, por lo cual presentan crestas transversales pronunciadas y surcos que se arquean a lo largo de la superficie, producto del flujo del hielo.
- Clase 5: Tienen un flujo reducido, debido al porcentaje limitado de hielo que poseen. Esta característica genera crestas y surcos lineales, que van en la dirección del flujo.
- Clase 6: Topografía superficial tenue porque el movimiento del glaciar de roca ha cesado.

2.2. Percepción Remota e Imágenes Satelitales

La Percepción Remota se presenta como una herramienta para la observación y análisis de nuestro entorno, utilizando tecnologías que nos permiten recopilar datos relevantes sobre la superficie terrestre. En este capítulo, se desarrollan tres temas específicos: La percepción remota en sí misma, las imágenes ópticas y las imágenes SAR (*Synthetic Aperture Radar*) con el fin de dar un marco general que sustente la información recopilada mediante imágenes satelitales en la detección de glaciares rocosos.

2.2.1. Percepción Remota

De manera general, podemos definir la percepción remota como los procedimientos y técnicas que se emplean para detectar información de un objeto sin tener contacto con él [13]. Con respecto a su origen, diferentes autores postulan la creación de la fotografía como el hito que inicia la percepción remota. Aunque, no fue hasta 1950, donde la geógrafa y oceanógrafa estadounidense Evelyn Pruitt, empleó por primera vez el termino *remote sensing*, que traducimos como *percepción remota*. Ya en 1957, con el lanzamiento por la URSS del primer satélite espacial (Sputnik) empieza la era para la percepción remota moderna en conjunto con la digitalización de las imágenes y desarrollo de técnicas de procesamiento para éstas.

Para detectar información de los objetos de interés, la percepción remota se basa en las interacciones electromagnéticas que pueda tener con los objetos de estudio. El comportamiento que tengan los objetos con la radiación depende de la naturaleza de éstos. Al recibir radiación los objetos pueden reflejarla, absorberla o transmitirla. También existen objetos que emiten radiación y otros que no.

Por otro lado, los procesos asociados a percepción remota son muy variados y se puede hacer una categorización en dos grandes etapas que corresponden por un lado a la captura y transmisión de los datos y por otro el análisis de éstos. Respecto a la primera categoría, tenemos los objetos físicos, los sensores y sus plataformas. En cuanto a la segunda, está la extracción de información, transformación y aplicación. Para lo que concierne a los sensores, estos obtienen datos cuando interactúan con el objeto a través de la radiación electromagnética. Estos datos pueden abarcar información sobre espectro no visible, resoluciones inusuales, etc. Los sensores podemos categorizarlos en:

- **Sensores Activos:** Emite radiación en la dirección del objetivo a investigar. Posteriormente, detecta y mide la radiación que se refleja o se dispersa desde el objetivo (Figura 2.2).
- **Sensores Pasivos** Detectan la radiación natural que emite o refleja el objeto a investigar (Figura 2.2).

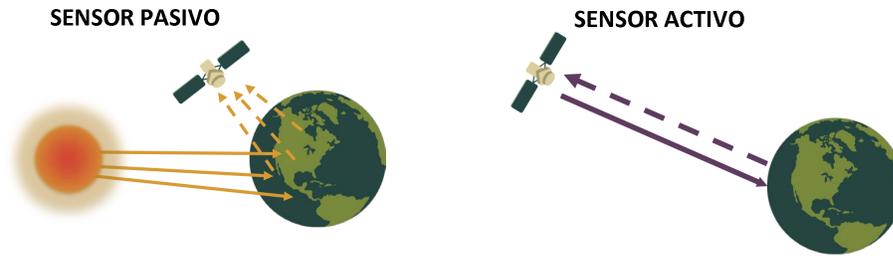


Figura 2.2: Comportamientos de sensores utilizados para la percepción remota. A la izquierda el sensor pasivo y a la derecha el sensor activo. Las líneas punteadas son las recibidas por el sensor y las continuas corresponden a radiación emitida ya sea por una fuente externa al sensor o el mismo sensor.

2.2.2. Imágenes Ópticas

En el ámbito de la percepción remota, las imágenes satelitales se generan mediante mediciones de la radiación electromagnética en rangos específicos, comúnmente denominados bandas (Figura 2.3). Estas mediciones se realizan a través de sensores equipados con un instrumento conocido como espectrómetro. El espectrómetro captura la radiación y genera una imagen digital, donde cada elemento de la imagen (píxel) tiene un valor discreto expresado en unidades de Número Digital (DN).

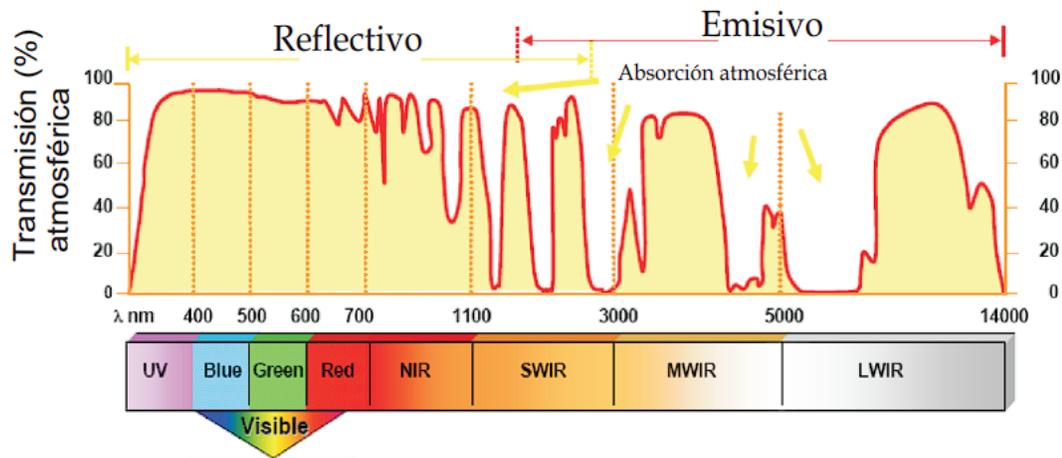


Figura 2.3: Representación de las bandas consideradas en imágenes espectrales. La figura abarcando diferentes rangos de longitud de onda. A la vez señala el Porcentaje de Transmisión Atmosférica el cual indica cómo la atmósfera interactúa con diferentes longitudes de onda.

Los espectrómetros de imagen se categorizan como multispectrales o hiperespectrales, dependiendo del ancho espectral de cada banda y del número de bandas recopiladas, siendo una banda definida como un intervalo específico de longitudes de onda (Figura 2.4). Cuando el espectro recopilado comprende múltiples bandas que son espectralmente anchas y separadas, se clasifican como datos multispectrales, con un número de bandas típicamente en el orden de "decenas". En cambio, si el espectro recopilado consiste en bandas cuasi continuas y estrechas desde el punto de vista espectral, los datos se denominan hiperespectrales, con un

número de bandas que suele situarse en el orden de centenas".

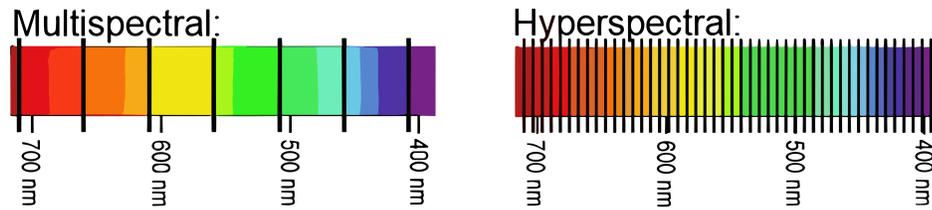


Figura 2.4: Proceso de separación de bandas en imágenes multiespectrales e hiperspectrales. Las imágenes multiespectrales generalmente capturan la radiación en varias bandas discretas y amplias. Como se muestra en la figura, estas bandas pueden incluir segmentos específicos del espectro electromagnético, como el visible y el infrarrojo cercano. Por otro lado, las imágenes hiperspectrales tienen la capacidad de capturar información en una mayor cantidad de bandas, con intervalos de longitud de onda más estrechos y numerosos

Un aspecto importante de las imágenes satelitales es su resolución, que puede variar según el sensor utilizado. Las principales resoluciones incluyen:

- Resolución Espacial: Define el tamaño medio de los píxeles en la imagen.
- Resolución Espectral: Se refiere al número y la ubicación de las bandas en el espectro electromagnético que el sensor puede captar.
- Resolución Radiométrica: Se mide en bits (dígitos binarios) y representa el rango de valores de brillo disponibles en la imagen.
- Resolución temporal: Indica el tiempo necesario para que el satélite realice una revisita a la misma área de la Tierra.

De los anterior podemos desprender de que cuanto más alta sea la resolución espacial, menos área cubre un solo píxel.

2.2.3. Imágenes SAR e Interferometría

Radar de Apertura Sintética - SAR

El Radar de Apertura Sintética, conocido como SAR por sus siglas en inglés (Synthetic Aperture Radar), es un sistema de radar que utiliza técnicas de procesamiento de señales para generar imágenes de alta resolución de la superficie terrestre.

A diferencia del radar convencional, que emite pulsos de radiofrecuencia y registra la señal de retorno reflejada, el SAR emite una señal continua y cartografía un área mientras se mueve. Esto se logra sintetizando una apertura de antena virtual más grande, lo que permite obtener una mayor resolución espacial y una mejor calidad de imagen. La principal ventaja del SAR es su capacidad para adquirir imágenes independientemente de las condiciones meteorológicas o la iluminación solar (Figura 2.5). El radar puede penetrar las nubes, la lluvia e

incluso la vegetación, lo que permite obtener imágenes de la superficie terrestre en diferentes condiciones.

Las imágenes SAR resultantes ofrecen una alta resolución espacial y una rica información sobre la topografía, la estructura y las propiedades de la superficie terrestre. Estas imágenes son utilizadas en una amplia gama de aplicaciones, como el monitoreo ambiental, la gestión de desastres, la agricultura, la cartografía, la vigilancia y la seguridad.

En forma simplificada, las mediciones realizadas por el radar SAR capturan dos características principales: la amplitud y la fase de las ondas asociadas. La amplitud representa la intensidad de la señal de retorno y está relacionada con las propiedades eléctricas, la forma y la orientación de los objetos en la escena observada. Por otro lado, la fase de la señal se refiere al cambio en la posición relativa de la onda electromagnética medida en términos de longitud de onda.

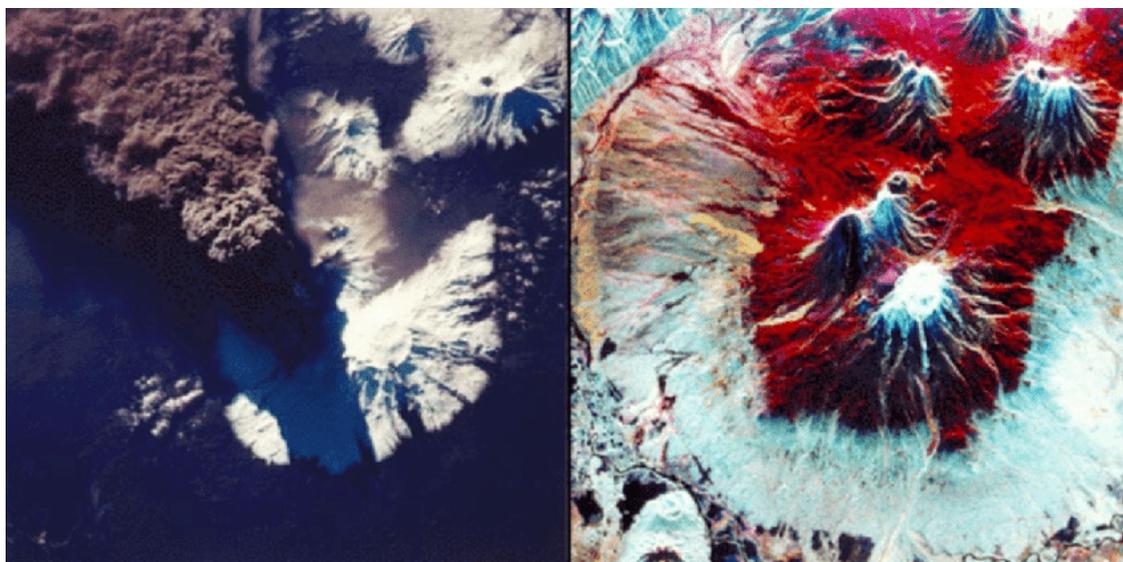


Figura 2.5: Ejemplo de imágenes ópticas e imágenes SAR ante un evento de erupción volcánica. A la izquierda, ejemplo de imagen óptica y a la derecha ejemplo de imagen SAR para el volcán Kamchatka ubicado en Rusia [31].

Imágenes Single-Look Complex (SLC)

Las imágenes Single Look Complex (SLC) reciben su nombre debido a que cada píxel muestra el resultado de una única medición de la reflexión de la señal de microondas emitida desde el satélite hacia la superficie terrestre, lo cual se conoce en inglés como *single measurement*. La caracterización de *complex* se refiere al hecho de que esta medición se analiza durante el procesamiento en tierra para obtener tanto la fase como la amplitud de la onda reflejada, y estas dos cantidades se representan como un número complejo. Cabe agregar que para que el tamaño de la imagen generada sea manejable, se reduce el rango dinámico de los números complejos.

Las imágenes SLC conservan las mediciones originales del sensor y no posee mayor in-

terferencia de artefactos de interpolación o problemas de proyección lo que les otorga la característica de tener una mayor fidelidad respecto a los otros productos de imagen SAR. Estas imágenes se almacenan en la dirección correspondiente a la línea de visión del satélite y se producen en una geometría Doppler cero. La geometría Doppler cero, corresponde a una corrección de precisión para compensar el efecto Doppler y mantener una relación constante de frecuencia entre la señal transmitida y la recibida. Esto se logra ajustando el tiempo de muestreo de la señal de microondas recibida durante el procesamiento de la imagen. Al eliminar el efecto Doppler, se obtiene una geometría más estable y precisa en la imagen SLC, lo que facilita la interpretación y el análisis de los datos.

En general, los sistemas almacenan la información de la imagen en matrices binarias. Se utilizan enteros de 16 bits para contener por separado la información de la componente real y la componente imaginaria, donde los píxeles son válidos si no están marcados como NaN. En la Figura 2.6 se presenta la estructura de los datos binarios para el almacenamiento de las imágenes SLC.

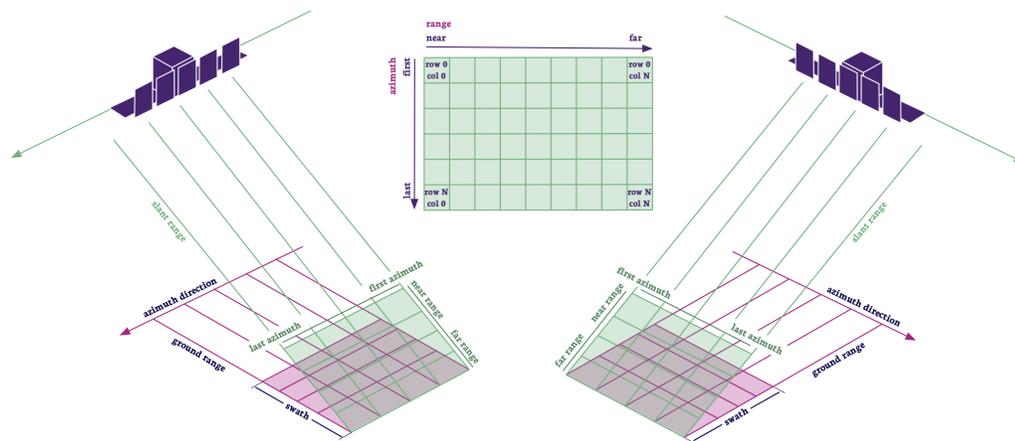


Figura 2.6: Estructura de los datos binarios para el almacenamiento de imágenes SLC (Single Look Complex Image) [19].

Principios Generales de InSAR

La interferometría SAR (InSAR, por sus siglas en inglés) es una técnica utilizada en el análisis de imágenes de radar de apertura sintética (SAR) para obtener información detallada sobre la superficie terrestre. InSAR combina pares de imágenes SAR adquiridas en diferentes momentos para medir cambios en la fase de retorno de la señal del radar. Estos cambios de fase se utilizan para extraer información sobre el desplazamiento de la superficie y otros fenómenos geofísicos.

El procesamiento de InSAR se basa en conceptos como la interferencia de ondas y el teorema de la convolución. La señal recibida en cada punto de la escena es una combinación de la señal reflejada por la superficie y otros factores como la contribución de la atmósfera y el ruido.

Matemáticamente, la señal recibida en un punto de la escena se puede representar como:

$$s(t) = A(t)e^{j(2\pi f(t)t + \phi(t))}$$

Donde:

- $s(t)$ es la señal recibida en el tiempo t ,
- $A(t)$ es la amplitud de la señal reflejada en el tiempo t ,
- $f(t)$ es la frecuencia instantánea de la señal en el tiempo t ,
- $\phi(t)$ es la fase de la señal en el tiempo t , j es la unidad imaginaria.

La interferencia se produce cuando se superponen dos señales recibidas en diferentes momentos. Si consideramos dos imágenes SAR adquiridas en momentos t_1 y t_2 , la señal interferométrica resultante se puede expresar como:

$$I(t) = s(t_1) \cdot s^*(t_2)$$

Donde $s^*(t_2)$ representa el conjugado complejo de la señal recibida en el momento t_2 . Al expandir y simplificar esta expresión, se llega a la siguiente ecuación:

$$I(t) = |A(t_1)| \cdot |A(t_2)| e^{j(2\pi\Delta f t + \Delta\phi)}$$

Donde:

- $\Delta f = f(t_1) - f(t_2)$ es la diferencia de frecuencia entre los dos momentos de adquisición,
- $\Delta\phi = \phi(t_1) - \phi(t_2)$ es la diferencia de fase entre los dos momentos de adquisición.

Esta ecuación muestra que la interferencia genera una nueva señal con una amplitud proporcional al producto de las amplitudes originales y una fase determinada por la diferencia de fase entre los momentos de adquisición.

En el procesamiento de InSAR, se busca obtener información sobre la fase interferométrica $\Delta\phi$, ya que está relacionada con los desplazamientos de la superficie y otros fenómenos geofísicos de interés. Para ello, se realizan técnicas de filtrado, corrección atmosférica y desempaquetado de fase para obtener mediciones precisas de los desplazamientos y deformaciones de la superficie.

Procesamiento de InSAR

La interferometría SAR (InSAR, por sus siglas en inglés) es una técnica utilizada en el análisis de imágenes de radar de apertura sintética (SAR) para obtener información detallada sobre la superficie terrestre. InSAR combina pares de imágenes SAR adquiridas en diferentes

momentos para medir cambios en la fase de retorno de la señal del radar. Estos cambios de fase se utilizan para extraer información sobre el desplazamiento de la superficie y otros fenómenos geofísicos.

El procesamiento de InSAR implica una serie de pasos para generar interferogramas y obtener mediciones precisas de los desplazamientos de la superficie. A continuación, se presentan los principales pasos involucrados en el procesamiento de InSAR:

- **Preprocesamiento de las imágenes:** Antes de generar el interferograma, es necesario realizar algunas correcciones y ajustes en las imágenes SLC. Esto puede incluir la compensación de movimiento *motion compensation*, corrección de fase topográfica y corrección del efecto del sensor, entre otros.
- **Selección de las imágenes adecuadas:** Selecciona dos imágenes SLC adecuadas para el proceso de interferometría. Estas imágenes deben ser adquiridas desde la misma plataforma de radar en momentos de tiempo cercanos y deben cubrir la misma área geográfica.
- **Registro de las imágenes:** Alinea y registra las dos imágenes seleccionadas para asegurar una coincidencia precisa de las características en ambas imágenes. Esto se hace para compensar cualquier diferencia en la posición y orientación entre las imágenes.
- **Interferometría:** Realiza el proceso de interferometría utilizando las dos imágenes registradas. Esto implica combinar las señales de radar y calcular la diferencia de fase entre las dos imágenes en cada píxel. El resultado es un interferograma que representa las diferencias de fase entre las dos imágenes.
- **Filtrado y correcciones adicionales:** Aplica filtros y realiza correcciones adicionales en el interferograma para eliminar el ruido y corregir artefactos. Estas correcciones pueden incluir la remoción de la fase atmosférica y la compensación del desplazamiento de fase debido a la topografía.
- **Desempaquetado de la fase:** Desempaqueta la fase del interferograma para obtener una medida continua de los desplazamientos de fase. Esto implica resolver el problema del "ambiguo de fase" que ocurre cuando la diferencia de fase excede el rango de $[-\pi, \pi]$.
- **Interpretación y análisis:** Analiza el interferograma desempaquetado para obtener información sobre los desplazamientos y deformaciones en la superficie observada. Esto puede incluir la detección de movimientos del terreno, deformaciones tectónicas o cambios en el movimiento de glaciares, entre otros fenómenos.

Es importante tener en cuenta que estos son pasos generales y el proceso exacto puede variar según el software y las técnicas utilizadas. Además, la generación de interferogramas de alta calidad está condicionada a diferentes factores, tales como la resolución de las imágenes SAR, nivel conocimientos especializados en el campo de la interferometría de radar, entre otros.

Interferometría SAR de Series Temporales

Las metodologías TS-InSAR (Interferometría SAR de Series Temporales) son técnicas avanzadas de procesamiento de datos SAR que permiten extraer información sobre la deformación del suelo a lo largo del tiempo. Si bien, la exactitud y la precisión de la medición generada por la técnica puede variar debido a la falta de correlación de la señal SAR causada, por ejemplo, a los errores de desenvolvimiento de fase, al retraso atmosférico o los cambios en el mecanismo de retrodispersión de los objetivos a lo largo del tiempo ha demostrado ser un enfoque valioso en aplicaciones geofísicas y de gestión de riesgos, como el monitoreo de volcanes, la evaluación de peligros geológicos y la vigilancia de subsidencias en áreas urbanas debido a su bajo costo de implementación[22].

Para la generación de series de tiempo, existen diferentes algoritmos (Figura 2.7 tales como el método *persistent scatterer* (PS) que se utiliza para identificar y medir de manera precisa los puntos de referencia persistentes en una serie de imágenes SAR. Estos puntos de referencia, también conocidos como *scatters* o 'puntos permanentes', son objetivos en la superficie terrestre que presentan una retrodispersión coherente y estable a lo largo del tiempo. El análisis mediante esta técnica se realiza en la escala completa de resolución espacial y centrándose en los píxeles que contienen un único dispersor dominante [33].

Junto a este, existe el método de *small baseline subset* o SBAS, el cual utiliza un subconjunto de imágenes SAR con líneas de base pequeñas para generar series de tiempo de deformación del suelo. La línea de base se refiere a la distancia perpendicular entre las trayectorias de vuelo del satélite en diferentes adquisiciones.

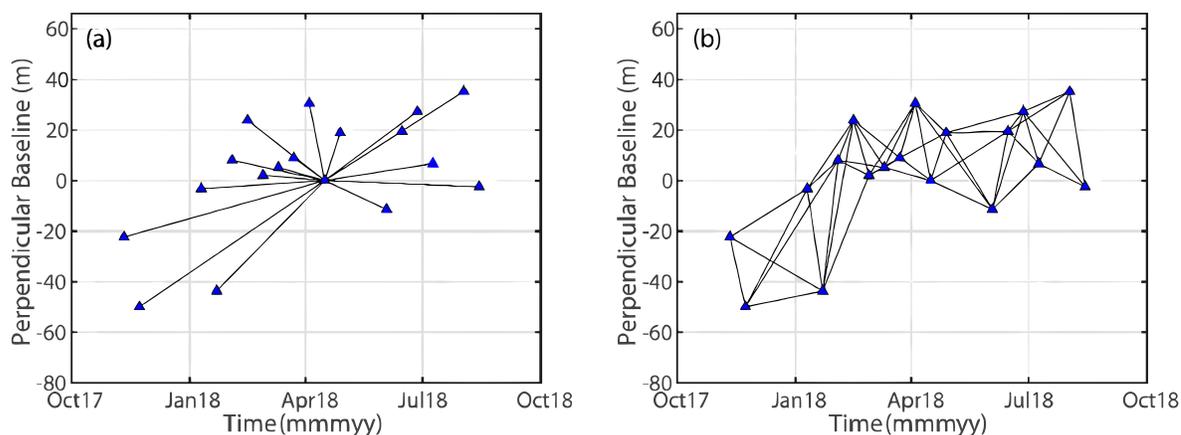


Figura 2.7: Gráfico de línea base de ejemplo para el método PSI y SBAS.

Durante el últimos años, los algoritmos SBAS han sido ampliamente utilizados para monitorear los desplazamientos lentos del terreno [24]. La base de ejecución para estos algoritmos implica identificar los puntos de alta coherencia en los interferogramas, donde la fase media supera un umbral determinado. Posteriormente, se elimina la fase de elevación residual con el objetivo de obtener los términos de velocidad.

2.3. Deep Learning y Visión Computacional

La inteligencia artificial (IA, por sus siglas en inglés), específicamente el aprendizaje profundo (*deep learning* o DL, en inglés), ha influido y modificado diversas disciplinas. Este capítulo proporciona una breve introducción a la inteligencia artificial y examina la interconexión entre subcampos como *machine learning* (ML), *neural network* (NN) y *deep learning*. Posteriormente, se revisa el empleo de técnicas de DL para la segmentación semántica de imágenes. Se profundiza cómo estas metodologías afectan la comprensión y el análisis de datos satelitales. Además de explorar cómo la segmentación semántica facilita la identificación de características en este tipo de imágenes.

2.3.1. Preliminares

El aprendizaje de máquinas, redes neuronales y aprendizaje profundo son subcampos de la inteligencia artificial (Figura 2.8). La inteligencia artificial (IA, por sus siglas en inglés) busca resolver problemas *intuitivos* utilizando técnicas que permiten a las máquinas imitar el comportamiento de los animales, especialmente los humanos [16].

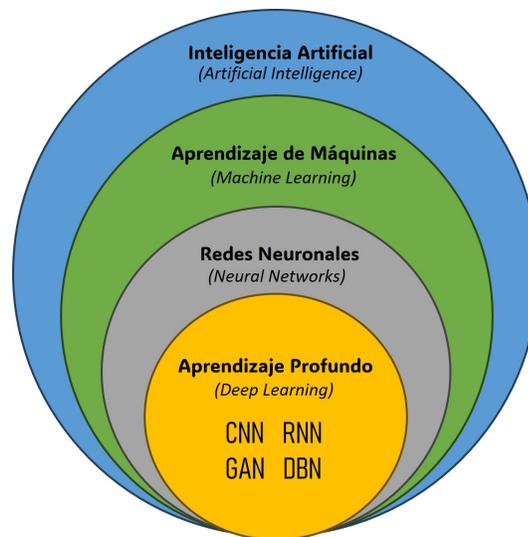


Figura 2.8: Esquema de relación entre aprendizaje de máquinas (*ML*), aprendizaje profundo (*DL*) y otras áreas de la inteligencia artificial (*IA*).

El campo del aprendizaje de máquinas (ML, por sus siglas en inglés) tiene como objetivo abordar problemas mediante el aprendizaje basado en la experiencia, utilizando datos de entrada. En este enfoque, los algoritmos de ML procesan la información de entrada, identifican patrones y realizan predicciones fundamentadas en los datos [16, 38]. Un algoritmo se considera que *aprende* de la experiencia \mathbf{E} cuando al llevar a cabo una tarea \mathbf{T} y evaluar su rendimiento a través de un conjunto de parámetros que definen una medida de desempeño \mathbf{P} 2.9. El desempeño de la tarea \mathbf{T} medido mediante \mathbf{P} , mejora de forma correlacionada con la experiencia \mathbf{E} [6].

Las tareas por realizar se determinan por el tipo de procesamiento que se aplicará a

los datos y se definen en función del tipo de predicción que se busca realizar (Figura 2.9). Por ejemplo, de acuerdo con el objetivo de la predicción, se pueden llevar a cabo tareas de regresión (predicción de una variable numérica), clasificación (etiquetado de una imagen), agrupamiento (basado en métodos de distancia), entre otras. En relación con la evaluación del rendimiento, se consideran métricas como la acierto (*accuracy*), la tasa de error (*error rate*), la sensibilidad (*specificity*), la divergencia de Kullback-Leibler, entre otras. La selección de estas métricas está vinculada al tipo específico de tarea que se está llevando a cabo. Además, las tareas pueden llevarse a cabo en diversos entornos de aprendizaje, que se clasifican de la siguiente manera:

1. Aprendizaje Supervisado: Se desarrollan algoritmos que requieren de ejemplos de entrenamiento para los cuales se conoce de antemano la salida correcta [38]. En concreto, se entrena un programa para aprender una función $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^d$, teniendo m prototipos de pares de entrada y salida (x, y) . Las tareas de clasificación se desarrollan en entornos de aprendizaje supervisados; aquí se le pide al programa que determine a cuál de las k categorías pertenece cierta entrada. Para resolver, generalmente se solicita al algoritmo que encuentre una función $f : \mathbb{R}^n \rightarrow \{1, \dots, k\}$. Cuando $y = f(x)$, el modelo asigna una entrada descrita por el vector x a una categoría identificada por y . Otro ejemplo es la tarea de regresión, la cual es similar la tarea de clasificación excepto porque se le pide al algoritmo que genere una función $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$, diferenciándose por el output [16].
2. Aprendizaje No Supervisado: Se conocen los datos de entrada x pero no se tienen variables de salida y . En esa línea, el objetivo de encontrar alguna función $f(x)$ que mejor ajuste $f(x)$ respecto a y se cambia a que los algoritmos traten de identificar patrones ocultos dentro de los datos de entrada. K-means y Análisis de Componentes Principales son ejemplos de tareas ejecutadas en entornos de aprendizaje no supervisados [16, 38].
3. Aprendizaje por Refuerzo: Este tipo de aprendizaje se centra en capacitar a un *agente* para que realice *acciones* apropiadas en función de un *estado* particular del *entorno*, utilizando una *política* con el objetivo de maximizar las *recompensas* a largo plazo [16, 32].



Figura 2.9: Esquema resumen para los tipos de tarea, métricas y experiencias en aprendizaje mediante algoritmo de *machine learning*.

2.3.2. Deep Learning

Las técnicas de *deep learning* (DL) son parte de la familia de métodos de *machine learning* que se basan en la aproximación de funciones no lineales desconocidas $f : X \rightarrow Y$, donde X representa el conjunto de posibles entradas e Y las salidas correspondientes. En el contexto de clasificación, Y puede ser un conjunto finito de clases.

comparando con los otros métodos de ML, los algoritmos de DL emplean redes neuronales artificiales con múltiples capas, también conocidas como redes neuronales profundas. Estas redes están diseñadas para aprender, ya sea de forma supervisada, no supervisado o por refuerzo, las representaciones jerárquicas de los datos; lo que les permite capturar características y patrones complejos en los conjuntos de entrada. A medida que los datos se propagan a través de las capas de la red, se aplican transformaciones no lineales y se ajustan los pesos de las conexiones entre las neuronas con el propósito de minimizar una función de pérdida, que mide la discrepancia entre las salidas predichas por la red y las salidas deseadas.

En concreto, para iniciar un algoritmo de DL se requiere un conjunto de observaciones $D = \{(x, f(x)) : x \in A\}$ que representa el fenómeno que se está estudiando. En la literatura, este conjunto también se conoce como conjunto de entrenamiento. Aquí, A es un subconjunto finito de X . El siguiente paso consiste en definir una familia de funciones candidatas $f_\theta : \theta \in \Xi$, donde buscamos una buena aproximación de la función f . En este caso, Ξ es un subconjunto de \mathbb{R}^κ para algún $\kappa \in \mathbb{N}$.

Continuando, la calidad de la aproximación se evalúa mediante una función de costo $L(\cdot; D) : \Xi \rightarrow \mathbb{R}$. Intuitivamente, buscamos encontrar el valor θ^* que minimiza la función de costo $L(\cdot; D)$ sobre el conjunto de candidatos Ξ . En otras palabras, seleccionamos f_{θ^*} como la aproximación elegida.

Por otro lado, una red neuronal convolucional (CNN) es una clase de red neuronal profunda que se utiliza principalmente para analizar imágenes. A diferencia de una red neuronal tradicional que realiza multiplicaciones de matrices, una CNN utiliza una técnica especial llamada convolución [24].

En matemáticas, la convolución es una operación que combina dos funciones para producir una tercera función que describe cómo la forma de una de ellas es modificada por la otra. En el contexto de las CNN, la convolución se aplica a la entrada de la red, que generalmente es una imagen, utilizando filtros o núcleos convolucionales. Estos filtros son pequeñas matrices de pesos que se deslizan por la imagen y calculan la convolución en cada región [24].

Si bien, las capas de una red neuronal convolucional (CNN) típica pueden variar según la arquitectura específica, se describen a continuación las más comunes (Figura 2.10).

1. Capa de entrada: Esta capa acepta las imágenes de entrada y las pasa a través de la red.
2. Capas convolucionales: Estas capas realizan la convolución de los filtros o núcleos en la imagen de entrada para extraer características importantes. Cada capa convolucional puede tener múltiples filtros y producir mapas de características.

3. Capas de activación: Después de la convolución, se aplica una función de activación, como ReLU (*Rectified Linear Unit*), para introducir no linealidad en la red y permitir la captura de relaciones más complejas en los datos.
4. Capas de agrupamiento (*Pooling*): Estas capas reducen la dimensionalidad espacial de los mapas de características mediante operaciones de pooling, como el max-pooling, que selecciona el valor máximo de una región determinada. El agrupamiento ayuda a disminuir el número de parámetros y la complejidad computacional. Una de las razones para realizar este proceso es evitar el sobreajuste.
5. Capas totalmente conectadas (*fully-connected*): Estas capas toman las características extraídas por las capas convolucionales y las transforman en una representación adecuada para la tarea específica. Cada neurona en estas capas está conectada a todas las neuronas de la capa anterior, lo que permite aprender relaciones más complejas. Sus activaciones se pueden calcular con una multiplicación de matrices seguida de un desplazamiento de sesgo.
6. Capa de salida: Esta capa produce las salidas finales de la red, que dependen del tipo de problema que se esté abordando. Puede ser una capa de clasificación con múltiples neuronas representando diferentes clases, o una capa de regresión que produce valores continuos.

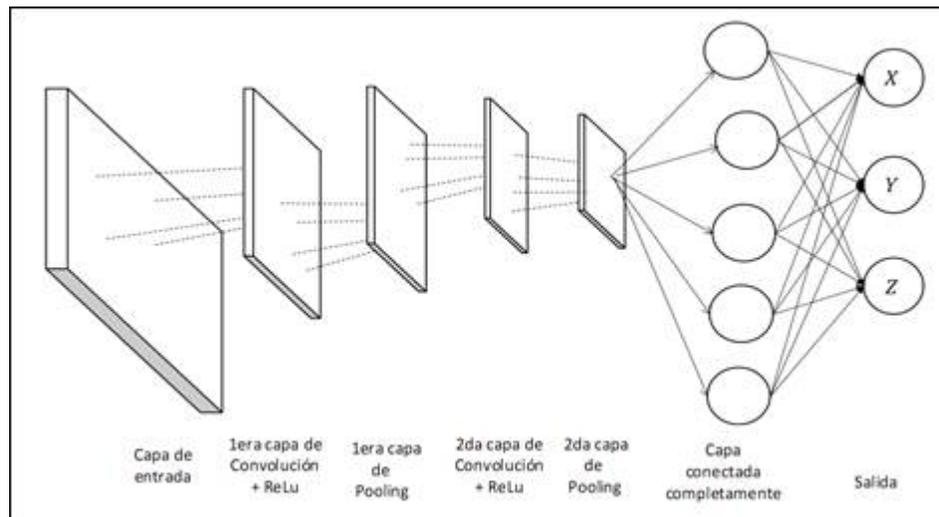


Figura 2.10: Ejemplo de capas que puede contener una CNN. Aprendizaje profundo para la extracción de aspectos en opiniones textuales [27].

Al tener capas más profundidad y conexiones con un alto nivel de complejidad entre neuronas, las CNN logran aprender características y patrones más sofisticados de las imágenes, lo que les permite realizar las tareas de detección, clasificación y segmentación con mejor precisión [39].

2.3.3. Visión Computacional

La visión por computadora es un campo de estudio que se centra en permitir a las máquinas comprender y procesar imágenes o vídeos de manera similar a como lo hacen los seres humanos. Dentro de la visión por computadora, hay varias tareas fundamentales que se realizan para analizar y extraer información de las imágenes. Las tres tareas principales son la clasificación, la detección y la segmentación (Figura 2.11).

1. **Clasificación:** La clasificación en visión por computadora implica asignar una etiqueta o categoría específica a una imagen. Se trata de determinar la clase a la que pertenece una imagen en particular. Por ejemplo, se puede entrenar un modelo de clasificación para reconocer si una imagen contiene un perro o un gato. Esto implica aprender características distintivas de las diferentes clases y utilizar algoritmos de aprendizaje automático para tomar decisiones de clasificación basadas en esas características.
2. **Detección:** La detección se centra en identificar la presencia y ubicación de objetos específicos dentro de una imagen. En lugar de simplemente clasificar una imagen o segmentarla en regiones, la detección implica detectar y localizar objetos específicos dentro de la imagen, dibujando cuadros delimitadores alrededor de ellos. Esta tarea es especialmente útil cuando se trata de localizar y seguir objetos en movimiento en vídeos o en aplicaciones de vigilancia.
3. **Segmentación:** La segmentación de imágenes se refiere a la tarea de dividir una imagen en regiones o segmentos más pequeños con características similares. En lugar de asignar una única etiqueta a toda la imagen, la segmentación permite identificar y delimitar áreas de interés en una imagen. Por ejemplo, en la segmentación semántica, cada píxel de la imagen se etiqueta con una clase correspondiente, lo que permite comprender la estructura y el contenido de la escena de manera más detallada.



Figura 2.11: Ejemplo para las tareas de clasificación, detección y segmentación de imagen en visión computacional.

Las redes neuronales convolucionales (CNN) han logrado un éxito extraordinario en diversos campos de la visión por computadora, incluyendo la segmentación de imágenes. La

segmentación tiene como propósito extraer información relevante de las imágenes para facilitar el análisis. En este caso, los píxeles de la imagen se etiquetan de manera que cada píxel comparte características como color, intensidad, textura, entre otros[39]. En particular, la segmentación semántica extrae información semántica significativa de una imagen, identificando y etiquetando cada píxel según la clase o categoría a la que pertenece [39]. Por ejemplo, en una imagen de una calle, la segmentación semántica puede etiquetar los píxeles correspondientes a automóviles, peatones, edificios, árboles, entre otros (Figura 2.12).

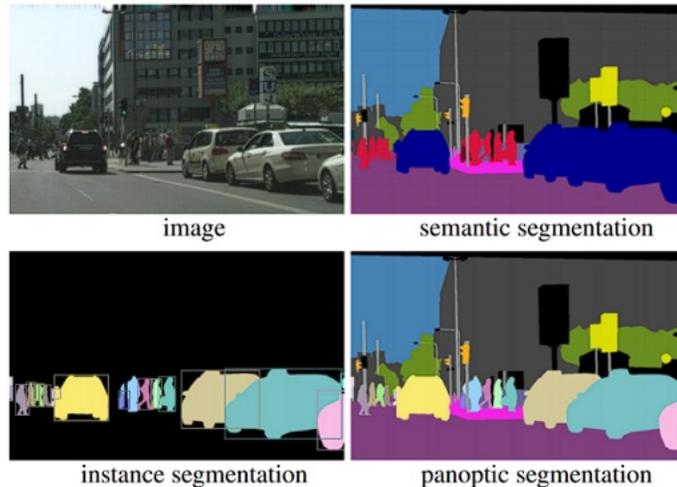


Figura 2.12: Ejemplo de los diferentes tipos de segmentación de imagen [39].

2.3.4. Aprendizaje por Transferencia

El aprendizaje por Transferencia, también conocida como *transfer learning*, tiene como objetivo mejorar el rendimiento de una tarea al aplicar el conocimiento adquirido en otra tarea relacionada, pero distinta. En este proceso, se busca aprovechar la experiencia previa de aprendizaje para mejorar el desempeño en la nueva actividad objetivo [46].

En el ámbito del *deep learning*, la transferencia de conocimiento se refiere a la aplicación de la experiencia adquirida en una tarea específica, utilizando los pesos aprendidos por un modelo como punto de partida para abordar otra tarea relacionada. Cuando se realiza la transferencia de aprendizaje, los pesos preentrenados de un modelo se comparten o inicializan en un nuevo modelo a través de dos enfoques. El primer, se congelan varias capas de la red y entrenan solo las capas finales. En el segundo enfoque, se entrena toda la red mientras se inicializan los pesos con los valores obtenidos del modelo preentrenado.

Este enfoque resulta particularmente beneficioso en situaciones donde se cuenta con conjuntos de datos limitados para la tarea específica que se está abordando. En tales casos, la transferencia de aprendizaje permite aprovechar la información previamente capturada por el modelo en una tarea relacionada, mitigando la escasez de datos y mejorando así la capacidad del modelo para generalizar y desempeñarse eficazmente en la nueva tarea, ahorrando en costos computacionales, pues se reduce el tiempo de entrenamiento y supera limitaciones en tareas que involucran datos escasos, donde la recopilación y etiquetado es costoso o son inac-

cesibles. La aplicación de esta técnica a tareas como clasificación de imágenes, clasificación de texto, la clasificación de sentimiento, entre otras ha sido de gran éxito [41].

2.3.5. Arquitecturas

Arquitectura U-NET

La arquitectura de U-Net se basa en una forma de U invertida, de ahí su nombre (Figura 2.13). Consiste en una ruta de contracción y una ruta de expansión. La ruta de contracción está compuesta por múltiples capas de convolución y *max-pooling*, lo que permite reducir progresivamente la resolución espacial mientras se aumenta la representación de características. Esta ruta captura características contextuales y de alto nivel.

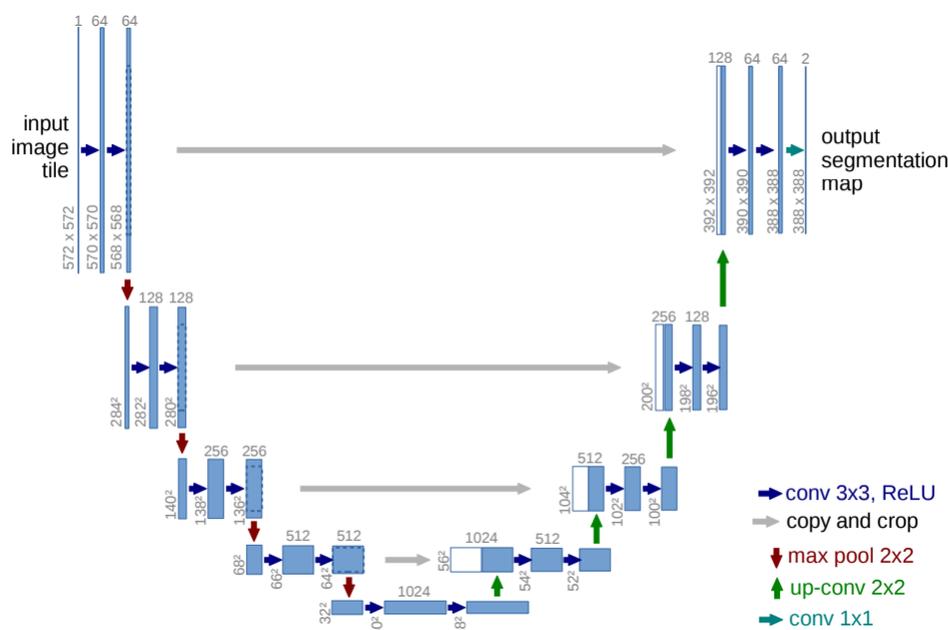


Figura 2.13: Ejemplo de arquitectura U-Net para 32x32 píxeles en la resolución más baja. Cada caja azul corresponde a un mapa de características multicanal. El número de canales se indica en la parte superior de la caja. El tamaño x-y se proporciona en la esquina inferior izquierda de la caja. Las cajas blancas representan mapas de características copiados.[25].

Por otro lado, la ruta de expansión utiliza capas de convolución transpuesta (también conocida como *up-convolución* o *up-sampling*) para aumentar gradualmente la resolución espacial. A medida que la información se propaga en la ruta de expansión, se realiza una concatenación con las características correspondientes de la ruta de contracción para combinar información contextual y espacial en cada nivel de la red.

El diseño de U-Net permite una fusión efectiva de características a diferentes escalas, lo que es crucial para la segmentación semántica [26]. Además, para evitar la pérdida de detalles espaciales en la ruta de expansión, se utiliza una técnica de salto de conexión llamada *skip connections*. Estas conexiones directas entre las capas de contracción y expansión permiten

que la información de baja y alta resolución se combine de manera efectiva, lo que mejora la precisión de la segmentación.

Arquitectura ResNet

ResNet (Redes Neuronales Residuales) es una arquitectura de redes neuronales convolucionales (CNN) que utiliza conexiones residuales en las capas convolucionales. Estas conexiones permiten que la información se propague directamente de una capa a otra, evitando así la degradación del rendimiento que puede ocurrir al agregar capas adicionales a una red profunda [17].

La arquitectura de ResNet está compuesta por bloques residuales que contienen múltiples capas convolucionales. Cada bloque tiene una conexión directa (conexión residual) que agrega la salida de una capa a la salida de otra capa más profunda. Esta conexión permite que la información fluya sin obstáculos a través de la red, facilitando el entrenamiento de redes más profundas y mejorando el rendimiento.

Al utilizar ResNet en la segmentación semántica, se pueden combinar las características extraídas de diferentes niveles de la red. La información contextual y espacial capturada por ResNet puede ser utilizada para identificar y distinguir diferentes objetos y regiones en una imagen, lo que contribuye a una segmentación semántica más precisa.

Arquitectura InceptionResNetV2-UNet

En la arquitectura InceptionResNetV2-UNet, se utiliza una estructura U-Net como base para la red. Esta estructura consta de una ruta de contracción y una ruta de expansión, con conexiones de salto (*skip connections*) que permiten combinar información de diferentes niveles de resolución. La arquitectura, en lugar de utilizar capas convolucionales estándar en la ruta de contracción y expansión, se utilizan bloques residuales basados en InceptionResNetV2. Estos bloques residuales introducen conexiones residuales para evitar la degradación del rendimiento a medida que se agregan capas (Figura 2.14).

Más precisamente, en la ruta de contracción de InceptionResNetV2-UNet utiliza bloques residuales en lugar de capas convolucionales simples. Esto permite que la información fluya a través de la red sin obstáculos, lo que facilita el entrenamiento de redes más profundas y mejora la capacidad de capturar características significativas. Con respecto a la ruta de expansión, se utilizan operaciones de convolución transpuesta (*up-convolution*) para aumentar la resolución espacial. Al igual que en la ruta de contracción, se utilizan bloques residuales para preservar y combinar las características de los diferentes niveles de la red.

La capacidad de preservar información espacial y capturar características profundas permite una segmentación con mejor precisión y detalle de los objetos y regiones presentes en la imagen.

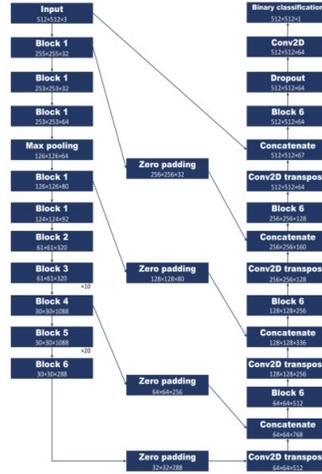


Figura 2.14: Ejemplo de arquitectura InceptionResNetV2-UNet. La estructura de InceptionResNet-UNet utiliza la estructura U-Net tradicional, que consta de partes de codificador (muestreo reducido) y decodificador (muestreo ascendente) con bloques residuales [1].

2.3.6. Métricas

Para medir el desempeño de los modelos, existen varios indicadores. Los más comunes derivan de la matriz de confusión. Se presentan a continuación diferentes tipos de métricas utilizadas para los modelos de segmentación semántica de imágenes.

- **Matriz de confusión:** Esta matriz organiza las predicciones del modelo en cuatro categorías distintas (Tabla 2.1), comparándolas con las clases reales de los datos. La primera clase corresponde a Verdaderos Positivos (TP) aquí se presentan los casos en los que el modelo predijo correctamente la clase positiva. En segundo lugar, la clase Falsos Negativos (FN) indica los casos en los que el modelo falló al predecir la clase positiva. En tercer lugar, la clase de Falsos Positivos (FP) se refiere a los casos en los que el modelo predijo incorrectamente la clase positiva. Por último, la clase Verdaderos Negativos (TN) representa los casos en los que el modelo predijo correctamente la clase negativa.

		Clase Predicha	
		Positiva	Negativa
Clase Real	Positiva	TP	FN
	Negativa	FP	TN

Tabla 2.1: Ejemplo de forma para la matriz de confusión para un problema de clase binaria

La matriz de confusión se utiliza para calcular diversas métricas de evaluación del rendimiento del modelo, como el valor de acierto (*accuracy*), precisión (*precision*), sensibilidad (*recall*), tasa de falsos positivos (*false positive rate*), entre otras.

- **Acierto (Accuracy):** Esta métrica mide el porcentaje de acierto del modelo, frente a todos los casos, en particular para segmentación semántica se llama *Pixel Accuracy*

e informa el porcentaje de píxeles de la imagen que se clasificaron correctamente. Se calcula como:

$$\text{accuracy} = \frac{TP + TN}{P + N} \quad (2.1)$$

donde $P = TP + FN$ y $N = FP + TN$.

- Precisión (Precision): Métrica que muestra qué porcentaje de las predicciones en que el modelo identifica a la clase positiva, pertenecen realmente a esa clase. Esta métrica se calcula mediante:

$$\text{precision} = \frac{TP}{TP + FP} \quad (2.2)$$

- Recall o True Positive Rate: Métrica que muestra qué porcentaje de casos que el modelo logra identificar correctamente como pertenecientes a la clase positiva. Esta métrica se calcula como:

$$\text{recall} = \frac{TP}{P} \quad (2.3)$$

- False Positive Rate: Métrica que muestra qué porcentaje de casos pertenecientes a la clase negativa, el modelo se equivoca. Esta métrica se calcula como:

$$\text{FPR} = \frac{FP}{N} \quad (2.4)$$

- F1-score: Métrica que corresponde a la media armónica entre precisión y recall. Se calcula de la siguiente forma:

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.5)$$

- Índice Jaccard: Esta métrica también conocida como IoU por su ciclo en inglés de *Intersection over Union*. En segmentación semántica es esencial para determinar el porcentaje de superposición entre la máscara objetivo y nuestra salida de predicción. La métrica mide el número de píxeles comunes entre la máscara original y de predicción, todo esto dividido por el número total de píxeles presentes en ambas máscaras.

$$\text{IoU} = \frac{\text{Área de Intersección}}{\text{Área de Unión}} \quad (2.6)$$

Donde el área de intersección es la cantidad de elementos comunes entre los dos conjuntos y el área de unión es la cantidad total de elementos en ambos conjuntos, contados sin duplicados. Cuanto mayor sea el valor de IoU, mejor será la superposición entre las regiones identificadas y las reales. La métrica IoU varía de 0 (sin superposición) a 1 (superposición completa) como se ve en la Figura 2.15.

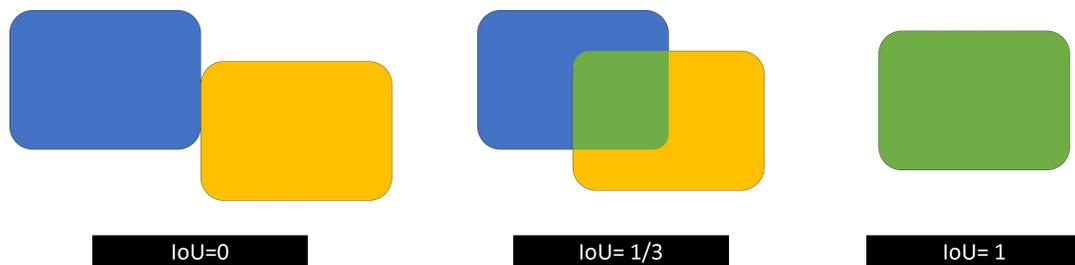


Figura 2.15: Ejemplo de predicción según el valor de IoU, donde en la izquierda, el índice es 0, en el centro es $1/3$ y en la derecha es 1.

2.4. Deep Learning y Detección de Glaciares Rocosos

En el ámbito de visión computacional, el análisis e interpretación automatizada de imágenes de detección remota, tanto satelitales como aéreas, ha ganado considerable interés en los últimos años. Durante este período, se ha abordado una variedad de problemas, incluyendo la detección de recursos naturales, la clasificación del uso del suelo y la segmentación de edificaciones [45].

En esta misma línea, la cartografía de glaciares ha experimentado mejoras significativas gracias al empleo de técnicas de visión computacional. Específicamente, las metodologías utilizadas para la identificación de glaciares descubiertos han demostrado ser altamente exitosas. Al combinar bandas en el espectro visible y no visible de misiones satelitales como Landsat, ASTER o Sentinel-2, junto con otros datos como coherencia topográfica e imágenes SAR, se logra reconocer de manera semiautomática tanto el hielo descubierta como el cubierta de escombros de estas criofomas, permitiendo una delimitación de su forma con un buen nivel de precisión [36]. No obstante, la cartografía de glaciares rocosos como de glaciares cubiertos no ha tenido el mismo éxito. Aunque en los últimos años ha aumentado el interés de catastrar estas formas criogénicas debido a la información que ofrecen para monitorear cambios en el permafrost y con ello tener un indicador del cambio climático actual; como también por la contribución que hacen a los sistemas hidrológicos de algunas cuencas, especialmente en regiones donde se proyecta un aumento en la demanda de agua, pero se anticipa una disminución en las precipitaciones [36]. Si bien la automatización en la cartografía de glaciares rocosos supera obstáculos como la hostilidad del terreno en el que se encuentran estos cuerpos, lo que implica costos económicos y de seguridad significativos en las expediciones, así como desafíos en términos de escalabilidad para abarcar grandes áreas y las variaciones en criterios y tiempos de cartografía asociados con la labor manual de diferentes expertos, también surgen nuevos desafíos debido a la naturaleza de los datos utilizados para su detección de ubicación y forma. [36, 12].

En los últimos años, se han creado diversas metodologías que combinan el *deep learning* con el análisis de imágenes satelitales para abordar diferentes objetivos y desafíos que involucra la cartografía de glaciares rocosos. Dentro las metodologías empleadas, Marcer (2020) introduce innovadoramente un modelo basado en una red neuronal convolucional (CNN), utilizando datos satelitales de uso comercial SPOT6 en escala de grises con una resolución de 2.5 metros. Se selecciona como zona de entrenamiento los Alpes franceses, de donde además se tiene un inventario previo realizado con imágenes de alta resolución, mientras que la vali-

dación se realiza los Alpes italianos donde se hace un inventario de forma manual utilizando las imágenes de SPOT6. El modelo logra identificar exitosamente entre el 60 % y el 70 % de los glaciares de roca de la región (cabe destacar que se realiza un filtro previo de los glaciares rocosos según atributos como tamaño, forma, presencia de sombras, entre otros) [30].

En paralelo, Robson et al. (2020) desarrollaron un modelo mixto que emplea redes neuronales convolucionales (CNN) y Análisis de Imágenes Basado en Objetos (OBIA) con el objetivo de extraer la ubicación y forma de glaciares rocosos en dos áreas de estudio: la cuenca de La Laguna en el norte de Chile y la cuenca de Poiqu en el Himalaya central. En este estudio, se utilizaron datos ópticos de la misión Sentinel-2, que incluían las bandas azul, verde, roja, NIR y SWIR, para generar índices como NDVI, MNDW y SAVI. También se incorporaron datos de coherencia SAR obtenidos de la misión Sentinel-1, así como imágenes de Pléiades, que presentan una resolución de 2 [m] en comparación con la resolución máxima de 10 [m] de Sentinel-2. Además, se emplearon contornos de glaciares de roca previamente catastrados para la generación de datos de entrenamiento. Los resultados revelaron una precisión del productor, definida como el porcentaje total de glaciares de roca correctamente clasificados mediante el método CNN+OBIA, del 75.4 % para la cuenca de La Laguna y del 75.0 % para la cuenca de Poiqu [36].

Por otro lado, Erharder et al. (2021) emplean una metodología que se basa en una arquitectura de red neuronal convolucional (CNN) tipo U-Net para llevar a cabo la segmentación de imágenes y generar un mapa de probabilidad píxel a píxel de la presencia de glaciares rocosos en toda el área de cartografía, correspondiente a los Alpes austríacos. Los datos utilizados incluyen un Modelo Digital de Elevación (DEM) con una resolución de 1×1 [m], ortofotos en escala de grises que se muestrean para lograr una resolución de 2 [m], y datos provenientes del inventario de glaciares rocosos de los Alpes austriaco. Al discretizar los resultados del modelo, se puede concluir que este identifica todos los glaciares en el conjunto de validación, aunque con variaciones en el porcentaje de área detectada. Específicamente, en el 22 % del total de glaciares, solo se logra detectar el 10 % o menos de su extensión. En contraste, para el 35 % del total de glaciares, se logra detectar el 90 % de su respectiva área [12].

Aunque todos los modelos propuestos (Tabla 2.2) han arrojado resultados prometedores, algunos de estos modelos son altamente sensibles a la resolución de las imágenes utilizadas y sus objetivos están dirigidos a respaldar el perfeccionamiento de los inventarios existentes de glaciares rocosos, por sobre la capacidad de identificar posibles nuevos glaciares no catastrados.

Año	Autor	Arquitectura	Zona de Estudio	Input	Resultados
2020	Marcer	CNN:	E: Alpes francés V: Alpes italiano	Imágenes SPOT6 (2.5 [m]) Inventario GR zona Alpes francés e italiano	70.0% de los Alpes italiano respecto al total de GR existentes
2020	Robson et al.	CNN+ OBIA	Cuenca La Laguna (Chile) Cuenca Poiqu (Himalaya)	Ópticos Sentinel -2 (10[m]) Mapa Coherencia SAR Imágenes Pléiade (2 [m]) Polígonos de GR de las zonas de estudio.	75.4% para La Laguna y 75.0% para Poiqu con respecto al total de GR existentes
2021	Erhardter et al.	CNN: U-Net	Alpes austriacos	DEM (1 × 1 [m]) Ortofotos (2 [m]) Inventario de GR de los Alpes austriacos	100 % del total de GR existentes con diferentes porcentajes para el área de extensión.

Tabla 2.2: Resumen de los diversos modelos de aprendizaje profundo desarrollados para la cartografía de glaciares rocosos. Estos modelos han sido aplicados en diferentes contextos geográficos y haciendo uso de diversas fuentes de datos, teniendo buenos resultados en la identificación y cartografía de glaciares rocosos en entornos montañosos.

Capítulo 3

Metodología

La metodología adoptada aborda la evaluación de la sensibilidad del modelo entrenado ante las características proporcionadas por los datos ópticos y los datos derivados de InSAR. Para llevar a cabo esta evaluación, se han preparado tres experimentos específicos:

- **Primer Experimento (Óptico):** Se realiza un entrenamiento exclusivamente con un conjunto de datos compuesto por imágenes ópticas.
- **Segundo Experimento (Conjunto (Óptico+SAR)):** Se entrena el modelo de segmentación semántica de imágenes con imágenes ópticas y datos de velocidad promedio de desplazamiento obtenido mediante InSAR.
- **Tercer Experimento (Conjunto SAR):** Se realiza un entrenamiento exclusivamente con un conjunto de datos formado a partir del mapa de velocidad promedio de desplazamiento obtenido mediante InSAR.

En el capítulo se entrega una descripción del área de estudio, la forma de recopilación de los datos, preprocesamiento, entrenamiento y postprocesamiento para la primera etapa. Además se describe el criterio elegido para corroborar los resultados de la segmentación semántica de imágenes con los datos de velocidad de desplazamiento.

3.1. Área de Estudio

Hasta el momento, en la Región Metropolitana se han registrado un total de 785 criofomas clasificadas como glaciares rocosos. El área de estudio (Figura 3.1) abarca una extensión aproximada de $2750 [km^2]$ (-70.11 a -69.78 ° longitud oeste, -34.04 a -33.24 ° de latitud sur). En esta área se encuentran contenidos un total de 392 glaciares clasificados como glaciares rocosos.

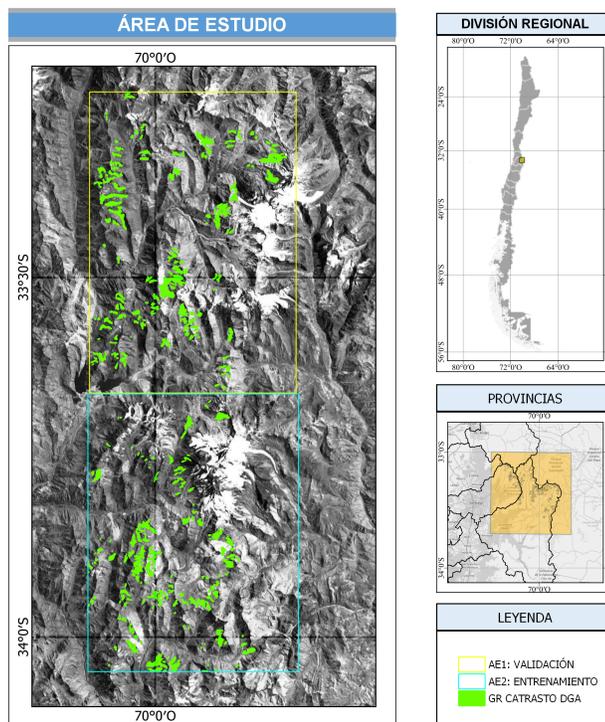


Figura 3.1: Extensión geográfica de la zona de estudio seleccionada. En esta área, se identificaron y catalogaron un total de 392 glaciares clasificados como glaciares rocosos. La región abarca aproximadamente $2750 [km^2]$ y se ubica entre las longitudes -70.11 a -69.78 al oeste, así como las latitudes -34.04 a -33.24 al sur.

3.2. Recopilación de datos

3.2.1. Ópticos

Los datos ópticos fueron adquiridos de la plataforma *Google Earth Engine* basándose en los siguientes criterios de búsqueda: Se eligieron las bandas 2 (Azul), 3 (Verde), 4 (Rojo) y 9 (Infrarrojo cercano - NIR) de la misión Sentinel-2, que ofrece una resolución espacial aproximada de $10 [m]$. Para cada banda, se creó una composición utilizando imágenes disponibles en el período comprendido entre el 1 de febrero de 2022 y el 20 de febrero de 2022 (Figura 3.2). El algoritmo empleó la función *median()* para calcular la mediana de los valores de píxeles correspondientes a las múltiples imágenes dentro del rango de fechas especificado. La elección de estas fechas se fundamentó en la baja presencia de nieve en la cordillera para ese periodo. Por último, para cada banda se seleccionaron las imágenes donde el porcentaje de píxeles de nubes fuera inferior al 10 % y se enmascararon los píxeles con un porcentaje de brillo inferior al 5 % como sombra. Debido a que la plataforma *Google Earth Engine* tiene un máximo de píxeles para las imágenes en descarga. Se dividió en dos polígonos rectangulares para poder bajar los datos de las imágenes. Se estableció denotan las dos áreas como: **área de validación** y **área de entrenamiento** (Figura 3.1).

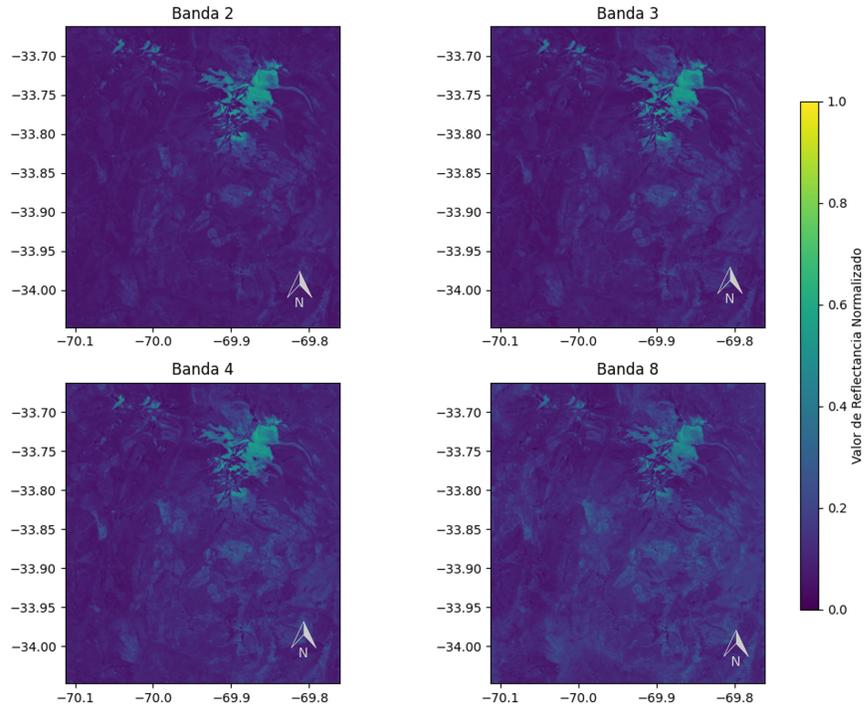


Figura 3.2: En la Figura se presentan las cuatro imágenes ópticas correspondientes a las bandas BGR (Azul - banda 1, Verde- banda 2, Rojo- banda 3) y NIR (Infrarrojo Cercano- banda 1) para el área de entrenamiento. Estas imágenes fueron obtenidas a partir de la misión Sentinel 2 y se utilizaron en el proceso de análisis.

3.2.2. Polígonos

Los polígonos fueron obtenidos del catastro oficial actualizado por la DGA el 2022 [11]. En este catastro se contemplan 8 categorías entregadas en la Tabla 3.1

Número	Clasificación Primaria	Forma
0	Misceláneo	Incierto
1	Hielo continental	Lengua: largo mayor que el ancho
2	Campo de hielo	Lobado: ancho mayor que largo
3	Casquete de hielo	Espatulado
4	Glaciar effluente	Coalescente
5	Glaciar de valle	Otras
6	Glaciar de montaña	
7	Glaciarete y manchón de nieve	
8	Plataforma de hielo	
9	Glaciar rocoso	

Tabla 3.1: Clasificación primaria de glaciares y forma considerada en el Inventario Público de Glaciares de Chile elaborado por la DGA 2022.

Mediante el software QGIS se toma un polígono con la extensión para la zona de estudio, con el fin de poder filtrar solo los polígonos correspondientes, por un lado, a glaciares rocosos,

los cuales se usaron para generar las máscaras binarias para el entrenamiento del modelo; y por otro lado, los polígonos de glaciares de valle que estén dentro de la extensión de la zona de estudio y pueden presentar características similares a las de los glaciares rocosos. Los datos de los polígonos son guardados en GeoJSON debido a que puede incluir atributos no espaciales junto con datos geoespaciales, evitando tener múltiples archivos.

3.2.3. SAR

En general, la información de imágenes satelitales SAR no es abierta y gratuita, salvo en el caso de los satélites Sentinel-1 del Programa Copernicus de la Agencia Espacial Europea (ESA), los cuales además consideran un plan de adquisición de imágenes permanente, facilitando el monitoreo de zonas específicas. Para la realización del proyecto se trabajó con imágenes satelitales de la misión Sentinel-1 A y B cuyas características técnicas son:

- Resolución espacial mínima es de 10 [m].
- Frecuencia de captura de imágenes sobre el mismo punto en promedio de 6 días.
- Banda “C” en modo IW.
- Órbita descendente
- Fecha de toma: 1/01/2020 al 1/03/2022.
- Ruta 83

Preprocesamiento

El preprocesamiento consiste en realizar un control de calidad a las imágenes SLC disponibles. Se generó una selección previa de imágenes para que en los procesos posteriores se logre obtener una buena coherencia entre pares interferométricos y que la información sea confiable. Se seleccionaron imágenes sin mosaico o borde, pues si la zona de interés se encuentra en los bordes de una imagen, es probable que entregue datos con alto ruido lo cual afectará la coherencia entre los pares interferométricos. Del total de imágenes disponibles, quedaron seleccionadas 121 imágenes.

Procesamiento InSAR

Para la generación de los interferogramas, se utilizó el software ISCE2, ya que permite la generación semi-automatizada de grandes cantidades de interferogramas y es de distribución gratuita. Además, su código de fuente puede ser modificado según las necesidades específicas de cada usuario. El software fue programado para abarcar 16 etapas en las cuales ocurren las siguientes acciones:

- Preprocesamiento de datos para la calibración, alineación y eliminación de artefactos o ruido presentes en los datos.

- Corregistro.
- Generación de interferograma.
- Desenvolvimiento de fase.
- Eliminación de fase topográfica.
- Corrección de fase atmosférica.
- Geocodificación.

De los datos SLC iniciales, en la primera etapa de procesamiento se descartaron imágenes debido a que no fueron encontradas sus respectivas órbitas. Quedando con 15 imágenes menos. En los siguientes pasos, el proceso se llevó con éxito. Se obtuvieron un total de 105 interferogramas compuesto por píxeles para los cuales se registró una coherencia por sobre 0.7.

Procesamiento Series de Tiempo

Para la generación de series de tiempo se utilizó el software MintPy, el cual es una herramienta de procesamiento de datos de radar de apertura sintética (SAR) que se emplea en combinación con ISCE2 para crear series de tiempo de interferometría SAR.

En particular, se aplicó la configuración de MitPy para georeferenciar los interferogramas desenvueltos a un mismo píxel; corrigen errores tales como de retrasos troposféricos, DEM, corrección de ionosfera y atmosférica para estimar la velocidad de desplazamiento promedio en la línea de visión de satélite. Ya que tanto ISCE2 como MintPy, permiten el procesamiento de grandes áreas, se procesó la zona de estudio como una sola. Posteriormente se extrajeron los datos con la misma extensión utilizada para la descarga de las imágenes ópticas, obteniéndose finalmente los polígonos de la Figura 3.3.

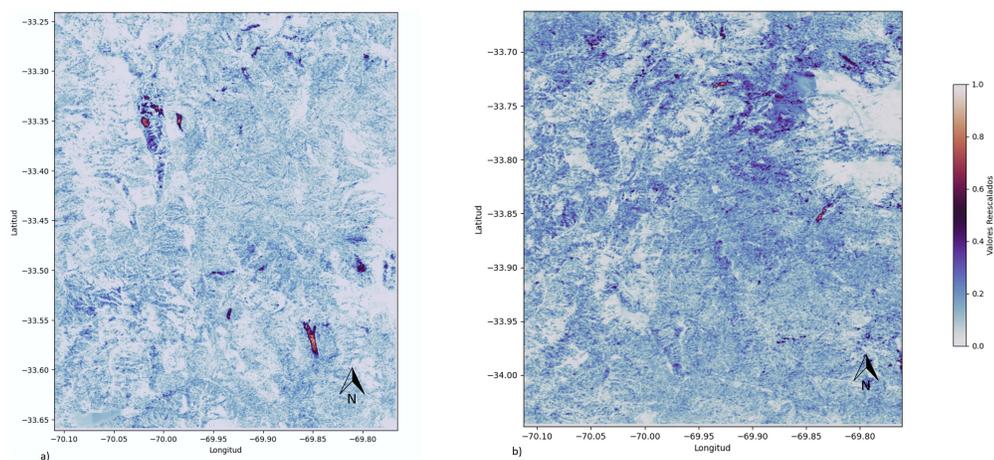


Figura 3.3: Mapa de velocidad promedio generados para el área de estudio. Donde a) mapa para el área de validación y b) mapa para el área de entrenamiento.

3.3. Muestreo

3.3.1. Máscara

Ya determinado los polígonos de glaciares rocosos para la zona de estudio completa se genera una máscara binaria tomando las clases como se expone en la Tabla 3.2:

Nombre	r	g	b
No Glaciar Rocoso (NGR)	255	255	255
Glaciar Rocoso (GR)	0	0	0

Tabla 3.2: Descripción de etiquetas y colores para las dos clases implementadas en el modelo.

3.3.2. Imágenes

En el caso del conjunto de datos ópticos que incluye las bandas RGB+NIR, se llevó a cabo un análisis de componentes principales (PCA) con el objetivo de reducir la dimensionalidad de 4 bandas a 3, conservando las características más relevantes. A partir de este análisis, se generó una nueva imagen que incorpora las tres primeras componentes resultantes del PCA como se muestra en la Figura 3.4.

En el caso del conjunto de datos compuesto por información óptica y SAR, se procedió a normalizar los datos de InSAR. Luego, se redujo la dimensionalidad del conjunto de datos de 5 bandas a 3 bandas utilizando un análisis de componentes principales (PCA) con el objetivo de retener las características más relevantes. Posteriormente, se generó una nueva imagen utilizando las tres primeras componentes proporcionadas por el PCA.

El conjunto de datos que involucra solo datos SAR (los datos de velocidad promedio de desplazamiento) no tuvo modificaciones en esta etapa.

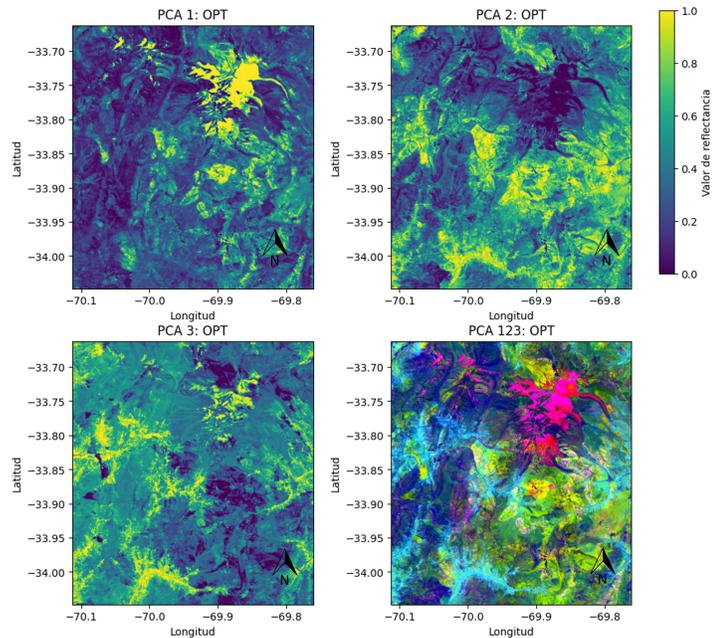


Figura 3.4: PCA para imágenes óptica del área de entrenamiento a utilizar en el primer experimento.

3.4. Dataset

En los tres experimentos propuestos, el conjunto de datos se compone de una serie de parches de imágenes, cada uno con su correspondiente máscara asociada (Figura 3.6). Para generar esta colección de parches junto con sus máscaras respectivas, se aprovechó la georreferenciación de las imágenes. Se implementó una rutina en *Python 3.6.9* que establece una grilla estática sobre la imagen y su máscara correspondiente, utilizando parches de tamaño uniforme de 256 píxeles de ancho y 256 píxeles de alto. Posteriormente, estos parches fueron recortados y tanto la imagen como su máscara se guardaron en formato PNG. Este proceso garantiza la consistencia y estandarización del conjunto de datos utilizado en el análisis sin realizar alteraciones sensoriales.

La Figura 3.5 ilustra la cuadrícula empleada para las imágenes utilizada como entrenamiento y validación. Inicialmente, se generaron 255 parches de imágenes junto con sus máscaras respectivas para el área de estudio de entrenamiento; sin embargo, se observa que alrededor del 60% de las máscaras binarias presentan solo píxeles blancos, por lo cual es altamente probable no encontrar glaciares rocosos. Para mejorar la probabilidad anterior, se toma la decisión de seleccionar únicamente los parches que contienen porciones de glaciares rocosos, identificados por su color azul en la Figura 3.5. Después de aplicar este filtro, se retienen 104 parches. Para el área de estudio de validación se obtuvieron 270 parches, los cuales están representados en la sección a) de la Figura 3.5.

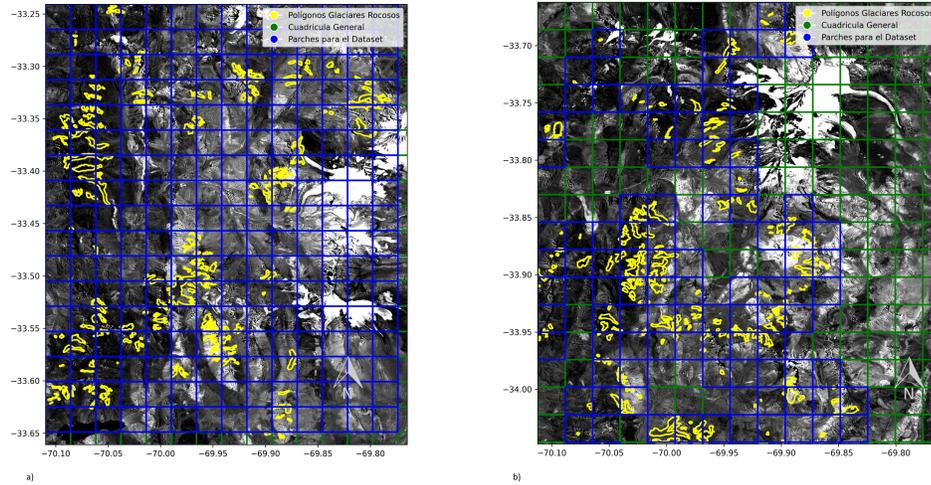


Figura 3.5: Parches para la creación del dataset. En la figura se muestra la grilla creada para obtener los parches de las imágenes provenientes de los datos Ópticos, Ópticos+SAR y SAR en el área de estudio ocupada como entrenamiento y la ocupada como validación. De estos, solo se tomaron los parches marcados en azul, para equilibrar el conjunto de datos de entrenamiento. La imagen a) corresponde a la zona utilizada como validación y la imagen b) la utilizada como entrenamiento.

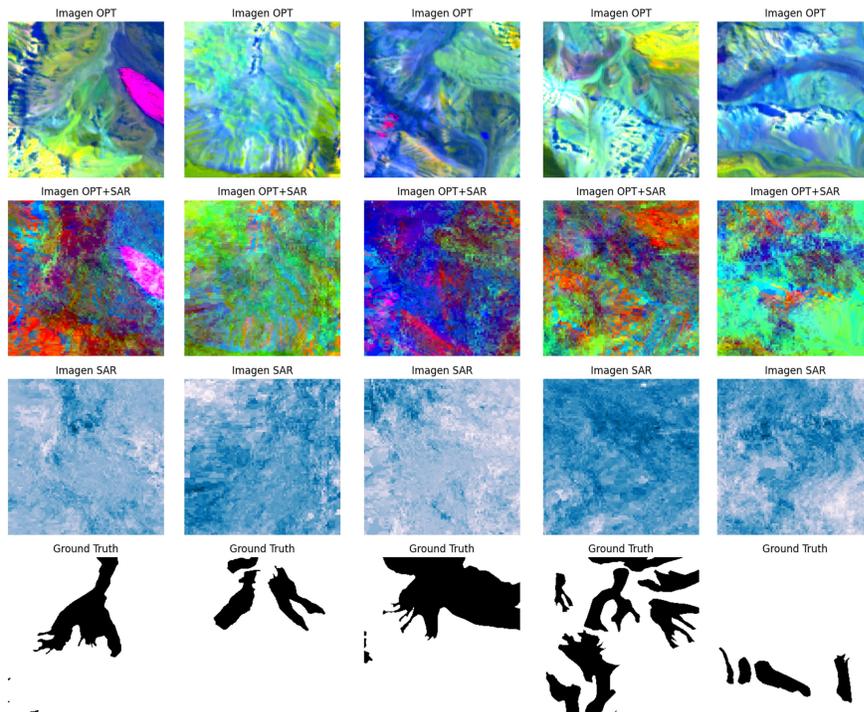


Figura 3.6: Ejemplo de parches (representados en falso color) generados para cada uno de los experimentos con su respectiva máscara.

En el contexto de este conjunto preliminar de entrenamiento de 104 parches, se implementó un proceso de aumento de datos para diversificar la información disponible para el

modelo. Este proceso consistió en aplicar varias transformaciones, incluyendo recortes aleatorios, volteos verticales y horizontales, así como rotaciones aleatorias en un rango entre 60° y 300°. Esta estrategia de aumento resultó en un conjunto de entrenamiento final ampliado a 615 parches de imágenes, cada uno acompañado de sus correspondientes máscaras.

3.5. Modelo, Entrenamiento y Postprocesamiento

El modelo de segmentación semántica implementado en los tres experimentos propuestos se basó en la arquitectura U-Net con bloques InceptionResNetV2 integrados debido a los buenos resultados obtenidos en comparación con otras arquitecturas para la segmentación semántica binaria de imágenes satelitales [37]. En concreto, se utilizó InceptionResNetV2 previamente entrenado con el conjunto de datos ImageNet en la etapa de codificación. Los bloques de convolución fueron configurados con la función de activación *ReLU*, mientras que la capa de salida utilizada se configuró la función de activación *softmax* para generar una distribución de probabilidad sobre las clases para cada píxel. Una descripción general de la arquitectura utilizada se puede revisar en la Tabla 3.3.

Por otro lado, las imágenes de entradas se configuraron como (256, 256, 3) correspondiendo a una imagen RGB de 256x256 píxeles. El tamaño de las máscaras se configuró como (256, 256, 1) siendo esta una imagen binaria, mientras que la salida se configuró como (256, 256, 2), donde 2 indica el número de clases definidas.

	Capa	Tamaño Salida
Entrada		(256, 256, 3)
Preentrenamiento	InceptionResNetV2	
	1	(256, 256, 3)
Codificador	2	(127, 127, 32)
	3	(62, 62, 80)
	4	(29, 29, 256)
Puente	5	(14, 14, 288)
	6	(32, 32, 256)
Decodificador	7	(64, 64, 80)
	8	(128, 128, 32)
	9	(256, 256, 64)
Salida		(256, 256, 2)

Tabla 3.3: Esquema general de la arquitectura InceptionResnetV2 - UNet utilizada para la segmentación semántica de imágenes.

Respecto al ajuste de hiperparámetros, se definieron 20 épocas con un tamaño de lote de 39 para el modelo. En paralelo, se ajustó el mecanismo de detención temprana para cancelar el entrenamiento y así evitar un sobreajuste. A su vez, para controlar la magnitud de los ajustes de los pesos durante el entrenamiento se configuró el optimizador *Adam* con una tasa de aprendizaje en 0.0001 y debido a la naturaleza de la segmentación, se utilizó la función de pérdida *binary crossentropy*. Se evaluó el rendimiento del modelo en cada uno de los experimentos mediante las métricas de F1-Score, IOU y Accuracy.

El modelo se implementó en *Python 3.10* utilizando *TensorFlow* con la librería *Keras* y se entrenó en una GPU NVIDIA T4, disponible en el entorno de *google collaboratory* (Google Colab).

Una vez obtenida las predicciones para las máscaras binarias, en cada experimento, se procedió generar imágenes PNG con los valores de probabilidad. Posteriormente, se llevó a cabo la georreferenciación de los parches para examinar los resultados del entrenamiento en las áreas identificadas como nuevas. Utilizando estos datos, se evaluó la capacidad del modelo para detectar porciones específicas del área con glaciar rocoso bajo diversos umbrales de probabilidad.

Después de comparar los resultados de las áreas predichas para los diferentes umbrales, se consideraron los píxeles con probabilidad mayor a 0.7 y se procedió a generar una salida discretizada. Este paso permitió realizar comparaciones detalladas entre la cantidad de glaciares rocosos detectados según las áreas que tienen intersección con el catastro y no, en vez de revisar el resultado según píxeles detectados.

Capítulo 4

Resultados y Discusión

En esta sección, comenzamos exponiendo la precisión del método de cartografía automatizado. A continuación, exploramos las características de todos los glaciares rocosos cartografiados desde considerando las características topográficas de pendiente y orientación, junto con el área de extensión.

4.1. Desempeño de la cartografía automatizada

Después de completar el proceso de entrenamiento para cada uno de los modelos, en el primer experimento se logró un Índice de Intersección sobre Unión (IOU) de 0.96. Este modelo identificó y delimitó con precisión 115 glaciares rocosos pertenecientes al catastro, generando adicionalmente 21 nuevos polígonos con áreas comprendidas entre 0.1 y 0.6 [km^2].

Modelo	F1-score	Accuracy	IoU
OPT	0.98	0.98	0.96
SAR	0.92	0.94	0.85
OPT_SAR	0.95	0.96	0.90

Tabla 4.1: Resultados del entrenamiento para los experimentos E1 (OPT), E2 (SAR) y E3 (OPT+SAR)

En el segundo experimento, el modelo alcanzó un IOU de 0.85, identificando y delimitando correctamente 77 glaciares rocosos del catastro. Este proceso resultó en la generación de 39 nuevos polígonos, con áreas variando entre 0.1 y 0.9 [km^2].

Por último, el tercer experimento exhibió un IOU de 0.90, identificando y delimitando con precisión 114 glaciares rocosos del catastro. Además, generó 25 nuevos polígonos, con áreas situadas entre 0.1 y 0.3 [km^2].

La Figura 4.1, muestra en el mapa principal los polígonos delimitados en cada uno de

los experimentos junto con algunos ejemplos de la precisión que se obtuvo en el delimitado automatizado con respecto a los polígonos registrados en el inventario de glaciares de la DGA. La delimitación azul corresponde a los resultados obtenido por los datos ópticos (experimento 1), el delineado verde corresponde al registrado por el modelo que ocupa datos procesados de imágenes SAR (experimento 2) y el delineado celeste corresponde a los resultados del modelo que fue entrenado con los datos mixtos (experimento 3). Observamos que para las muestras identificadas los tres modelos logran identificar los glaciares, el modelo entrenado mediante datos de imágenes SAR detecta algunas zonas de la geometría, a diferencia de los otros dos modelos que en varios ejemplos sobrestiman área del glaciar.

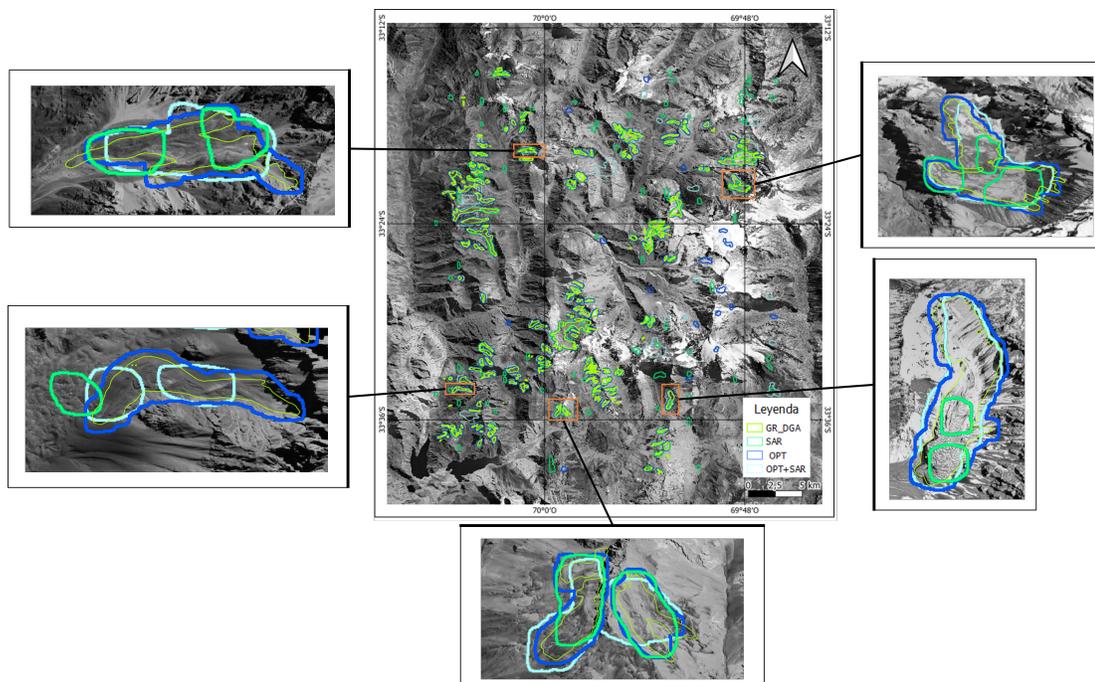
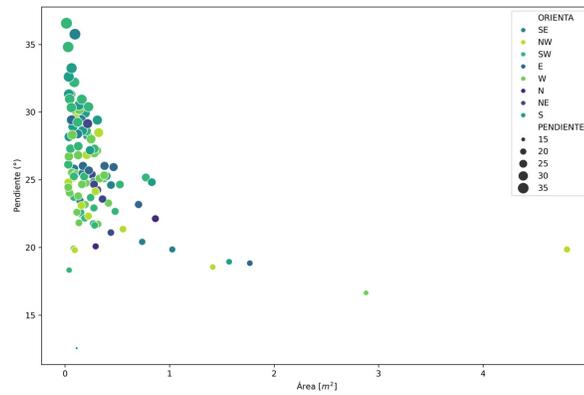


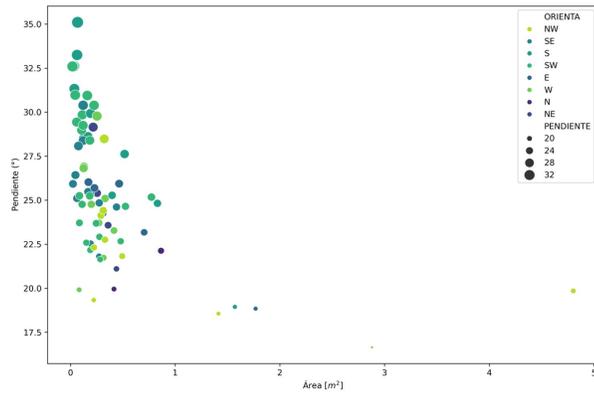
Figura 4.1: Comparación de la cartografía de glaciar rocosos (RG) generada por los modelos entrenados (E1: Azul, E2: Verde y E3: Celeste) y la proveniente del inventario de la DGA (verde lima).

4.2. Área, Pendiente y Orientación

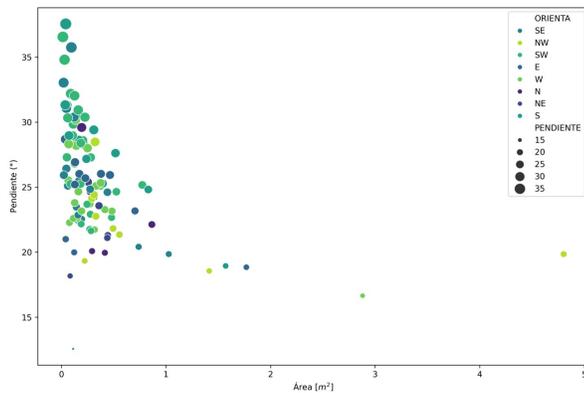
Al comparar los gráficos de dispersión (Figura 4.2), para áreas superiores a $1.0 [km^2]$, se obtienen resultados iguales en los tres experimentos. Sin embargo, en el caso de áreas menores, el primer experimento identifica una cantidad mayor de glaciares pequeños, los cuales en su mayoría presentan una orientación sur (S), sueste (SE) y suroeste (SW).



a)



b)



c)

Figura 4.2: Relación entre área, pendiente y orientación de los glaciares rocosos detectados en cada experimento: a) Resultados del experimento 1 utilizando imágenes ópticas. b) Resultados del experimento 2 basados en imágenes SAR. c) Resultados del experimento 3 utilizando imágenes mixtas.

Analizando la capacidad de detección de glaciares rocosos entre los tres experimentos propuestos con respecto a los glaciares registrado en el inventario. En el primer experimento, se logró identificar un 58.9% del total de glaciares presentes en la zona de estudio, muy similar el segundo experimento alcanzó un 58.4%, y el tercero logró un menor desempeño del 39.4% (Figura 4.3). Observamos que el primer modelo, identifica mejor los glaciares rocosos cuya

área es menor a $0.1 [km^2]$ (Figura 4.4). Por otro lado, para áreas mayores, el tercer experimento es levemente mejor identificando glaciares rocosos. El segundo experimento presentó el peor desempeño con respecto a los otros dos modelos para todos los intervalos de área propuestos.

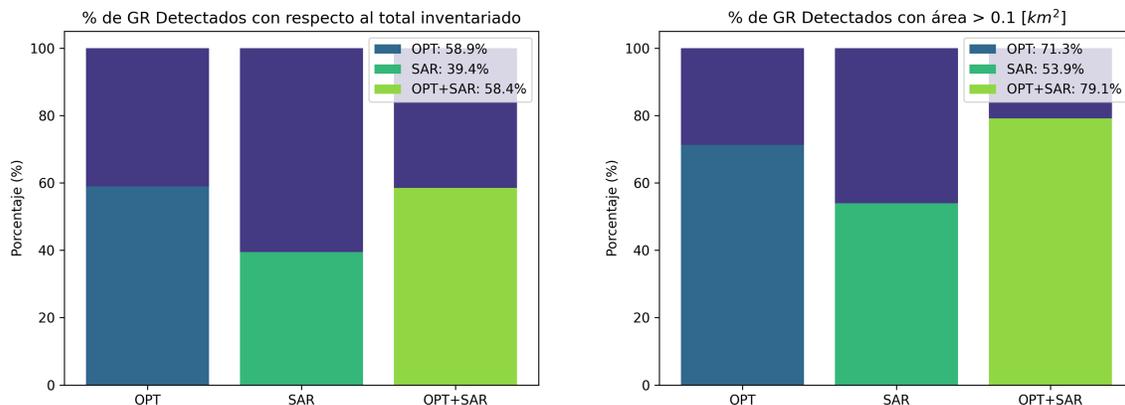


Figura 4.3: Porcentaje de glaciares de roca detectados respecto al primer experimento basado en imágenes ópticas, el segundo basado en datos procesados de imágenes SAR y el tercero, correspondiente datos ópticos y SAR

Al aplicar un filtro basado en el área de los glaciares, se observa un aumento significativo en el porcentaje de identificación (Figura 4.3). En concreto, se obtiene un 71.3 %, 53.9 %, y 79.1 % para el primer, segundo y tercer experimento, respectivamente, al identificar glaciares con una superficie superior a $0.1 [km^2]$.

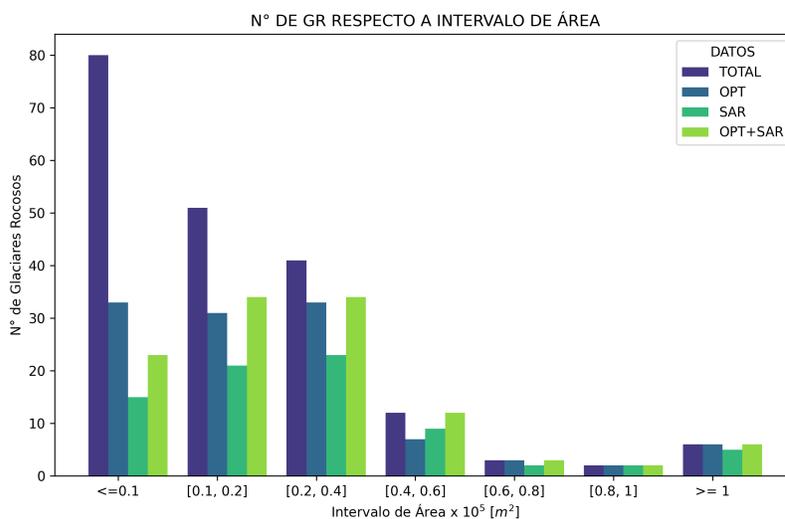


Figura 4.4: Comparación entre los glaciares rocosos detectado por cada experimento según su área.

En cuanto a la pendiente de los glaciares identificados, el primer experimento tiene mejores

resultados para las pendientes superiores a 24° . Mientras que el tercer experimento presenta mejores resultados para pendiente menores. No obstante, la diferencia entre la cantidad de glaciares detectados según sus pendientes para el primer y tercer experimento es baja (Figura 4.5). Sobre el segundo experimento, este tiene los resultados más bajos para todos los ángulos de pendiente registrados.

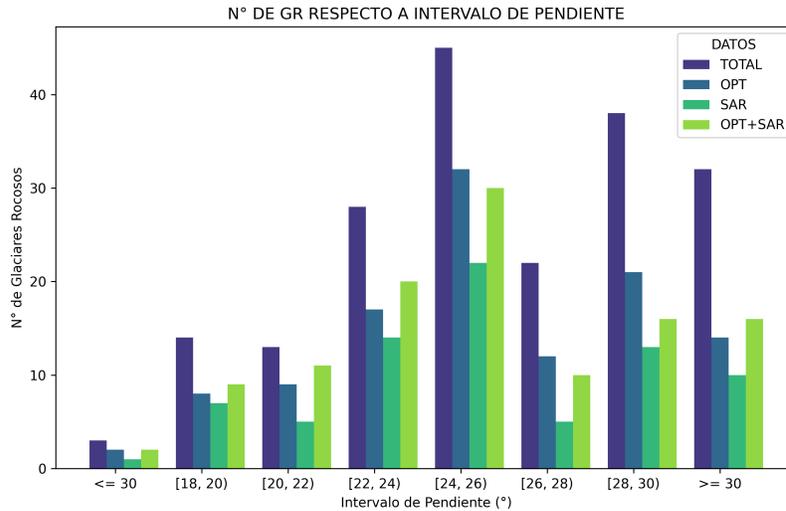


Figura 4.5: Comparación entre los glaciares rocosos detectado por cada experimento según su pendiente.

En otro aspecto, en relación con la orientación, se destaca que en los tres experimentos se logra identificar con un porcentaje superior al 50 % a los glaciares rocosos orientados hacia el noreste (NE), a pesar de lo pequeña que es la muestra (9 GR en total). Sin embargo, el primer experimento exhibe un rendimiento superior al identificar los glaciares en las direcciones sureste (SE: 40 %), sur (S: aprox. 40 %) y suroeste (SW: aprox. 50 %) (Figura 4.6). El segundo experimento, tiene un porcentaje de identificación de GR entre un 10 % a 30 % en las direcciones diferentes a NE, mientras que el tercer experimento varía entre el 10 % al 50 % en las otras orientaciones.

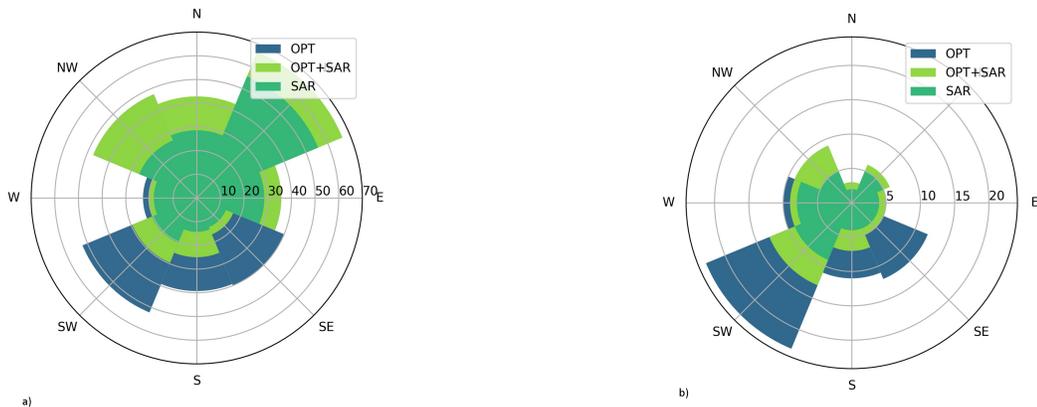


Figura 4.6: Comparación entre los glaciares rocosos detectado por cada experimento según su orientación.

De lo observado, se obtiene que el primer y tercer experimento presentan resultados comparables (Figura 4.2). Encontramos que si bien ambos identifican una cantidad similar de GR, 115 GR y 114 GR para el primer y tercer experimento respectivamente, estos reconocen diferentes glaciares (Figura 4.2). Al comparar los resultados de los glaciares diferentes identificados, se obtiene que el primer experimento como el tercero, reconocen una cantidad similar de glaciares (OPT:37 y OPT+SAR:36). De estos, el primero reconoce mejor de forma independiente los glaciares de área mejor a $0.1 [km^2]$, con orientación S, SW, W y NW; y pendientes mayores a 24° . Por otro lado, el experimento 3 reconoce levemente mejor glaciares de área entre 0.1 y $0.4 [km^2]$; y mucho mejor con áreas ente 0.4 y $0.6 [km^2]$. También se diferencia al reconocer más con orientación N, NE, E y SE, como glaciares con pendiente entre 18° y 24° .

RELACIÓN	OPT	OPT+SAR
INTERSECCIÓN	78	78
DIFERENCIA	37	36
GR TOTAL	115	114

Tabla 4.2: Número de glaciares rocosos distintos detectados por el primer experimento (OPT) y el tercer experimento (OPT+SAR), junto con la cantidad idéntica que identifican.

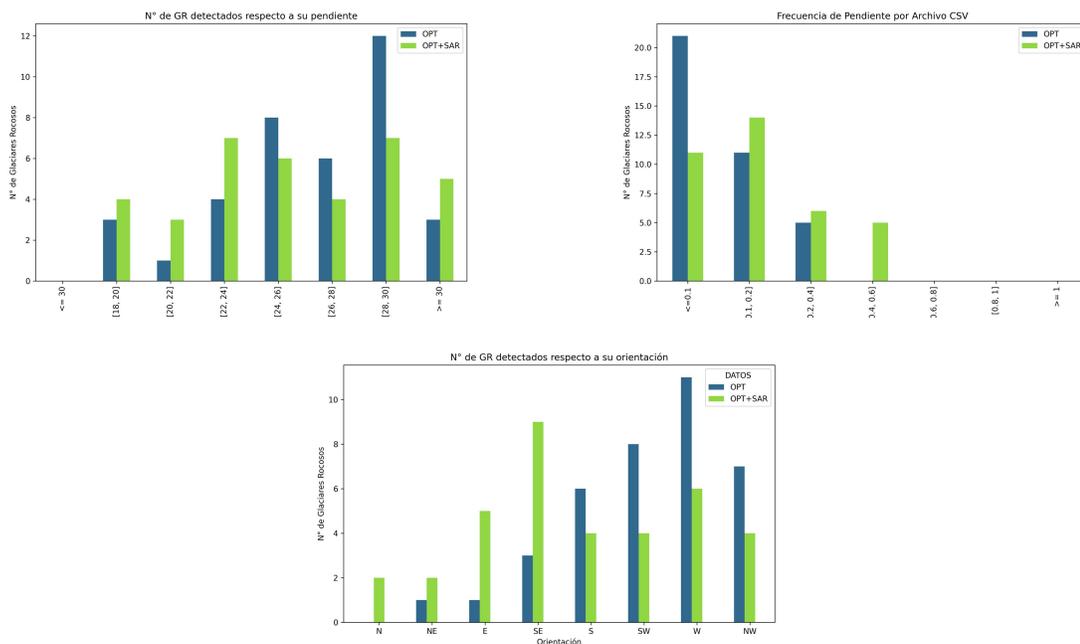


Figura 4.7: Comparación de las características de los glaciares detectados por el primer y tercer experimento.

Capítulo 5

Conclusiones

Los experimentos propuestos, que emplean diversos datos satelitales para entrenar una red mediante *transfer learning*, permiten obtener resultados aceptables en la identificación de glaciares de distintos tamaños. Sin embargo, al aplicar un filtro a los datos, enfocándonos en áreas superiores a $0.1 [km^2]$, se observa una mejora considerable en los resultados de todos los modelos, como es evidenciado por el caso del modelo óptico (OPT: $58.9\% \rightarrow 71.3\%$). Este hallazgo se alinea con resultados documentados en la literatura [36], [12], donde se destaca que la capacidad para detectar glaciares de menor tamaño está estrechamente relacionada con la resolución de las imágenes utilizadas (Sentinel: aproximadamente 10 metros).

Al comparar el desempeño de cada modelo, se observa que los experimentos entrenado solo datos ópticos, como el entrenado con datos mixtos (OPT+SAR), presentan resultados superiores al entrenado únicamente con imágenes SAR. Se observa que al integrar los datos SAR a los datos ópticos no incrementa de manera considerable la cantidad de GR detectados. No obstante, sí afecta el tipo de glaciares de roca identificados. Por lo tanto, ambos modelos proporcionan información complementaria.

Con respecto a los datos SAR, aunque estos por sí solos no exhiben resultados satisfactorios en la detección de glaciares mediante el método propuesto, es posible ampliar su utilidad al aplicar un postprocesamiento. Este postprocesamiento podría ser empleado para aumentar la probabilidad de que las áreas no cartografiadas pero generadas por el modelo sean identificadas como glaciares rocosos activos, ya que las nuevas zonas delimitadas presentan movimiento. Además, este enfoque podría proporcionar información valiosa para incorporar una nueva categoría en los glaciares pertenecientes al inventario, específicamente la distinción entre glaciares activos y no activos.

Por otro lado, al considerar los requerimientos de otros modelos encontrados en la literatura (Tabla 2.2) para la realización de la misma tarea. El modelo basado en datos ópticos es el que mejor se ajusta a las necesidades de operatividad y escalabilidad. Esto se debe a que la mayor parte de los pasos de recopilación de información, como el preprocesamiento, no conllevan altos costos ni económicos como computacionales, sumado a que la generación del conjunto de datos es sencilla. Esto permite obtener resultados preliminares satisfactorios para revisar zonas no cartografiadas en las que solo se disponen de pocos datos satelitales, con el objetivo de cartografiar glaciares rocosos con áreas superiores a $0.1 [km^2]$. Cabe agregar

que debido a su dependencia de la resolución de los datos, se espera que al aumentar esta característica, los resultados mejoren tanto en la detección como en la delimitación de los glaciares.

Aunque los modelos propuestos han demostrado ofrecer buenos resultados para un rango específico de áreas, presentan desafíos. Por un lado, surgen sesgos asociados al desequilibrio de clases, y por otro lado, muestran susceptibilidad a características de las imágenes, tales como el nivel de sombra, la presencia de nieve, y píxeles con morfologías similares pero de naturalezas diferentes, como las remociones en masa, entre otras. Ante estas consideraciones, se sugiere la creación de una base de datos más amplia, no solo en términos de la cantidad de datos relacionados con glaciares rocosos, sino también en la inclusión de nuevas clases, como áreas con glaciares descubiertos, glaciares de valle, detecciones de remociones en masa, etc. Este enfoque tiene como objetivo mejorar la capacidad de generalización del modelo.

Adicionalmente, se propone mejorar la descripción de la precisión del modelo al integrar la diferencia entre el área delimitada por el modelo y el área real, como se sugiere en la investigación de Robson et al. (2020) [36].

La tesis contribuye mediante la creación de una base de datos en formato PNG de la región metropolitana, acompañada de sus respectivas máscaras. Además, se automatiza el proceso de generación de la base de datos, entrenamiento y postprocesamiento para los modelos propuestos. Esto se traduce en el desarrollo de una herramienta que sirve como referencia, tanto para la confirmación de áreas ya cartografiadas como para la exploración de nuevas zonas con una alta probabilidad de ser glaciares rocosos.

Bibliografía

- [1] S. Aghayari, Ahmad Hadavand, Saeed Niazi, and Mohammad Omidalizarandi. Building detection from aerial imagery using inception resnet unet and unet architectures. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-4/W1-2022:9–17, 01 2023.
- [2] Robert S. Anderson, Leif S. Anderson, William H. Armstrong, Matthew W. Rossi, and Sarah E. Crump. Glaciation of alpine valleys: The glacier – debris-covered glacier – rock glacier continuum. Geomorphology, 311:127–142, 6 2018.
- [3] Soumya Balasubramanya and David Stifel. Viewpoint: Water, agriculture y poverty in an era of climate change: Why do we know so little? Food Policy, 93:101905, 2020.
- [4] Francesco Brardinoni, Riccardo Scotti, Rudolf Sailer, and Volkmar Mair. Evaluating sources of uncertainty and variability in rock glacier inventories. Earth Surface Processes and Landforms, 44:2450–2466, 9 2019.
- [5] Alexander Brenning and Guillermo F. Azócar. Minería y glaciares rocosos: impactos ambientales, antecedentes políticos y legales, y perspectivas futuras. Revista de Geografía Norte Grande, pages 143–158, 12 2010.
- [6] Andriy Burkov. The hundred page machine learning. Computer, 2005:414, 1997.
- [7] James B. Campbell and Randolph H. Wynne. Introduction to remote sensing. The Guilford Press, pages 3–30, 2011.
- [8] Jason Charng, Di Xiao, Maryam Mehdizadeh, Mary Attia, Sukanya Arunachalam, Tina Lamey, Jennifer Thompson, Terri McLaren, John Roach, David Mackey, Shaun Frost, and Fred Chen. Deep learning segmentation of hyperautofluorescent fleck lesions in stargardt disease. Scientific Reports, 10:16491, 10 2020.
- [9] J Cogley, Regine Hock, L Rasmussen, Anthony Arendt, A Bauder, R Braithwaite, Peter Jansson, G Kaser, Marco Möller, Lindsey Nicholson, and M. Zemp. Glossary of glacier mass balance and related terms. International Hydrological Programme (IHP) of the United Nations Educational, Scientific and Cultural Organization (UNESCO), 01 2011.
- [10] Flavia A. Croce and Juan P. Milana. Internal structure and behaviour of a rock glacier in the arid andes of argentina. Permafrost and Periglacial Processes, 13:289–299, 10 2002.

- [11] Dirección General de Aguas. Unidad de Glaciología y Nieves. Estrategia nacional de glaciares, 2009.
- [12] Georg H. Erharder, Thomas Wagner, Gerfried Winkler, and Thomas Marcher. Machine learning. an approach for consistent rock glacier mapping and inventorying. example of austria. ApGS, 16:100093, 12 2022.
- [13] WA FISCHER, HEMPHILL WR, and KOVER A. Progress in remote sensing (1972-1976). PROGRESS IN REMOTE SENSING (1972-1976)., 1976.
- [14] F. Francisco. Sobre la distribución de glaciares rocosos en chile, análisis de la situación y reconocimiento de nuevas localizaciones. Investigaciones Geográficas, 54:127–144, 12 2017.
- [15] René D. Garreaud, Juan P. Boisier, Roberto Rondanelli, Aldo Montecinos, Hector H. Sepúlveda, and Daniel Veloso-Aguila. The central chile mega drought (2010–2018): A climate dynamics perspective. International Journal of Climatology, 40:421–439, 1 2020.
- [16] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [17] Kaiming He, X. Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770–778, 2015.
- [18] Yan Hu, Lin Liu, Lingcao Huang, Lin Zhao, Tonghua Wu, Xiaowen Wang, and Jiaxin Cai. Mapping and characterizing rock glaciers in the arid western kunlun mountains supported by insar and deep learning. Journal of Geophysical Research: Earth Surface, 128:e2023JF007206, 9 2023.
- [19] ICEYE. Single look complex image - iceye product documentation.
- [20] Jason R. Janke, Antonio C. Bellisario, and Francisco A. Ferrando. Classification of debris-covered glaciers and rock glaciers in the andes of central chile. Geomorphology, 241:98–121, 7 2015.
- [21] Richard S. Williams Jr. and Jane G. Ferrigno. State of the earth’s cryosphere at the beginning of the 21st century : glaciers, global snow cover, floating ice, and permafrost and periglacial environments: Chapter a in satellite image atlas of glaciers of the world. Professional Paper, 2012.
- [22] Kleanthis Karamvavis and Vassilia Karathanassi. Performance analysis of open source time series insar methods for deformation monitoring over a broader mining region. Remote Sensing, 12:1380, 04 2020.
- [23] Martin P. Kirkbride. Debris-covered glaciers. Encyclopedia of Snow, Ice and Glaciers, pages 190–192, 2011.
- [24] Li. Aplicación de la tecnología insar en peligro de deslizamientos de tierra: Avances y perspectivas. Natl. Sensores remotos Bull, 25(02):614–629, 2021.

- [25] Michelle Livne, Jana Rieger, Orhun Aydin, Abdel Aziz Taha, Ela Akay, Tabea Kossen, Jan Sobesky, John Kelleher, Kristian Hildebrand, Dietmar Frey, and Vince Madai. A u-net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease. Frontiers in Neuroscience, 13, 02 2019.
- [26] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. CoRR, abs/1411.4038, 2014.
- [27] Dionis Lopez Ramos and Leticia Arco Garcia. Aprendizaje profundo para la extraccion de aspectos en opiniones textuales (deep learning for aspect extraction in textual opinions). Revista Cubana de Ciencias Informaticas, 13:105–145, 07 2019.
- [28] Jinna Lv, Qi Shen, Mingzheng Lv, Yiran Li, Lei Shi, and Peiyong Zhang. Deep learning-based semantic segmentation of remote sensing images: a review. Frontiers in Ecology and Evolution, 11:1201125, 7 2023.
- [29] H. Madsen, D. Lawrence, M. Lang, M. Martinkova, and T. R. Kjeldsen. Review of trend analysis and climate change projections of extreme precipitation and floods in europe. Journal of Hydrology, 519:3634–3650, 11 2014.
- [30] Marco Marcer. Rock glaciers automatic mapping using optical imagery and convolutional neural networks. Permafrost and Periglacial Processes, 31:561–566, 10 2020.
- [31] Ravi Teja Nallapu, Yinan Xu, and Jekan Thangavelautham. Spacecraft swarm applications for monitoring active earth features. Advances in the Astronautical Sciences, 01 2021.
- [32] Santanu Pattanayak. Pro deep learning with tensorflow 2.0. Pro Deep Learning with TensorFlow 2.0, 2023.
- [33] Antonio Pepe and Fabiana Calò. A review of interferometric synthetic aperture radar (insar) multi-track approaches for the retrieval of earth’s surface displacements. Applied Sciences, 7(12), 2017.
- [34] Yadu Pokhrel, Farshid Felfelani, Yusuke Satoh, Julien Boulange, Peter Burek, Anne Gädeke, Dieter Gerten, Simon N. Gosling, Manolis Grillakis, Lukas Gudmundsson, Naota Hanasaki, Hyungjun Kim, Aristeidis Koutroulis, Junguo Liu, Lamprini Papadimitriou, Jacob Schewe, Hannes Müller Schmied, Tobias Stacke, Camelia Eliza Telteu, Wim Thiery, Ted Veldkamp, Fang Zhao, and Yoshihide Wada. Global terrestrial water storage and drought severity under climate change. Nature Climate Change 2021 11:3, 11:226–233, 1 2021.
- [35] Andrés Rivera, Claudio Bravo, and Giulia Buob. Climate change and land ice. International Encyclopedia of Geography, pages 1–15, 3 2017.
- [36] Benjamin Aubrey Robson, Tobias Bolch, Shelley MacDonell, Daniel Hölbling, Philipp Rastner, and Nicole Schaffer. Automated detection of rock glaciers using deep learning and object-based image analysis. Remote Sensing of Environment, 250:112033, 12 2020.
- [37] Batuhan Sariturk and Dursun Zafer Seker. A residual-inception u-net (riu-net) approach and comparisons with u-shaped cnn and transformer models for building segmentation from high-resolution satellite images. Sensors, 22(19), 2022.

- [38] Lavanya Sharma and Pradeep Kumar Garg. Artificial intelligence. Artificial Intelligence, 10 2021.
- [39] Farhana Sultana, Abu Sufian, and Paramartha Dutta. Evolution of image segmentation using deep convolutional neural network: A survey. Knowledge-Based Systems, 201-202:106062, 2020.
- [40] Dario Trombotto, Pablo Wainstein, and Likas Arenson. Guía terminológica de la geocriología sudamericana. terminological guide of the south american geocryology", 2014.
- [41] Karl Weiss, Taghi Khoshgoftaar, and DingDing Wang. A survey of transfer learning. Journal of Big Data, 3, 05 2016.
- [42] Gerfried Winkler, Thomas Wagner, and Karl Krainer. Entwässerungsdynamik aktiver blockgletscher-abflusskomponententrennung basierend auf einem multidisziplinären ansatz. Alpine Hydrologie, 2019.
- [43] Jakub Wojkowski, Andrzej Walega, Dariusz Mlynski, Artur Radecki-Pawlik, Tomás Lepeska, Mikolaj Piniewski, and Zbigniew W. Kundzewicz. Are we losing water storage capacity mostly due to climate change – analysis of the landscape hydric potential in selected catchments in east-central europe. Ecological Indicators, 154:110913, 10 2023.
- [44] Lin Yang, Guangju Zhao, Xingmin Mu, Yanli Liu, Peng Tian, Puqiong, and Danzengbandian. Historical and projected evolutions of glaciers in response to climate change in high mountain asia. Environmental Research, 237:117037, 11 2023.
- [45] Wenhui Ye, Wei Zhang, Weimin Lei, Wenchao Zhang, Xinyi Chen, and Yanwen Wang. Remote sensing image instance segmentation network with transformer and multi-scale feature representation. Expert Systems with Applications, 234:121007, 12 2023.
- [46] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. A comprehensive survey on transfer learning, 2020.