



UNIVERSIDAD
NACIONAL
DE COLOMBIA

**Modelo de segmentación semántica de imágenes
satelitales basado en redes neuronales
convolucionales para la clasificación de cobertura
de la tierra en páramos**

Nohora Marcela Reyes Quintana

Universidad Nacional de Colombia
Facultad de Ciencias Agrarias
Bogotá, Colombia
2024

Modelo de segmentación semántica de imágenes satelitales basado en redes neuronales convolucionales para la clasificación de cobertura de la tierra en páramos

Nohora Marcela Reyes Quintana

Trabajo de investigación presentada como requisito parcial para optar al título de:
Magister en Geomática

Director:
(Ph.D.) Iván Alberto Lizarazo Salcedo

Línea de Investigación:
Desarrollo de tecnologías geoespaciales
Grupo de Investigación:
Análisis espacial del territorio y del cambio global (AET-CG)

Universidad Nacional de Colombia
Facultad de Ciencias Agrarias
Bogotá, Colombia
2024

Dedicatoria

Esta investigación está dedicada a mis padres, hermanos y novio, quienes han sido una fuente inagotable de inspiración y motivación para mí. Su apoyo incondicional ha sido la luz que ha iluminado mi camino a lo largo de este viaje académico.

A mis padres, hermanos y novio, con amor y gratitud, este logro es tan suyo como mío.

Agradecimientos

Quiero expresar mi más sincero agradecimiento a todas las personas e instituciones que contribuyeron a la realización de esta investigación y culminación de mi *Maestría en Geomática*. Sus esfuerzos y apoyo han sido fundamentales en cada paso de este arduo pero gratificante camino.

A mi familia, novio y amigos, les agradezco por su constante apoyo emocional y comprensión durante este proceso. Su aliento ha sido mi motivación en los momentos más desafiantes.

Agradezco sinceramente a mi director Iván Alberto Lizarazo Salcedo, mi guía y mentor en este desafío académico. Su experiencia, dedicación y orientación fueron esenciales para dar forma a esta investigación y expandir mi comprensión en el campo de la geomática.

Agradezco a los miembros de mi comité evaluador por su tiempo, revisión crítica y comentarios constructivos. Sus aportes han enriquecido la calidad de esta investigación.

Expreso mi reconocimiento a la Universidad Nacional de Colombia por proporcionar un entorno propicio para el aprendizaje y la investigación en geomática. La calidad de la educación recibida ha sido fundamental para mi desarrollo profesional.

Resumen

Modelo de segmentación semántica de imágenes satelitales basado en redes neuronales convolucionales para la clasificación de cobertura de la tierra en páramos

Los ecosistemas de páramo son cruciales para la regulación hídrica y la biodiversidad, pero enfrenta amenazas significativas debido el cambio climático y las actividades humanas. La falta de delimitación clara y sistemas de monitoreo continuo dificultan su protección, resaltando la necesidad de técnicas de clasificación de coberturas con alta exactitud temática y tiempos de procesamiento cortos. En esta investigación evaluó dos arquitecturas de redes neuronales convolucionales (*CNN*), *U-Net++* y *DeepLabV3* en su capacidad para realizar la segmentación semántica de imágenes satelitales para clasificar las coberturas del complejo de páramos Tota-Bijagual-Mamapacha (TBM), ubicado en Boyacá y Cundinamarca. Utilizando imágenes *Landsat 8* del periodo 2017 a 2019 junto con etiquetas de coberturas del mapa nacional de coberturas 1:100.000 del IDEAM de 2018. Los resultados indican una *exactitud global (OA)* de 68.07% y un *kappa* de 0.60 para *U-Net++* y un OA de 67.29% y un *kappa* de 0.59 para *DeepLabV3+*. En coberturas de páramo, *U-Net++* alcanzó un *F1* del 78.43% para *Herbazal* y del 79.22% para *Bosques*, mientras que *DeepLabV3+* logró un 75% y 74.27%, respectivamente. Ambos modelos se entrenaron en una estación de trabajo equipada con una NVIDIA Quadro RTX™ 5000, presentando tiempos de procesamiento similares. Los hallazgos confirman el potencial de las CNN para la clasificación de coberturas en ecosistemas de páramo y destacan los desafíos asociados al desbalance de clases y la necesidad de datos fuente con alta consistencia lógica y exactitud. Estos resultados establecen una base sólida para investigaciones futuras.

Palabras clave: Teledetección, páramo, redes neuronales convolucionales, segmentación semántica, cobertura del suelo, exactitud temática.

Abstract

Semantic segmentation model of satellite images based on convolutional neural networks for land cover classification in páramos

Paramo ecosystems are fundamental for water regulation and biodiversity but face significant threats due to climate change and human activities. Their protection is hindered by the lack of clear delineation and a continuous monitoring system, highlighting the need to develop land cover classification techniques that offer high thematic accuracy and efficient processing times. This research evaluates two convolutional neural network (CNN) architectures, U-Net++ and DeepLabV3, in their ability to perform semantic segmentation of satellite images to classify the land covers of the Tota-Bijagual-Mamapacha (TBM) paramo complex, located in Boyacá and Cundinamarca. Using Landsat 8 images from the period 2017 to 2019 along with land cover labels from the 2018 national land cover map 1:100,000 of IDEAM. The results indicate an overall accuracy (OA) of 68.07% and a kappa of 0.60 for U-Net++ and an OA of 67.29% and a kappa of 0.59 for DeepLabV3+. In paramo covers, U-Net++ achieved an F1 score of 78.43% for Herbazal and 79.22% for Forests, while DeepLabV3+ achieved 75% and 74.27%, respectively. Both models were efficiently trained on a workstation equipped with an NVIDIA Quadro RTX™ 5000, presenting similar processing times. The findings confirm the potential of CNNs for land cover classification in paramo ecosystems and highlight the challenges associated with class imbalance and the need for source data with high logical consistency and accuracy. These results provide a solid foundation for future research in land cover classification and ecosystem monitoring.

Keywords: *Remote Sensing, páramo, convolutional neural networks, semantic segmentation, land cover, thematic accuracy.*

Contenido

	Pág.
Resumen	VI
Lista de figuras	X
Lista de tablas	XIII
1. INTRODUCCIÓN	15
1.1 Antecedentes.....	15
1.2 Planteamiento del problema	16
1.3 Aproximación conceptual y metodológica.....	18
1.4 Pregunta de investigación	21
1.5 Objetivos	22
1.5.1 Objetivos específicos.....	22
1.6 Justificación	22
1.7 Organización del documento.....	23
2. MARCO TEÓRICO Y ESTADO DEL ARTE.....	25
2.1 Marco teórico	25
2.1.1 Ecosistema de páramos	25
2.1.2 Clasificación de cobertura de la tierra en páramos.....	27
2.1.3 Red neuronal convolucional (CNN)	27
2.1.4 Evaluación de exactitud.....	33
2.2 Estado del arte.....	35
2.2.1 U-Net.....	36
2.2.2 DeeplabV3+.....	38
2.2.3 Retos en la aplicación de clasificación de coberturas en el análisis de imágenes de teledetección.....	40
3. MATERIALES Y MÉTODOS	41
3.1 Área de estudio.....	41
3.2 Datos.....	44
3.2.1 Imágenes satelitales	44
3.2.2 Modelo digital de elevación (DEM)	46
3.2.3 Etiquetas de referencia.	47
3.3 Métodos.....	51
3.3.1 Preprocesamiento (PRE).....	51
3.3.2 Procesamiento (PRO).....	55

3.3.3	Análisis (<i>ANA</i>).....	60
3.3.4	Herramientas	61
4.	RESULTADOS	63
4.1	Selección de Hiperparámetros	63
4.2	Mitigación del desbalance de clases.....	65
4.3	Definición del número de bandas.....	70
4.4	Selección del modelo	72
5.	DISCUSIÓN	81
5.1	Preguntas de investigación	81
5.1.1	¿Cómo abordar el desbalance de las diferentes clases de coberturas de páramos presentes en el área de estudio?	82
5.1.2	¿Cuál es el impacto de incrementar el número de bandas utilizadas como variables explicativas en la exactitud temática de los modelos?	84
5.1.3	¿Cuáles coberturas de páramos muestran una segmentación de mayor exactitud temática mediante el uso de los modelos U-Net++ y DeepLabv3+?	86
5.2	Aportes a la geomática.....	87
5.3	Ventajas y limitaciones de esta investigación	89
5.3.1	Ventajas:.....	89
5.3.2	Limitaciones:.....	89
6.	CONCLUSIONES	91
A.	Anexo: Clasificación de las áreas de estudio.....	93
B.	Anexo: Glosario.....	97
7.	BIBLIOGRAFÍA.....	99

Lista de figuras

	Pág.
Figura 1-1: Coberturas naturales más representativas en los páramos del estudio del IDEAM et al., (2014).	18
Figura 1-2: Una red neuronal artificial emula la red de neuronas de un cerebro. Esta incluye la capa de entrada, capas ocultas y la capa de salida que se conectan entre sí con diferentes pesos. Tomado de Davins Jovells, (2021).....	19
Figura 1-3: Segmentación semántica de la (a) Composición <i>RGB</i> , (b) imagen clasificada píxel a píxel asignándole una clase. Tomada de Hoerer & Kuenzer, (2020).....	20
Figura 1-4: Una red de <i>CNN</i> se divide en dos etapas: codificación, que contiene la capa de convolucional, agrupamiento y exclusión; y decodificación, que comprende la capa de salto, convolucional transpuesta y la función de pérdida. Tomado de Martin-Abadal et al., (2018).	21
Figura 2-1: Las coberturas más representativas del páramo TBM son <i>Herbazales</i> , <i>Áreas agrícolas heterogéneas</i> y <i>Bosques</i> . Elaboración Propia basada en el estudio del (IDEAM et al., 2014).	27
Figura 2-2: La arquitectura <i>CNN</i> está compuesta por la capa convolucional, capa de agrupamiento y la capa de clasificación. Adaptado de Altamirano (2021).....	28
Figura 2-3: La capa convolucional procesa una imagen de entrada para extraer características importantes. Adaptado de Igar (2020).....	29
Figura 2-4: La capa de agrupación, simplifican la información reduciendo la resolución espacial de las características extraídas. Tomada de Igar, (2020).....	30
Figura 2-5: Función de activación <i>Sigmoid</i> . Adaptado de Torralba, (2022).	30
Figura 2-6: Función de activación <i>Tanh</i> . Adaptado de Torralba, (2022).	31
Figura 2-7: Función de activación <i>ReLU</i> . Adaptado de Torralba, (2022).	31
Figura 2-8: Ajuste del resultado del modelo (a) Subajustado, (b) Sobreajuste y (c) Balanceado. Tomado de Alzubaidi et al., (2021).	32
Figura 2-9: En la arquitectura <i>U-Net</i> , los cuadros azules simbolizan los mapas de características, con el número de canales en la parte superior y las dimensiones x-y en la parte inferior izquierda. Los cuadros blancos representan copias de los mapas de características. Las operaciones se identifican mediante flechas que los conectan. Adaptado de Navab et al., (2015).	37
Figura 2-10: Comparación entre <i>Max-pooling</i> y <i>Max-unpooling</i> . Tomada de Altamirano P, (2021).	38

Figura 2-11	<i>Kernel 3 x 3 con rate = 1, rate = 2 y un rate = 3. Tomada de Chen et al., (2018).</i>	38
Figura 2-12:	Arquitectura <i>DeepLabV3 +</i> , mostrando las fases del codificador y decodificador con sus componentes claves, incluyendo la Convolución <i>Atrous</i> y la Agrupación de Pirámides Espaciales <i>Atrous</i> . Adaptado de (Liu et al., 2021).	39
Figura 3-1:	Principales Coberturas presentes en el páramo TBM: (a) <i>Rastrojos y Pastos</i> ; (b) <i>Herbazales</i> . Tomado de lapensadera.com (2021).	42
Figura 3-2:	Área de estudio se muestra en color rojo. Fuente de datos MADS e IGAC, composición propia.	44
Figura 3-3:	Composición <i>RGB</i> en color natural del área de estudio capturada por el satélite <i>Landsat 8</i> , utilizando la combinación de bandas 4, 3 y 2. Esta representación visual permite una observación de la superficie terrestre similar a la visión humana. Fuente de datos USGS, composición propia.	45
Figura 3-4:	DEM del área de estudio. Fuente de datos NASA/USGS/JPL-Caltech, composición propia.	46
Figura 3-5	Coberturas de la tierra en el Área de estudio. Elaboración Propia basada en el estudio del IDEAM et al., (2021).	47
Figura 3-6:	Mapa de referencia del que extraen las Etiquetas del área de estudio con los colores <i>RGB</i> para el entrenamiento. Fuente de datos IDEAM, composición propia	50
Figura 3-7:	Diagrama de la metodología general de la investigación. Fuente propia.	51
Figura 3-8:	Diagrama de la metodología detallada de la investigación. Fuente propia.	52
Figura 3-9:	Distribución de las grillas: (a) Imagen de 256*256 píxeles en falso color <i>RGB-543</i> . (b) Etiquetas <i>CLC 2018</i> de 256*256 píxeles. (c) Imagen de 128*128 píxeles en verdadero color <i>RGB-432</i> . (d) Etiquetas <i>CLC 2018</i> de 128*128 píxeles.	55
Figura 4-1.	Resultados de la clasificación de <i>U-Net++_4_4</i> con la imagen <i>img_432</i> . (a) Imagen, (b) Etiqueta y (c) Inferencia. Fuente propia.	66
Figura 4-2:	Resultados de la clasificación del modelo <i>DLV3+_4_3</i> con la imagen <i>img_432</i> . (a) Imagen, (b)Etiqueta y (c)Inferencias. Fuente propia.	68
Figura 4-3	<i>F1</i> de las coberturas representativas usando los modelos <i>U-Net++_4</i> y <i>DLV3+_4</i> . Fuente propia.	69
Figura 4-4:	Métricas de las coberturas del modelo <i>U-Net++_4_4</i> (a) <i>F1</i> y (b) <i>IoU</i> . En el eje x se indican los códigos de las diferentes coberturas. Fuente propia.	71
Figura 4-5:	Métricas de las coberturas del modelo <i>DLV3+_4_3</i> (a) <i>F1</i> y (b) <i>IoU</i> . En el eje x se indican los códigos de las diferentes coberturas. Fuente propia.	72
Figura 4-6:	Comparación del entrenamiento y la validación (a) Exactitud (b) Entropía cruzada ponderada para <i>U-NET++_4_4</i> . Fuente propia	74
Figura 4-7:	Comparación del entrenamiento y la validación (a) Exactitud (b) Entropía cruzada ponderada para <i>DLV3+_4_3</i> . Fuente propia	75
Figura 4-8:	Matriz de confusión del modelo (a) <i>U-Net++_4_4</i> (b) <i>DLV3+_4_3</i> . Fuente propia.	76
Figura 4-9:	Resultados de la clasificación del modelo <i>U-Net++_4_4</i> . (a) Imagen, (b)Etiqueta y (c) Inferencia. Fuente propia.	78

Figura 4-10:	Resultados de la clasificación del modelo <i>DLV3+_4_3</i> . (a) Imagen, (b) Etiqueta y (c) Inferencia. Fuente propia.	79
Figura 5-1:	Representatividad de las coberturas en el área de estudio. Fuente propia.	83
Figura 5-2:	Representatividad de las coberturas vs F1 de los resultados. (a) Modelo <i>U-Net++_4</i> y (b) Modelo <i>DLV3+_4_3</i> . Fuente propia.	84
Figura 5-3:	Comparación de las métricas generales de (a) <i>OA</i> y (b) <i>kappa de U-Net++_4</i> y <i>DLV3+_4_3</i> . Fuente propia.	85
Figura 5-4:	Comparación de las métricas por coberturas de (a) <i>F1</i> y <i>IoU de U-Net++_4</i> y de (b) <i>F1</i> y <i>IoU de DLV3+_4_3</i> . Fuente propia.	87
Figura 6-1:	Etiquetas del área de estudio con los colores <i>RGB</i> para el entrenamiento. Adapta del IDEAM et al., (2021)	93
Figura 6-2:	Resultados de la clasificación del modelo <i>U-Net++_4</i> . Fuente propia.....	94
Figura 6-3:	Resultados de la clasificación del modelo <i>DLV3+_4_3</i> . Fuente propia.....	95

Lista de tablas

	Pág.
Tabla 2-1. Matriz de error con los tipos de clasificaciones.....	33
Tabla 3-1. Características principales de los páramos más importantes de Colombia.....	43
Tabla 3-2. Información de las Bandas <i>OLI</i> de <i>Landsat 8</i> . Fuente <i>GEE</i> , (2007).....	45
Tabla 3-3. Información de las Bandas de <i>Landsat 8</i> . Fuente (IDEAM, 2010).....	48
Tabla 3-4. Imágenes utilizadas en el entrenamiento, validación y pruebas. Fuente propia.....	53
Tabla 3-5. Coberturas, nivel de representatividad y el color asignado. Fuente propia.....	54
Tabla 3-6. Hiperparámetros usados para la segmentación semántica de imágenes. Fuente propia.....	57
Tabla 3-7. Herramientas y software y sus propósitos respectivos. Fuente propia.....	61
Tabla 4-1. Hiperparámetros de <i>U-Net++</i> y <i>DeepLabV3+</i> . En color verde se resaltan los modelos seleccionados. Fuente propia.....	64
Tabla 4-2. Métricas de las coberturas de <i>U-Net++-4</i> y <i>DLV3+-4</i> . Fuente propia.....	64
Tabla 4-3. Métricas generales de las iteraciones de los pesos de las coberturas en el modelo <i>U-Net++-4</i> . Fuente propia.....	65
Tabla 4-4. Métricas por coberturas de las iteraciones de los pesos de las coberturas en el modelo <i>U-Net++-4</i> . Fuente propia.....	66
Tabla 4-5. Métricas generales de las iteraciones de los pesos de las coberturas en el modelo <i>DLV3+-4</i> . Fuente propia.....	67
Tabla 4-6. Métricas por coberturas de la iteración de los pesos de las coberturas en el modelo <i>DLV3+-4</i> . Fuente propia.....	67
Tabla 4-7. Pesos asignados a las diferentes coberturas para afrontar el desbalance de clases. Fuente propia.....	69
Tabla 4-8. Desempeño computacional de <i>U-Net++-4-4</i> y <i>DLV3+-4-3</i> utilizando diferentes imágenes. Fuente propia.....	72
Tabla 4-9. Comparación de las métricas de <i>U-Net++-4-4</i> y <i>DLV3+-4-3</i> usando la imagen <i>img_765432NPE</i> . Fuente propia.....	73
Tabla 4-10. Métricas por coberturas de <i>U-Net++-4-4</i> y <i>DLV3+-4-3</i> con la imagen <i>img_765432NPE</i> . Fuente propia.....	73
Tabla 4-11. Desempeño computacional de <i>U-Net++-4-4</i> y <i>DLV3+-4-3</i> . Fuente propia.....	80
Tabla 4-12. Comparación del desempeño computacional entre el área de estudio y Colombia. Fuente propia.....	80

1. INTRODUCCIÓN

1.1 Antecedentes

Los páramos son ecosistemas tropicales de alta montaña cruciales para la regulación hídrica y la conservación de la biodiversidad (Brück et al., 2023; Vásquez, A., Buitrago, 2011), situados entre el límite superior del bosque andino y el límite inferior de los glaciares según el MADS, (2002). Ubicados principalmente en la cordillera de los Andes, se extienden desde Venezuela hasta Colombia y también están presentes en la Sierra Nevada de Santa Marta y Costa Rica (Garavito, 2016). Los páramos se caracterizan por sus bajas temperaturas y alta humedad debido a las frecuentes lluvias y neblina (Gómez Sánchez et al., 2002). Estos ecosistemas, abarcan aproximadamente tres millones de hectáreas en Colombia, que representan alrededor del 50% de los páramos del mundo (MADS, 2021). Su importancia radica en sus funciones ecológicas claves como la captura de carbono atmosférico y la regulación del ciclo del agua (Brück et al., 2023).

A pesar de su importancia, los páramos enfrentan amenazas significativas por actividades humanas, lo que ha impulsado la necesidad de métodos más efectivos para su clasificación y monitoreo (Brück et al., 2023). Tradicionalmente en Colombia, la clasificación de coberturas de la tierra se ha llevado a cabo mediante técnicas de teledetección basadas en la interpretación visual a cargo de entidades como el Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM et al., 2021). Sin embargo, los avances en teledetección y el incremento en la capacidad de procesamiento de datos han fomentado la adopción de técnicas más sofisticadas como el aprendizaje profundo particularmente las redes neuronales convolucionales (*CNN*).

Las *CNN* conocidas por su estructura de múltiples capas¹, y su capacidad para procesar eficientemente datos estructurados en cuadrícula como las imágenes, representan una mejora significativa sobre los métodos tradicionales debido a su exactitud y eficiencia en tiempo en la clasificación de coberturas (Yuan et al., 2021). Investigaciones recientes han empezado a evaluar las *CNN* para la clasificación de coberturas de páramos, enfocándose en su aplicación para la conservación y manejo sostenible de estos ecosistema (Castelo-Cabay et al., 2022). La implementación de estas tecnologías plantea mejorar la gestión ambiental y facilitar la toma de decisiones, basada en información con buena exactitud y actualizada.

1.2 Planteamiento del problema

Los páramos son ecosistemas de alta montaña de vital importancia ecológica que enfrentan una creciente amenaza debido a prácticas dañinas como la agricultura intensiva, la ganadería, la explotación minera y la tala indiscriminada, que han llevado a una degradación ambiental y socioeconómica considerables (Brück et al., 2023). Aunque existen regulaciones como la Ley 1930 de 2018 emitida el Ministerio de Ambiente y Desarrollo Sostenible (MADS) y la Resolución 1294 de 2021, del MADS y el Ministerio de Agricultura y Desarrollo Rural (MADR), que buscan proteger estos ecosistemas y establecer lineamientos para el manejo sostenible de actividades de alto impacto, los desafíos persisten

En Boyacá, más del 50% de los páramos estas siendo degradados por actividades humanas, como la minería (IGAC, 2015), que introduce sustancias tóxicas como mercurio y cianuro a sus fuentes de agua. Un ejemplo de esto es el otorgamiento del título minero GEI-151 por la Agencia Nacional de Minería (ANM) el 19 de noviembre de 2020, dentro del complejo de páramos Tota-Bijagual-Mamapacha (TBM) (Retavisca G., 2021). Investigaciones como Garavito, (2016) y Loaiza-Usuga et al., (2021) resaltan la urgente necesidad de fortalecer las políticas de conservación para desarrollar estrategias más efectivas que permitan proteger los páramos de las amenazas antropogénicas. Mostrando la importancia de reforzar las medidas de protección

¹ Las capas en el contexto de CNN son componentes claves que estructuran la red y definen el procesamiento de datos para tareas como clasificación, detección de objetos y segmentación de imágenes. Cada capa ejecuta operaciones específicas, facilitando el aprendizaje de características jerárquicas a partir de los datos de entrada.

actuales para garantizar la conservación de estos ecosistemas críticos para el equilibrio ambiental de Colombia.

El IDEAM ha utilizado métodos tradicionales de interpretación visual de imágenes para producir mapas nacionales de coberturas del suelo con alta exactitud temática, pero algunos autores, como Suárez Londoño et al., (2017), han sugerido que estos métodos pueden ser costosos y dispendiosos. Además, la frecuencia con la que se actualizan las coberturas es relativamente baja, con una periodicidad media de aproximadamente cuatro años. Esta situación se refleja en las fechas de publicación de los mapas de coberturas de la tierra, como el Mapa de coberturas de la tierra línea base (2000-2002), Mapa de coberturas de la tierra periodo (2005-2009), Mapa de coberturas de la tierra periodo (2010-2012), Coberturas de la tierra en zonas de páramo priorizados y el Mapa Nacional de Coberturas de la Tierra, Periodo 2018 (IDEAM, 2015; IDEAM et al., 2014, 2021).

En respuesta a la necesidad de métodos eficientes y actualizaciones más frecuentes, se ha explorado técnicas de clasificación digital. Métodos como los bosques aleatorios (*RF*), red neuronal artificial (*ANN*) y máquinas de soporte vectorial (*SVM*) han demostrado ser efectivos en la clasificación de coberturas y podrían ofrecer soluciones ágiles con alta exactitud temática (Talukdar et al., 2020). Un ejemplo es el estudio de Ayala-Izurieta et al., (2017), donde se empleó *RF* para la clasificación de alta montaña en los Andes ecuatorianos, logrando una *OA* del 95%.

En el contexto colombiano, diversos estudios han evaluado técnicas de clasificación digital, incluyendo los trabajos de Coca Castro et al., (2021); y Espejo O., (2017). Para coberturas de páramos, el estudio realizado por Poveda-Sotelo et al., (2022), exploró diferentes métodos de clasificación en los páramos de Merchán y Telecom, este estudio reveló que las técnicas de *RF*, *ANN* y *SVM* proporcionaron resultados muy prometedores, con un Kappa de 92.53%, 91.51% y 91.32% respectivamente.

A pesar de estos resultados, es relevante explorar técnicas de clasificación digital para evaluar su desempeño en las coberturas de páramos en Colombia. Por esta razón, se ha considerado el potencial de las técnicas de aprendizaje profundo, en particular las *CNN*, que han demostrado ser eficientes en la clasificación de coberturas. La exploración de diversas arquitecturas de *CNN*, respaldada por diversos estudios previos como los de Arcos Serrano, (2021); Kussul et al., (2017); Martin-Abadal et al., (2018) sugieren un enfoque prometedor.

Esta investigación buscó profundizar en el uso de *CNN* para mejorar la clasificación de coberturas en los páramos colombianos, y facilitar un monitoreo más frecuente, para contribuir a la protección de estos ecosistemas vitales. También se abordaron los desafíos asociados con este método, incluyendo la clasificación de áreas de transición en los subpáramos, la escasez de etiquetas de alta calidad y actualizadas, además de la necesidad de disponer de un conjunto de datos fuentes adecuados, con las correspondencia tanto de etiquetas como de imágenes, para todas las coberturas de los páramos.

1.3 Aproximación conceptual y metodológica

Esta investigación se centró en dos aspectos fundamentales: (i) la clasificación de coberturas en páramos y (ii) la segmentación semántica mediante técnicas basadas en aprendizaje profundo. La clasificación de coberturas implica identificar y categorizar elementos como vegetación, construcciones humanas, áreas rocosas, suelos expuestos y cuerpos de agua (Di Gregorio, 2016).

El IDEAM ha proporcionado datos de coberturas a escala 1:100.000 para los periodos 2000-2002, 2005-2009, 2010-2012, 2018 (IDEAM, 2015; IDEAM et al., 2014, 2021). Además, para los páramos colombianos, se dispone de una clasificación de coberturas realizada en 2014, que abarco 21 complejos y fue realizada en colaboración entre el IDEAM y el Instituto de Investigación de Recursos Biológicos Alexander von Humboldt (IAVH), a escala 1:25.000 (IDEAM et al., 2014).

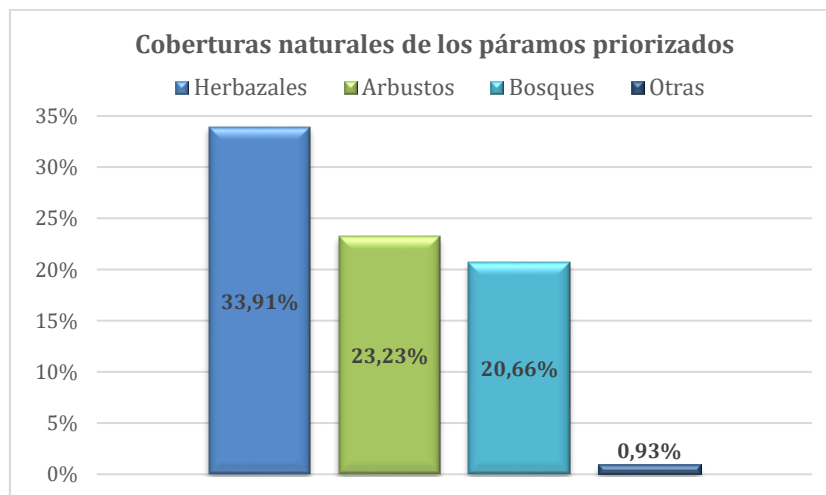


Figura 1-1: Coberturas naturales más representativas en los páramos del estudio del IDEAM et al., (2014).

Según los resultados mostrados en la **¡Error! No se encuentra el origen de la referencia.**, las coberturas naturales más representativas en los páramos son *Herbazales* con un 33,91%, *Arbustales* con un 23,23% y *Bosques* con un 20,66%. En *Otras coberturas* naturales se cuenta con *Turberas*, *Afloramientos rocosos*, y *Cuerpos de agua*, con una menor representatividad de 0,93%.

El aprendizaje profundo es una variante del Aprendizaje Automático, que se basa en redes neuronales artificiales interconectadas y agrupadas en capas para procesar datos y extraer información relevante, como se muestra en la **Figura 1-2**. Estos modelos de aprendizaje automático están diseñados para identificar y comprender patrones complejos, haciéndolos idóneos para tareas que requieren un análisis detallado de grandes conjuntos de datos. En la primera capa, la red aprende conceptos simples, que luego se envían a las capas ocultas para recopilar y combinar datos, generando información cada vez más compleja e infiriendo la categoría más probable. Este proceso se repite ajustando los pesos de cada conexión neuronal con el objetivo de obtener la mejor respuesta posible (Arcos Serrano, 2021).

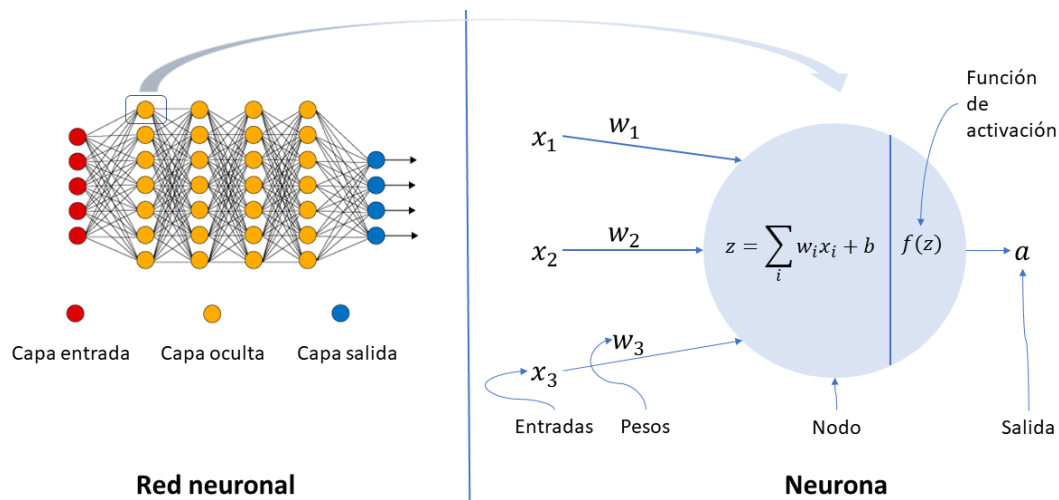


Figura 1-2: Una red neuronal artificial emula la red de neuronas de un cerebro. Esta incluye la capa de entrada, capas ocultas y la capa de salida que se conectan entre sí con diferentes pesos. Tomado de Davins Jovells, (2021)

La segmentación semántica de imágenes es un proceso en el campo del aprendizaje profundo que consiste en dividir una imagen en múltiples segmentos, asignando una etiqueta cualitativa a cada píxel (Osco et al., 2021). Este procedimiento se utiliza para identificar clases de objetos en una imagen y requiere de un conjunto de etiquetas para el entrenamiento. Cada píxel de la imagen debe tener una etiqueta que identifique una

clase con un color, como se puede ver en la **Figura 1-3**. Es crucial definir correctamente estas delimitaciones para lograr un entrenamiento efectivo y obtener un modelo que caracterice cada cobertura correctamente (Altamirano P, 2021).

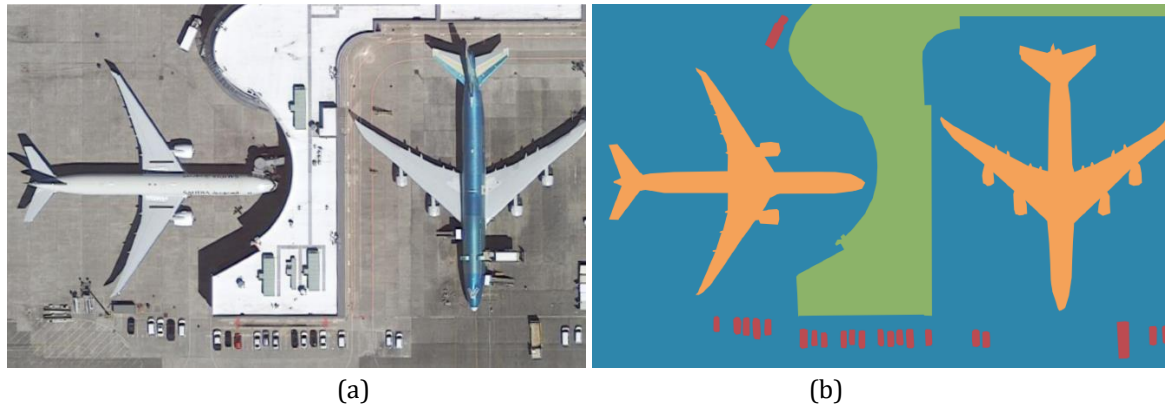


Figura 1-3: Segmentación semántica de la (a) Composición *RGB*, (b) imagen clasificada píxel a píxel asignándole una clase. Tomada de Hoeser & Kuenzer, (2020).

El Aprendizaje Automático se divide en tres áreas: supervisado, no supervisado y aprendizaje por refuerzo (Prince, 2023). El aprendizaje supervisado se enfoca en tareas de regresión y de clasificación. La clasificación es método útil en teledetección para categorizar coberturas terrestres en imágenes satelitales, como identificar agua, bosques o zonas urbanas en cada píxel, utilizando patrones y características espectrales aprendidas (Lecun et al., 2015; Prince, 2023). Dentro del aprendizaje supervisado, las *CNN* son reconocidas por su aplicación en teledetección, particularmente para la clasificación automática de imágenes, demostrando una notable eficiencia en el reconocimiento de objetos y escenas en imágenes satelitales (Deepan & Sudha, 2020).

El proceso de segmentación semántica usando *CNN*, empieza con una imagen *RGB* de entrada. En la fase de codificación, se extraen características esenciales de la imagen a través de capas de convolución y agrupación. Cada filtro de convolución se centra en identificar aspectos específicos de la imagen, como bordes, texturas o formas, generando un mapa de características que refleja estos detalles. Posteriormente, en la fase de decodificación, la imagen se reconstruye usando convolución traspuesta. Este paso es crucial para obtener la clasificación de cada píxel, resultando en una imagen clasificada del mismo tamaño que la imagen original (Martin-Abadal et al., 2018). En la **Figura 1-4**, se muestran las diversas capas que conforman una *CNN*, destacando la cantidad y el tamaño de las bandas que van almacenando rasgos característicos.

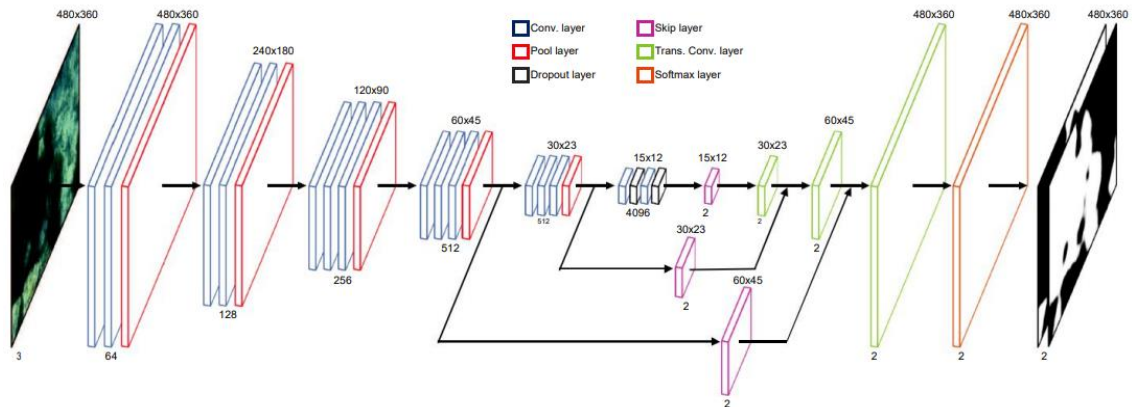


Figura 1-4: Una red de *CNN* se divide en dos etapas: codificación, que contiene la capa de convolucional, agrupamiento y exclusión; y decodificación, que comprende la capa de salto, convolucional transpuesta y la función de pérdida. Tomado de Martin-Abadal et al., (2018).

Existen diversas arquitecturas de redes profundas diseñadas específicamente para tareas de segmentación de imágenes, entre las que se encuentran: *Fully Convolutional Networks (FCN)*, *Pyramid Scene Parsing Network (PSP)*, *SegNet*, *U-Net*, *DeepLab* y *DenseNet* (Yuan et al., 2021). Para esta investigación, se han seleccionado dos arquitecturas destacadas: *U-Net* y *DeepLab* utilizando sus versiones más avanzadas, *U-Net++* y *DeepLabv3+*, respectivamente.

U-Net++ es reconocida por su eficiencia en la segmentación semántica, particularmente en la identificación de objetos pequeños y en la delineación de estructuras detalladas (Altamirano P, 2021, Zhou et al., 2018). Por otro lado, *DeepLabv3+* se destaca en la captura de características contextuales a diferentes escalas, ideal para la segmentación en imágenes de alta resolución debido a que combina las fortalezas de la convolución *atrous* y las estructuras de codificador-decodificador (Ayhan et al., 2020; L.-C. Chen et al., 2017). En conjunto, estas dos arquitecturas representan dos enfoques reconocidos en la segmentación semántica.

1.4 Pregunta de investigación

Esta investigación se enfocó en responder la pregunta: ¿Cuál de los modelos de segmentación semántica de imágenes satelitales, *U-Net++* o *DeepLabV3+*, demuestra el mejor rendimiento en la clasificación de cobertura de la tierra en páramos utilizando imágenes satelitales?

La expresión *rendimiento* involucra dos aspectos cruciales del proceso de clasificación de coberturas: (i) la exactitud temática de las categorías temáticas y; (ii) el desempeño computacional expresado en tiempo de entrenamiento y de inferencia. El caso de estudio seleccionado fue el complejo de páramos Tota-Bijagual-Mamapacha (TBM), un área de gran importancia por su suministro de agua ubicado en los departamentos de Boyacá y Cundinamarca, utilizando imágenes *Landsat 8* del periodo 2017-2019 y etiquetas de cobertura del mapa nacional de coberturas 1:100.000 del IDEAM de 2018.

Para lograr este objetivo, se plantearon las siguientes preguntas específicas:

- ¿Cómo abordar el desbalance de las diferentes clases de coberturas de páramos presentes en el área de estudio?
- ¿Cuál es el impacto de incrementar el número de bandas utilizadas como variables explicativas en la exactitud temática de los modelos?
- ¿Cuáles coberturas de páramos muestran una segmentación de mayor exactitud temática mediante el uso de los modelos *U-Net++* y *DeepLabv3+*?

1.5 Objetivos

Evaluar y comparar el rendimiento de los modelos de segmentación semántica de imágenes satelitales basados en las arquitecturas *U-Net++* y *DeepLabv3+* en la clasificación de coberturas en el complejo de páramos TBM.

1.5.1 Objetivos específicos

Los objetivos específicos que guían esta investigación son:

- Identificar y aplicar estrategias para abordar el desbalance de cubrimiento espacial de las clases de interés.
- Evaluar y cuantificar el impacto del aumento del número de bandas espectrales en la exactitud temática de los modelos *U-Net++* y *DeepLabv3+*.
- Comparar la exactitud temática de los modelos *U-Net++* y *DeepLabv3+* en la segmentación de coberturas de páramos.

1.6 Justificación

Los páramos en Colombia son fundamentales, suministrando agua al 70% de los municipios aledaños y sirviendo como la principal fuente de importantes cuerpos de

agua (Vásquez, A., Buitrago, 2011). Sin embargo, el conocimiento sobre la cobertura del suelo y su dinámica es limitado debido a baja frecuencia de actualización de los datos (Garavito, 2016). En el páramo TBM los estudios previos del IDEAM et al., (2014) y el IGAC, (2015), han documentado los efectos negativos de actividades antrópicas como la agricultura, la reforestación, la minería, la ganadería extensiva bovina y las quemadas. Sin embargo, falta una evaluación actualizada del impacto de estas actividades.

Esta investigación busca explorar métodos de clasificación alternativos que mejoren la comprensión de la dinámica de la cobertura del suelo en los páramos, esenciales para su conservación y el monitoreo ambiental frecuente. Tradicionalmente, esta tarea se ha abordado mediante métodos de interpretación visual de imágenes satelitales y aéreas (Suárez Londoño et al., 2017). Sin embargo, estos métodos son insuficientes en cuanto a la frecuencia de actualización, lo que limita el entendimiento y capacidad de respuesta ante cambios significativos (Foody, 2001). Por tanto, es fundamental adoptar técnicas de clasificación digital avanzadas, como los algoritmos de aprendizaje profundo, que ofrecen clasificaciones más rápidas y con alta exactitud temática (Yuan et al., 2021). Estas técnicas modernas han demostrado globalmente su eficiencia en la clasificación de coberturas y en el análisis de cambios ecológicos (Arcos Serrano, 2021; Kussul et al., 2017; Martin-Abadal et al., 2018).

El enfoque de esta investigación se centra en explorar el uso de *CNN* para mejorar la frecuencia en la clasificación de coberturas de páramos. Aprovechando los avances tecnológicos y la disponibilidad de imágenes satelitales más detalladas, el objetivo es desarrollar una metodología que no solo mejore la conservación de estos ecosistemas críticos, sino que también proporcione una comprensión más profunda de las ventajas y limitaciones de las *CNN* en la clasificación de coberturas de páramos. Se espera que los resultados del estudio beneficien a los ecosistemas de páramos y a las comunidades locales que dependen de estos recursos naturales, ofreciendo perspectivas valiosas para la aplicación futura de tecnologías de inteligencia artificial en la conservación ambiental.

1.7 Organización del documento

La presente investigación está estructurada de la siguiente manera para facilitar la comprensión del estudio. En el capítulo 2, se exploran conceptos fundamentales y estudios previos relacionados con la clasificación de coberturas en páramos utilizando *CNN*. Este capítulo establece el marco teórico y contextual necesario para la

investigación. En el capítulo 3, se detalla el área de estudio seleccionada, los datos utilizados y la metodología empleada para desarrollar y aplicar los modelos de *CNN* en la clasificación de coberturas. En el capítulo 4, se presentan y analizan los resultados obtenidos de los modelos, enfocándose en la exactitud temática y el tiempo de ejecución de las clasificaciones. En el capítulo 5, se discuten los resultados obtenidos, se responden las preguntas de investigación, y se examinan las contribuciones, ventajas y limitaciones del estudio. Finalmente, en el capítulo 6, se concluye con una síntesis de las lecciones aprendidas a lo largo de la investigación y se proponen recomendaciones para futuros estudios.

2. MARCO TEÓRICO Y ESTADO DEL ARTE

Este capítulo se organiza en dos secciones principales para proporcionar una comprensión integral de la clasificación de coberturas de páramo utilizando tecnologías avanzadas: la primera se centra en los conceptos claves para la clasificación de coberturas de páramo, con un énfasis particular en la aplicación de las *CNN* en segmentación semántica. Además, se analizan las métricas de exactitud utilizadas para evaluar la eficiencia de estas técnicas. En la segunda parte, se revisan estudios previos que han utilizado aprendizaje profundo en la clasificación de coberturas usando imágenes satelitales. Se detallan dos arquitecturas específicas de *CNN* y se abordaron los desafíos asociados con su uso.

2.1 Marco teórico

2.1.1 Ecosistema de páramos

Los páramos en Colombia son ecosistemas de alta montaña con una predominante vegetación tipo matorral o arbustiva, esenciales para el abastecimiento del agua y la conservación de la biodiversidad (Vásquez, A., Buitrago, 2011). Estos ecosistemas actúan como fuentes primarias de agua y desempeñan un papel decisivo en la retención de carbono atmosférico (Morales M, et al., 2007). Sin embargo, la fragilidad de los páramos frente a la actividad humana y el cambio climático plantea una amenaza significativa para su capacidad de mantener estos servicios ambientales.

Colombia alberga aproximadamente el 50% de los páramos del mundo, representando el 1,3% de la superficie continental del país, con una extensión de 14.434 km² (Garavito, 2016), distribuidos en 37 complejos a lo largo de las tres cordilleras, la Sierra Nevada de Santa Marta y el Nudo de los Pastos. Según Corpoboyacá, el departamento de Boyacá se destaca por contener la mayor extensión de páramos en Colombia, abarcando el 24% de los páramos del mundo, con una extensión de 684.837,086 hectáreas aproximadamente.

La biodiversidad en estos ecosistemas es notable, albergando aproximadamente 3.379 especies de plantas; 70 especies de mamíferos, 154 especies de aves y 90 especies de anfibios (Cabrera & Ramírez, 2014). Según el Mapa de Coberturas Vegetales, Ocupación y Uso del Espacio de Colombia elaborado por el IDEAM en 1996, Gómez Sánchez et al., (2002), los páramos se dividen en cuatro categorías principales:

- **Subpáramos:** Se encuentran a altitudes entre 2.400 y 3.200 msnm con una temperatura promedio de 10°C. Presentan áreas de transición con la vegetación andina, caracterizadas por arbustos y árboles bajos.
- **Páramo:** Se extienden entre 3.200 y 4200 msnm. Principalmente presentan coberturas de pajonales, pastizales y frailejonales. También se encuentran otras coberturas cerca de cursos de agua, y en ocasiones, mezcladas con bosques altoandinos y andinos, como chitales, puyales, turberas y matorrales, lo que dificulta su interpretación en imágenes satelitales. Las principales coberturas en los páramos incluyen pajonales (herbáceas con numerosas especies de gramíneas que generalmente alcanzan los 60 centímetros de altura, ubicadas por encima de los 3.000 msnm), pastizales (de origen mayoritariamente antrópico y establecidos para pastoreo) y frailejonales (caracterizados por la presencia de frailejones, excepcionalmente ubicados por debajo o cerca de los 2.600 msnm).
- **Superpáramo:** Se extiende aproximadamente de 4.200 hasta 4.700 msnm, con vegetación escasa debido a condiciones extremas de viento, precipitaciones y temperaturas por debajo de 0°C.
- **Zona nival:** Generalmente ubicada por encima de los 4.700 msnm, corresponde a la región más elevada del país, permanentemente cubierta de nieve y con temperaturas extremadamente bajas, inferiores a 0°C.

Esta investigación se centra en la clasificación de coberturas de páramos, incluyendo herbazales, arbustales y bosques, así como las áreas de subpáramo donde predominan las coberturas en transición. Este enfoque busca contribuir al monitoreo y preservación de estos ecosistemas, proporcionando datos valiosos para la gestión de los recursos hídricos y la protección de la biodiversidad. Al entender mejor la distribución y el estado de las coberturas de páramos, se facilita la toma de decisiones informadas para salvaguardar los servicios ambientales críticos que estos ecosistemas proporcionan.

2.1.2 Clasificación de cobertura de la tierra en páramos

En el estudio del 2014 se llevó a cabo una clasificación detallada de los páramos a escala 1:25,000 utilizando la metodología Corine Land Cover (CLC) adaptada para Colombia, abarcando 21 complejos de páramos (IDEAM et al., 2014). Para este estudio se empleó la interpretación visual de imágenes satelitales, principalmente de los satélites RapidEye y SPOT, que ofrecen una resolución espacial de 10 metros. Como resultado, se determinó que las coberturas más predominantes en el complejo TBM eran los *Herbazales* con un 42% del área, seguidos de *Áreas agrícolas heterogéneas* con un 23% y *Bosques* con un 13%, como se detalla en la **Figura 2-1**.

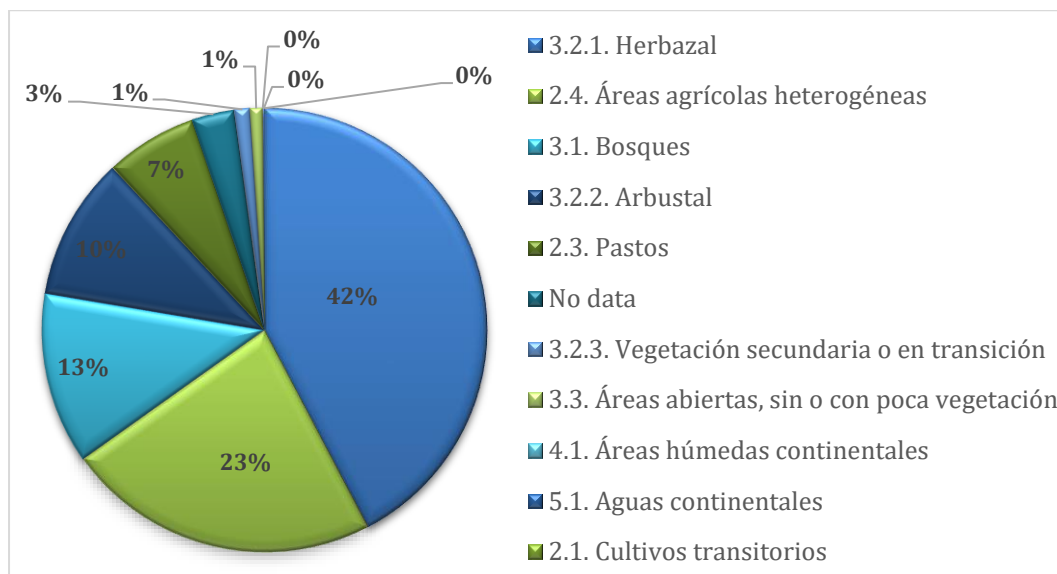


Figura 2-1: Las coberturas más representativas del páramo TBM son *Herbazales*, *Áreas agrícolas heterogéneas* y *Bosques*. Elaboración Propia basada en el estudio del (IDEAM et al., 2014).

Las coberturas del páramo TBM evidenciaron un alto grado de deterioro debido a actividades humanas como la agricultura, plantaciones forestales, minería, ganadería extensiva bovina y quemas. Esto ha contribuido a una pérdida significativa de biodiversidad en la región, según un estudio de suelos realizado por el Instituto Geográfico Agustín Codazzi (IGAC) en 2015.

2.1.3 Red neuronal convolucional (CNN)

Las *CNN* son modelos de aprendizaje profundo que simulan el proceso de visión humana para identificar y analizar patrones en imágenes (Saca & Moreno, 2019). Estas

redes son particularmente efectivas en la clasificación supervisada (Lecun et al., 2015), capaces de asignar etiquetas automáticamente a entradas basadas en patrones aprendidos durante el entrenamiento. Las *CNN* identifican características en las imágenes procesando capas sucesivas; las primeras capas detectan elementos simples como líneas y curvas, mientras que capas más profundas reconocen objetos y formas más complejas (Arcos Serrano, 2021).

En el campo de la teledetección, las *CNN* han demostrado ser una herramienta valiosa para la clasificación de coberturas y la interpretación de imágenes satelitales (Alzubaidi et al., 2021). Son útiles para analizar imágenes de distintos tipos de sensores, incluyendo *RGB*, multiespectral, hiperespectral y *LiDAR* (Arcos Serrano, 2021; Yuan et al., 2021). Lo que las hace ideales para la clasificación de coberturas en los páramos, permitiendo su delimitación para su seguimiento y protección efectiva.

El uso de *CNN* en la clasificación de coberturas enfrenta varios desafíos, como la necesidad de un gran volumen de datos etiquetados para el entrenamiento efectivo y de una buena capacidad computacional para procesar los datos (Yuan et al., 2021). Las *CNN* están compuestas por la capa convolucional que extrae características, la capa de agrupamiento que simplifican la información, y la capa de clasificación que generan las inferencias finales, ver en detalle en la **Figura 2-2**:

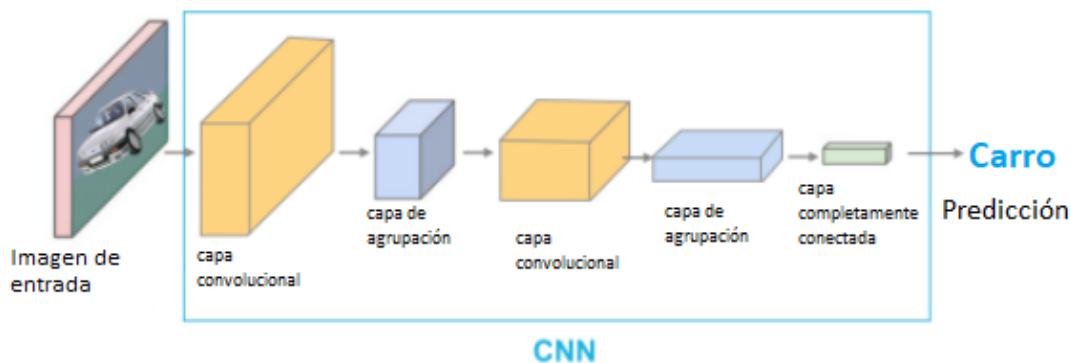


Figura 2-2: La arquitectura *CNN* está compuesta por la capa convolucional, capa de agrupamiento y la capa de clasificación. Adaptado de Altamirano (2021)

- **Capa Convolucional:** Esta capa está formada por un conjunto de filtros convolucionales, que se aplican a la imagen de entrada de múltiples dimensiones para extraer características esenciales como bordes, texturas y formas. La aplicación de estos filtros produce un mapa de características que representa

visualmente estos detalles, esencial para el análisis posterior de la imagen (Alzubaidi et al., 2021), ver en la **Figura 2-3**:

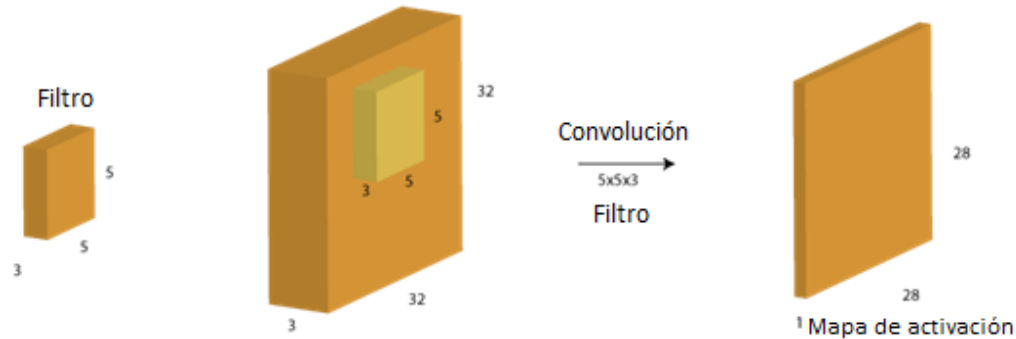


Figura 2-3: La capa convolucional procesa una imagen de entrada para extraer características importantes. Adaptado de Iaar (2020)

- ✓ **Kernel:** Una pequeña matriz de valores que se desplaza sobre la imagen de entrada, realizando operaciones de convolución para detectar patrones específicos. Los pesos de estos *kernels* se optimizan durante el entrenamiento para capturar con exactitud características como bordes, esquinas y texturas.
- ✓ **Padding:** Consiste en añadir píxeles alrededor de la imagen de entrada antes de aplicar la convolución. Este proceso ayuda a preservar la dimensión de la imagen, asegurando que las características en los bordes y esquinas sean adecuadamente evaluadas y no se pierdan durante la convolución.
- ✓ **Stride:** Se refiere a la cantidad de píxeles por los que el *kernel* se desplaza en cada movimiento a lo largo de la imagen. Un *stride* mayor resulta en un mapa de características de menor tamaño, ya que el *kernel* salta sobre más píxeles de la imagen en cada paso.
- ✓ **Conectividad dispersa:** Esta característica de las *CNN* implica que cada neurona en una capa está conectada solo con un subconjunto de neuronas en la capa siguiente, lo que reduce significativamente la cantidad de conexiones neuronales y, por ende, la demanda de memoria, aumentando la eficiencia del modelo.
- **Capa de agrupación (Pooling):** Esta capa reduce las dimensiones de los mapas de características generados por las operaciones de convolución. Este proceso ayuda a preservar las características esenciales mientras reduce la cantidad de datos, lo que facilita y acelera el procesamiento posterior, ver en la **Figura 2-4**. Los métodos más utilizados son la agrupación máxima, mínima y *GAP*. La

agrupación máxima selecciona el valor máximo de un área predeterminada, mientras que la agrupación mínima selecciona el valor mínimo. Por otro lado, *GAP* calcula el promedio de todas las características en un mapa, proporcionando una visión general y compacta de las características destacadas (Alzubaidi et al., 2021).



Figura 2-4: La capa de agrupación, simplifica la información reduciendo la resolución espacial de las características extraídas. Tomada de Iaar, (2020).

- **Función de activación (no linealidad):** la función de activación es fundamental para mapear entradas a salidas. Esta introduce no linealidades en el modelo, lo cual es crucial para aprender y modelar relaciones complejas en los datos (Alzubaidi et al., 2021). Entre las funciones de activación comúnmente utilizadas en *CNN* se encuentran:
 - ✓ **Sigmoid:** Esta función restringe la salida a un rango entre cero y uno, aceptando números reales como entrada, útil para modelos donde se requieren probabilidades como salida. La curva de la función sigmoide tiene una forma de S. Ver la **Figura 2-5**:

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

Figura 2-5: Función de activación *Sigmoid*. Adaptado de Torralba, (2022).

- ✓ **Tanh:** Similar a la función sigmoide, *Tanh* acepta números reales como entrada, pero su salida está restringida a un rango entre -1 y 1, proporcionando una mejor normalización para algunas aplicaciones. Ver la **Figura 2-6**:

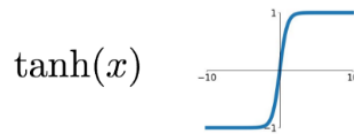


Figura 2-6: Función de activación *Tanh*. Adaptado de Torralba, (2022).

- ✓ **ReLU (Rectified Linear Unit):** Es la función de activación más utilizada en *CNN*. Convierte todos los valores de entrada negativos a cero, manteniendo los valores positivos sin cambios. Su principal ventaja sobre otras funciones es su baja carga computacional. Ver la **Figura 2-7:**

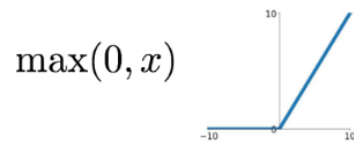


Figura 2-7: Función de activación *ReLU*. Adaptado de Torralba, (2022).

- **Clasificación:** La última etapa de una *CNN* es la capa de clasificación, una capa totalmente conectada que consolida las características identificadas por las capas convolucionales y de agrupación. El tamaño de las neuronas en la salida depende de la cantidad de clases a clasificar (Altamirano P, 2021). Esta capa utiliza la información recopilada para clasificar la entrada en categorías definidas, basándose en el aprendizaje realizado durante el entrenamiento (Alzubaidi et al., 2021).
- **Funciones de pérdida:** Son esenciales durante la etapa de aprendizaje de una *CNN*. Estas funciones evalúan la exactitud del modelo comparando las inferencias de la red con las etiquetas reales. El objetivo de la optimización en el entrenamiento de la *CNN* es minimizar esta pérdida, ajustando los pesos de la red para mejorar su rendimiento en tareas como la clasificación (Alzubaidi et al., 2021). Una de las más destaca es la Función de Pérdida de Entropía Cruzada (también conocida como *Softmax*)
 - ✓ **Softmax:** es ampliamente utilizada para evaluar el rendimiento de los modelos de *CNN*, especialmente en problemas de clasificación multiclase. La función *Softmax* mide la diferencia entre dos distribuciones de probabilidad: la inferencia de la red y la distribución real de las etiquetas. Su salida es una probabilidad $p \in \{0, 1\}$, indicando la probabilidad asignada a cada clase. (Alzubaidi et al., 2021).

$$p_i = \frac{e^{u_i}}{\sum_{k=1}^N e_k^a} \quad (2.1)$$

- **Regularización:** Para evitar el sobreajuste en modelos *CNN*, se utilizan técnicas de regularización, con el fin de tener un modelo balanceado, como se muestra en la **Figura 2-8**. Estas técnicas ajustan el proceso de aprendizaje para asegurar que el modelo no solo se ajuste a los datos de entrenamiento, sino que también generalice bien los nuevos datos (Alzubaidi et al., 2021).
 - ✓ **Regularización L1/L2:** Penaliza los pesos de la red durante el entrenamiento para mantenerlos pequeños.
 - ✓ **Dropout:** Apaga aleatoriamente un número de neuronas durante el entrenamiento para evitar la dependencia excesiva en cualquier neurona particular.

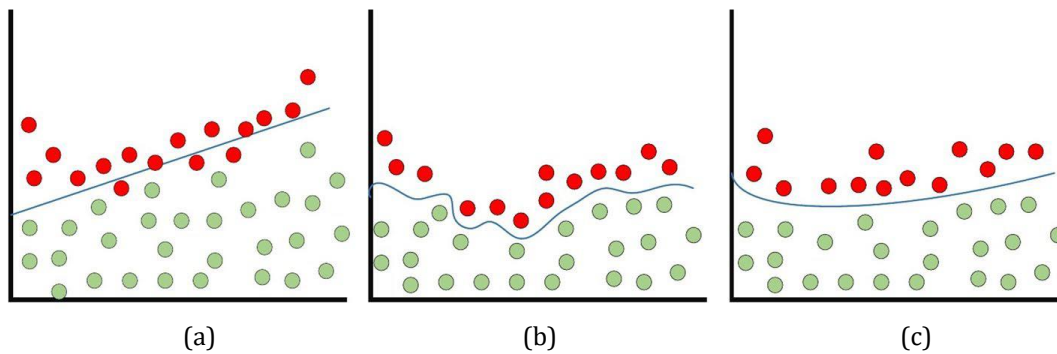


Figura 2-8: Ajuste del resultado del modelo (a) Subajustado, (b) Sobreajuste y (c) Balanceado. Tomado de Alzubaidi et al., (2021).

Al aplicar modelos de *CNN* para la clasificación de coberturas en imágenes multiespectrales, como en el caso de imágenes satelitales, se enfrentan varios desafíos importantes (Yuan et al., 2021):

- **Variabilidad espectral y espacial:** Las coberturas en las imágenes multiespectrales presentan variaciones significativas tanto en su composición espectral como en sus patrones espaciales. Esto significa que diferentes áreas con el mismo tipo de cobertura pueden aparecer distintas en las imágenes debido a condiciones de iluminación, estaciones y otros factores ambientales, lo que complica el proceso de clasificación.

- **Desbalance de clases:** Las *CNN* requieren grandes conjuntos de datos etiquetados de todas la etiquetas para entrenar eficientemente, lo que implica disponer de una cantidad sustancial de imágenes satelitales correctamente etiquetadas. La preparación de estos datos puede ser un proceso extenso y costoso.
- **Desarrollo de modelos personalizados para diferentes bandas espectrales:** Las imágenes multiespectrales contienen múltiples bandas, cada una puede contener información vital sobre diferentes características de la cobertura. Por lo que se requiere adaptar las *CNN* para que se procesen eficientemente múltiples bandas para mejorar la exactitud en la clasificación.
- **Necesidad de plataformas computacionales de alto rendimiento:** El procesamiento y análisis de grandes volúmenes de datos multiespectrales requieren plataformas computacionales potentes, especialmente cuando se utilizan modelos de aprendizaje profundo como las *CNN*. La necesidad de hardware especializado, como *GPUs* o *clusters* de alto rendimiento, puede generar altos costos.

2.1.4 Evaluación de exactitud

La exactitud temática es crucial para medir qué tan exactas son las inferencias de las cobertura realizadas por un modelo comparadas con la realidad observada (Foody, 2001). Se utiliza comúnmente una matriz de error o confusión para describir el rendimiento de modelos de clasificación supervisada (Sánchez,2016), ver la **Tabla 2-1**. La matriz de confusión es fundamental para comprender y evaluar de manera integral la exactitud y eficiencia de un modelo de clasificación (Radoux & Bogaert, 2017).

Tabla 2-1. Matriz de error con los tipos de clasificaciones.

		+	-
		+	-
Clasificación inferida	+	TP	FP
	-	FN	TN
		Clasificación real	

Los errores de omisión (*FN*) son casos en los que una clase verdadera no se identifica correctamente, mientras que los errores de comisión (*FP*) son casos donde se clasifica erróneamente una unidad en una clase a la que no pertenece:

- ✓ Verdaderos Positivos (*TP*): Clasificaciones correctas.
 - ✓ Verdaderos Negativos (*TN*): Clasificaciones correctas de no pertenencia.
 - ✓ Falsos Positivos (*FP*): Clasificaciones incorrectas asignadas a una clase.
 - ✓ Falsos Negativos (*FN*): Clasificaciones no asignadas a su clase verdadera.
- **Métricas de Evaluación:** Existen diversas métricas de evaluación para medir el desempeño de los modelos (Ej: Belyadi & Haghighat, 2021):

- ✓ **Overall Accuracy (OA):** Evalúa el porcentaje de píxeles correctamente clasificados en comparación con el total de píxeles. Se determina dividiendo la cantidad de píxeles correctamente clasificados por el número total de píxeles.

$$OA = \frac{TN+TP}{TN+FN+TP+FP} \quad (2.2)$$

- ✓ **Kappa:** Evalúa la concordancia entre las clasificaciones del modelo y los datos de referencia, considerando también la concordancia que podría ocurrir por azar. Para su interpretación Monserud & Leemans, (1992) propusieron una escala donde los valores menores a 40% se clasifican como pobres, 40-55% suficientes, 55-70% buenos, 70-85% muy buenos y mayores a 85% excelentes.

$$Kappa = \frac{N \sum_{ij=1}^m D_{ij} - \sum_{ij=1}^m R_i * C_j}{N^2 - \sum_{ij=1}^m R_i * C_j} \quad (2.3)$$

- ✓ **Intersection over Union (IoU):** Mide la superposición entre la inferencia del modelo y la referencia. Se calcula como la intersección de las áreas predichas y reales dividida por su unión.

$$IoU = \frac{TP}{TP+FP+FN} \quad (2.4)$$

- ✓ **F1 - score (F1):** Combina las métricas de *precision* y *recall* en una métrica única para resumir la capacidad de clasificación del modelo para cada clase.

$$precision = \frac{TP}{TP+FP} \quad (2.5)$$

$$recall = \frac{TP}{TP+FN} \quad (2.6)$$

$$F1 = 2 * \frac{precision * recall}{precision + recall} \quad (2.7)$$

Estas métricas de evaluación proporcionan en detalle el rendimiento del modelo, destacando no solo la exactitud global sino también la eficiencia en la clasificación por clases, lo que es crucial para aplicaciones en teledetección.

2.2 Estado del arte

La segmentación semántica de imágenes satelitales usando *CNN* ha experimentado avances significativos recientemente, particularmente en teledetección y clasificación de coberturas de tierra. (Arcos Serrano, 2021; Ayhan et al., 2020; Kemker et al., 2018, Kussul et al., 2017). Estas tecnologías han resuelto eficientemente problemas complejos de procesamiento de imágenes, destacándose en diversos estudios (Audebert et al., 2017; Ayhan et al., 2020a; Coca Castro et al., 2021a; Huang et al., 2018; Kemker et al., 2018).

Entre las arquitecturas de *CNN* se encuentran *FCN*, *PSP* y *SegNet*, destacándose *U-Net* y *DeepLab* por su eficiencia (L. C. Chen et al., 2018; Fan et al., 2022). Especialmente, las versiones más avanzadas *U-Net++* y *DeeplabV3+* (Ayhan et al., 2020; Zhou et al., 2018, 2019). Estas arquitecturas se han adaptado y optimizado para enfrentar desafíos específicos en teledetección y clasificación de coberturas, demostrando una notable capacidad para manejar la complejidad y variabilidad de las imágenes satelitales.

A pesar de estos avances, la investigación en el campo de la teledetección enfrenta grandes desafíos. Yuan et al., (2021) destaca aspectos críticos como la escasez de grandes conjuntos de datos de teledetección con imágenes correctamente etiquetadas, que son esenciales para el entrenamiento de los modelos de aprendizaje profundo. Además, los altos requerimientos computacionales y la adaptación de las arquitecturas a imágenes multiespectrales. También persiste el desafío del desbalance de clases en la clasificación de ecosistemas diversificados (Z. Chen et al., 2022; Iyer et al., 2021).

La aplicación de *CNN* en la clasificación de coberturas de páramos es prometedora. Un ejemplo es el estudio de Castelo-Cabay et al., (2022) en un páramo ecuatoriano, en el cual se utilizaron diferentes métodos de clasificación como el análisis de imágenes basado en píxeles (*PBIA*), el análisis de imagen basado en objetos geográficos (*GEOBIA*) y *CNN*. Además, se enriqueció el estudio incluyendo el Índice de Vegetación de

Diferencia Normalizada (*NDVI*) y el índice de suelo desnudo (*BSI*) como variables explicativas. Los resultados del modelo *CNN* fueron prometedores, con una Exactitud Global (*OA*) de 87.43%. Además, el desempeño en las coberturas de *Pastos, Rastrojos, Cultivos y Herbazales* fueron de un *F1* de 91%, 88%, 84% y 82% respectivamente. Estos resultados subrayan la capacidad de las *CNN* para ofrecer clasificaciones con buena exactitud y eficientes en la evaluación de coberturas de ecosistemas críticos como los páramos

En Colombia, el campo del Aprendizaje Profundo todavía se encuentra en una fase inicial. A pesar de esto, existen estudios como el de Suárez Londoño et al., (2017), que utilizaron estas técnicas para clasificar coberturas de *Bosques, Áreas con vegetación herbácea y/o Arbustiva, Áreas abiertas con poca o ninguna vegetación, y Áreas húmedas continentales*, utilizando imágenes *Landsat 5 TM* del Parque Nacional Tuparro, ubicado en la región de la Orinoquía. Obteniendo buenos resultados con una *OA* de 91.02% en la validación. Estos resultados resaltan la eficiencia de las *CNN* en tareas de clasificación y subrayan la importancia de avanzar en la investigación para desarrollar modelos adaptados a las características únicas de los ecosistemas colombianos, como los páramos.

2.2.1 U-Net

U-Net, desarrollada por Navab et al., (2015) es una arquitectura de *CNN* destacada para la segmentación semántica, inicialmente enfocada en imágenes médicas, ganó el desafío de segmentación de estructuras neuronales en el Simposio Internacional IEEE sobre Imágenes Biomédicas (*ISBI*) en 2015. *U-Net* ha demostrado ser eficiente en otros campos como la teledetección y la clasificación de coberturas. Su éxito en estas áreas se debe a su capacidad para capturar con alta exactitud tanto las características globales como las locales de las imágenes. *U-Net++* es una versión avanzada de *U-Net* que incorpora conexiones más complejas entre la codificación y la decodificación de la red, que ayudan a reducir las diferencias en la interpretación de las imágenes, mejorando la claridad de los detalles finos al unir la información. Esta característica lo hace especialmente útil para definir con exactitud los elementos dentro de una imagen (Zhou et al., 2019).

Su estructura en forma de U consta de dos fases: codificación y decodificación, simétricas entre sí, como se muestran en la **Figura 2-9**. En la fase de codificación, se utiliza una arquitectura convolucional que captura el contexto de la imagen a través de varios bloques. Cada bloque consiste en una capa convolucional de 3x3, activación *ReLU* y *Max Pooling* de 2x2 con un *stride* de 2 (Navab et al., 2015). Este diseño permite

duplicar los mapas de características en cada paso, capturando así detalles globales de la imagen. Durante la fase de decodificación, la imagen se expande mediante convolución traspuesta para clasificar cada píxel. Esta fase también incluye la concatenación de las características recortadas de la fase de codificación con una capa de convolución traspuesta, seguida de *Max Pooling* y convoluciones adicionales de 3x3 con activación *ReLU*. Finalmente, una convolución de 1x1 asigna a cada vector la clase correspondiente (Navab et al., 2015).

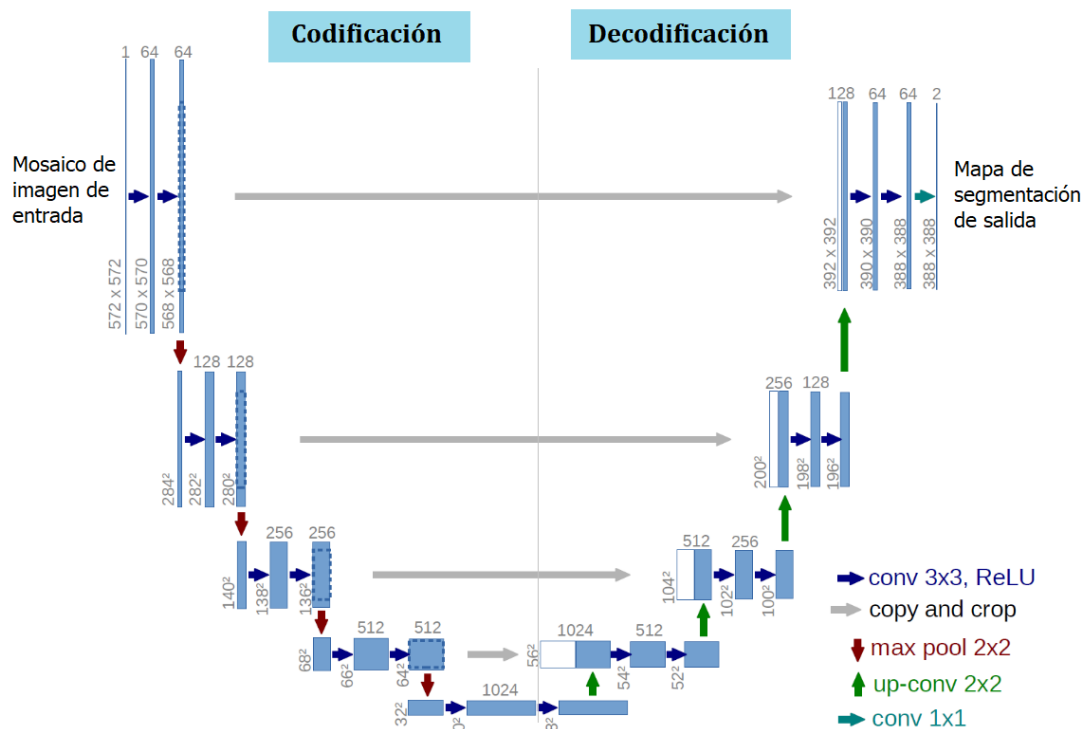


Figura 2-9: En la arquitectura *U-Net*, los cuadros azules simbolizan los mapas de características, con el número de canales en la parte superior y las dimensiones x-y en la parte inferior izquierda. Los cuadros blancos representan copias de los mapas de características. Las operaciones se identifican mediante flechas que los conectan. Adaptado de Navab et al., (2015).

La técnica de *Max-unpooling* utilizada en *U-Net* expande los mapas de características, permitiendo una segmentación más detallada, como se muestra en la **Figura 2-10**. Esta técnica duplica las dimensiones del mapa, seleccionando el valor máximo para incluir en la nueva matriz (Altamirano P, 2021).

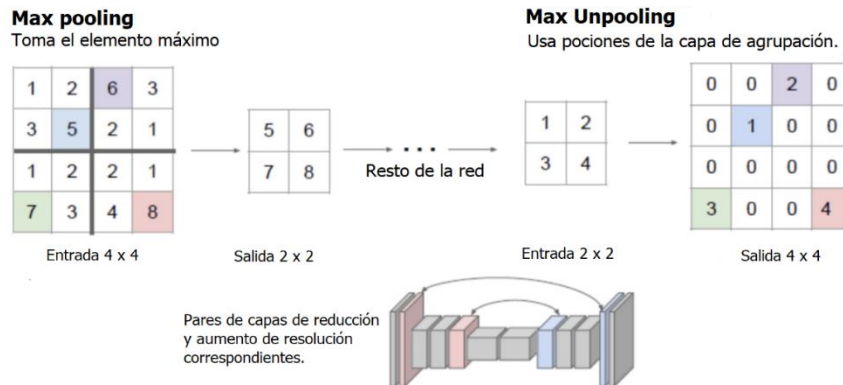


Figura 2-10: Comparación entre *Max-pooling* y *Max-unpooling*. Tomada de Altamirano P, (2021).

2.2.2 DeeplabV3+

DeepLabV3+, desarrollada por *Google* (Sahu, 2019), representa un avance significativo en la segmentación semántica mediante el uso de la Convolución *Atrous* y la Agrupación de Pirámides Espaciales *Atrous*. La Convolución *Atrous*, con su parámetro *Rate* ajustable como se muestra en la **Figura 2-11**, el cual expande el campo receptivo permitiendo a la red capturar información detallada sin comprometer la resolución de la imagen. Esta técnica es fundamental para manejar la complejidad y variabilidad en las imágenes. La Agrupación de Pirámides Espaciales *Atrous*, por otro lado, analiza las características convolucionales en múltiples escalas. Esto mejora el rendimiento de *DeepLabV3+* al codificar el contexto en diferentes niveles, lo que es crucial para la segmentación exacta en diversas aplicaciones, incluyendo la clasificación de coberturas de tierra (Chen et al., 2018).

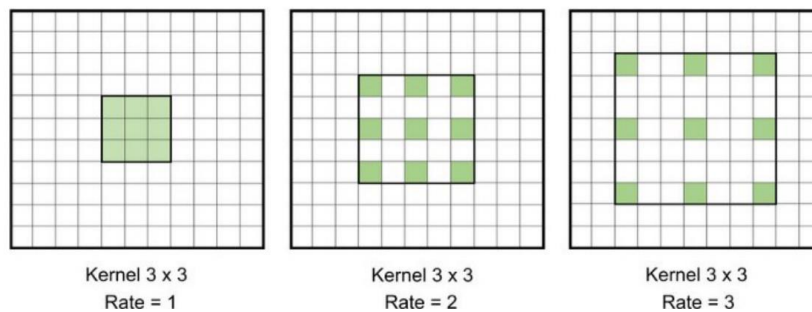


Figura 2-11 *Kernel 3 x 3* con *rate = 1*, *rate = 2* y un *rate = 3*. Tomada de Chen et al., (2018).

Desde su primera versión hasta *DeepLabV3*, la arquitectura ha experimentado evoluciones significativas. La versión inicial mejoró el modelo de red totalmente convolucional, mientras que *DeepLabV2* introdujo la Agrupación de Pirámides

Espaciales Atrous, considerando diferentes escalas de objetos para mejorar su capacidad de entender las relaciones de un objeto con su medio a distintas escalas (L. C. Chen et al., 2018). *DeepLabV3* modificó la manera de capturar información contextual de múltiples escalas a través de características analizadas con filtros y operaciones de agrupación, incluyendo la Convolución *Atrous* (Sahu, 2019).

La arquitectura *DeepLabV3+* consta de un codificador y un decodificador, como se aprecia en la **Figura 2-12**. El codificador reduce los mapas de características y captura información semántica, mientras que el decodificador recupera la información espacial y genera segmentaciones más nítidas (Sahu, 2019). En *DeepLabV3+*, se añadió un decodificador simple y eficiente para refinar aún más la clasificación. Este proceso comienza con la reducción del tamaño de la imagen mediante convoluciones *Atrous* y la generación de mapas de características a diferentes escalas. Tras la aplicación de una convolución 1×1 y la Agrupación de Pirámides Espaciales *Atrous*, se realiza una concatenación y se aplican convoluciones 3×3 (Chen et al., 2018). Finalmente, se efectúa un muestreo hacia arriba por un factor de 4, resultando en una salida del mismo tamaño que la imagen de entrada (Sahu, 2019).

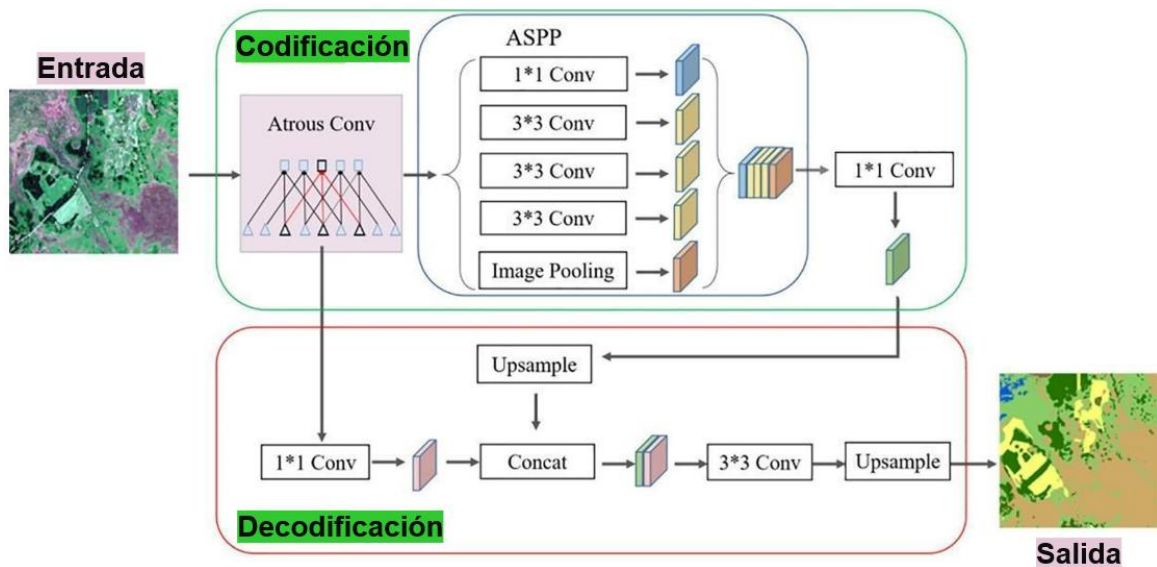


Figura 2-12: Arquitectura *DeepLabV3+*, mostrando las fases del codificador y decodificador con sus componentes claves, incluyendo la Convolución *Atrous* y la Agrupación de Pirámides Espaciales *Atrous*. Adaptado de (Liu et al., 2021).

2.2.3 Retos en la aplicación de clasificación de coberturas en el análisis de imágenes de teledetección

Las *CNN* han demostrado ser herramientas poderosas en el procesamiento de imágenes de teledetección, especialmente en la clasificación de coberturas (Zhu et al., 2017, Ayhan et al., 2020b, Yuan et al., 2021). Sin embargo, enfrentan varios desafíos (Yuan et al., 2021) :

- **Requisitos de datos etiquetados:** La eficiencia de las *CNN* depende en gran medida de la disponibilidad de datos etiquetados de alta calidad. El proceso de generación de etiquetas con buena exactitud y que sean representativas para cada clase en la imagen es intensivo en cuanto a recursos y tiempo. La calidad de las etiquetas impacta directamente en la exactitud de los modelos entrenados.
- **Alto requerimiento computacional:** Debido a su complejidad y al gran volumen de datos implicados, los modelos *CNN* requieren una considerable capacidad computacional. El uso de unidades de procesamiento gráfico (*GPU*) ayuda a acelerar este proceso, aunque su costo y disponibilidad pueden ser una limitante.
- **Adaptación a imágenes multispectrales:** se requiere modificaciones en la arquitectura y el entrenamiento de las *CNN* para que trabajen eficientemente con imágenes multispectrales comunes en teledetección. Estas imágenes presentan desafíos adicionales debido a un mayor número de bandas espectrales en comparación con las imágenes *RGB* estándar.
- **Desbalance de clases:** La distribución desigual de las coberturas en las imágenes puede resultar en un sesgo hacia las clases más representadas, afectando la capacidad del modelo para clasificar con exactitud las clases con menos frecuencia. Para abordar este desafío, se emplean técnicas como la ponderación de clases en la función de pérdida o la generación sintética de datos para las clases menos representadas (Z. Chen et al., 2022; Iyer et al., 2021).

A pesar de estos desafíos, el potencial de estas técnicas para proporcionar análisis detallados y con una buena exactitud temática de las coberturas siguen siendo significativos, ofreciendo valiosas oportunidades para el avance en el monitoreo y la gestión de ecosistemas a través de la teledetección.

3. MATERIALES Y MÉTODOS

Este capítulo está organizado en cuatro secciones principales. La primera sección describe detalladamente el área de estudio, proporcionando un contexto esencial para entender el enfoque del estudio. La segunda sección detalla los datos utilizados, incluyendo sus características y el proceso de selección o limpieza aplicada. La tercera expone la metodología empleada, describiendo los procesos aplicados, desde la preparación de los datos hasta los métodos de análisis y modelado utilizados. Finalmente, la cuarta, expone las herramientas y software utilizados en las diferentes fases de la investigación.

3.1 Área de estudio

El complejo de páramos TBM, ubicado en la cordillera Oriental de Colombia, es el cuarto más grande del país y comprende aproximadamente el 60% de los páramos de Colombia. Se extiende por 151.247 hectáreas según la resolución 1771 de 2016 (Gil, 2017) y atraviesa 19 municipios de Boyacá y Cundinamarca. Este complejo es un área crucial para el suministro de agua, influenciando varias subcuencas hidrográficas y proporcionando servicios ecosistémicos esenciales (Cabrera & Ramírez, 2014).

El paisaje del complejo varía entre 2.900 a 4.000 msnm de altitud, caracterizado por un clima frío y húmedo, con temperaturas que oscilan entre 6°C a 12°C y precipitaciones anuales que oscilan entre 1.000 a 2.500 mm. Las lluvias son más frecuentes de marzo a junio y de octubre a diciembre, mientras que la temporada seca se extiende de julio a septiembre y enero a febrero. Estas condiciones climáticas favorecen la diversidad de coberturas vegetales y la presencia de especies endémicas. Sin embargo, aproximadamente el 57,3% el paisaje ha sido modificado por actividades humanas como la agricultura y la ganadería, impactando tanto la biodiversidad como los servicios ecosistémicos (Gil, 2017). Las principales coberturas en el páramo TBM se muestran en la **Figura 3-1**.



Figura 3-1: Principales Coberturas presentes en el páramo TBM: (a) *Rastrojos y Pastos*; (b) *Herbazales*. Tomado de lapensadera.com (2021).

Dentro del complejo de páramos TBM, existen 3 áreas protegidas que abarcan 46,256 ha, lo que representa el 31% del total del complejo. Estas áreas incluyen el Parque Regional Natural Siscunsi-Ocetá, la Reserva Forestal Protectora Nacional de la Cuchilla de Sucuncuca y el Distrito de Manejo Integrado Lago de Sochagota, cruciales para la conservación de la biodiversidad y la protección de los recursos hídricos (Bocarejo et al., 2014). No obstante, gran parte del Páramo TBM se encuentra fuera de las áreas protegidas, lo que subraya la urgencia de implementar estrategias de monitoreo basadas en datos actualizados para su protección.

El Páramo TBM se distingue de otros páramos importantes de Colombia como Sumapaz, Chingaza, Santurbán y Cocuy, por su singular combinación de ecosistemas de alta montaña y significativos cuerpos de agua. Mientras que los páramos de Sumapaz y Chingaza son esenciales para el suministro de agua y la conservación de la biodiversidad en Bogotá; el Cocuy es famoso por sus majestuosos glaciares y como destino turismo de alta montaña (Vásquez, A., Buitrago, 2011).

En cambio, el TBM se destaca por su enfoque en la conservación integrada de ecosistemas acuáticos y terrestres, lo que demuestra la rica diversidad y riqueza de los paisajes naturales de Colombia (Bocarejo et al., 2014; IDEAM et al., 2014; Vásquez, A., Buitrago, 2011). En la **Tabla 3-1** se resumen las características más importantes de dichos páramos:

Tabla 3-1. Características principales de los páramos más importantes de Colombia.

Páramo	Ubicación	Tamaño ha (Aprox.)	Tipo de Páramo	Ecosistema	Importancia Ecológica
Sumapaz	Cundinamarca	333.420	Subpáramo, Páramo	Alta montaña, humedales. Con una biodiversidad de: Frailejones, cóndores.	Fuente principal de agua para Bogotá, conservación de biodiversidad.
Chingaza	Cundinamarca	109.956	Subpáramo, Páramo	Bosques, humedales. Con una biodiversidad de: Venados, zorros	Suministro de agua para Bogotá, biodiversidad
Santurbán	Santander	142.608	Páramo	Alta montaña, humedales. Con una biodiversidad de: Pájaro carpintero, frailejones	Protección de cuencas hidrográficas, biodiversidad
Cocuy	Boyacá	271.033	Páramo, Superpáramo	Alta montaña, glaciares. Con una biodiversidad de: Lobos de páramo, frailejones	Turismo de alta montaña, fuentes hídricas
TBM	Boyacá	151.247	Páramo	Alta montaña, lago. Con una biodiversidad de: Especies de aves, flora endémica	Regulación hídrica, biodiversidad

El área de estudio abarca 1.914.973,76 hectáreas, centrándose en el páramo TBM y sus alrededores, como se muestra en la **Figura 3-2**. La elección de este páramo no solo es por su gran extensión y rica biodiversidad, sino también por su importancia en el suministro de agua y en la regulación del clima regional. Un análisis detallado de las diversas coberturas del páramo TBM contribuirá en la toma de decisiones relacionadas con su conservación y el uso sostenible de sus recursos naturales.

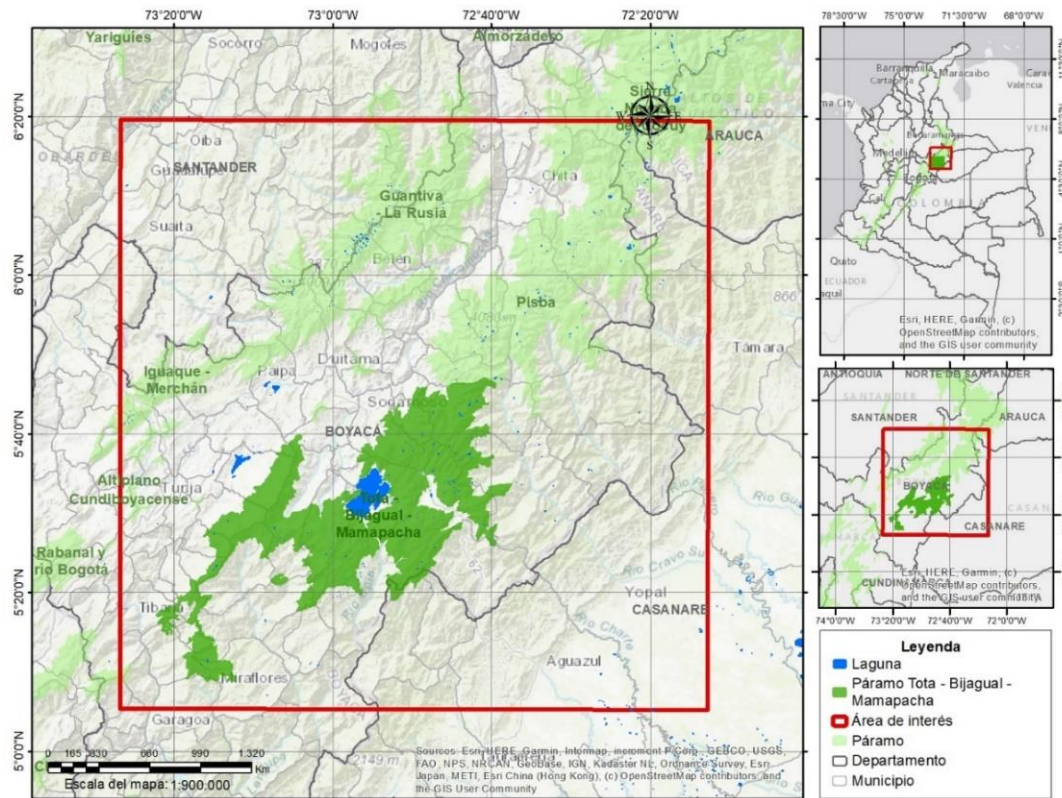


Figura 3-2: Área de estudio se muestra en color rojo. Fuente de datos MADS e IGAC, composición propia.

3.2 Datos

En esta investigación, se utilizaron imágenes satelitales *Landsat 8*, modelo digital de elevación y el Mapa nacional de coberturas 1:100.000 del IDEAM de 2018.

3.2.1 Imágenes satelitales

Se utilizaron imágenes del satélite *Landsat 8*, abarcando el periodo comprendido entre 01 de enero de 2017 al 31 de julio de 2019. Estas imágenes, con una resolución espacial de 30 metros, cuentan con reflectancia superficial corregida atmosféricamente proporcionada por los sensores *OLI/TIRS*. En la **Tabla 3-2** se proporcionan los detalles específicos sobre las bandas del satélite *Landsat 8*. La composición en color natural del área de estudio se presenta en la **Figura 3-3**.

- **Resolución Espacial:** 30 metros
- **Resolución Temporal:** 16 días
- **Cubrimiento:** 170 km x 183 km
- **Resolución Espectral:** 8 bandas que consisten en 5 bandas visibles e infrarrojas cercanas (*VNIR*), 2 bandas infrarrojas de onda corta (*SWIR*) y dos bandas infrarrojas térmicas (*TIR*).

Tabla 3-2. Información de las Bandas *OLI* de *Landsat 8*. Fuente *GEE*, (2007).

Nombre	Tamaño píxel(m)	Longitud de onda (μm)	Descripción
B1	30	0,43 - 0,45	Costera - Aerosoles
B2	30	0,45 - 0,51	Azul
B3	30	0,53 - 0,59	Verde
B4	30	0,64 - 0,67	Rojo
B5	30	0,85 - 0,88	Infrarrojo cercano (<i>NIR</i>)
B6	30	1,57 - 1,65	Infrarrojo de onda corta 1 (<i>SWIR 1</i>)
B7	30	2,11 - 2,29	Infrarrojo de onda corta 2 (<i>SWIR 2</i>)
B10	30	10,60 - 11,19	Infrarrojo térmico 1
B11	30	11,50 - 12,51	Infrarrojo térmico 2

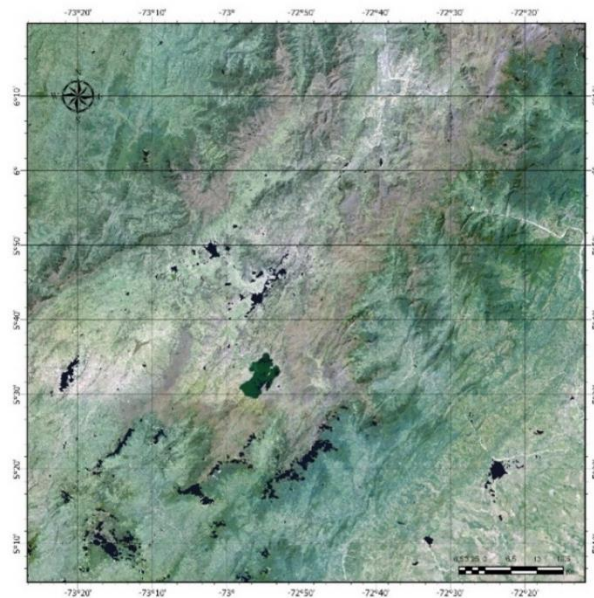


Figura 3-3: Composición *RGB* en color natural del área de estudio capturada por el satélite *Landsat 8*, utilizando la combinación de bandas 4, 3 y 2. Esta representación visual permite una observación de la superficie terrestre similar a la visión humana. Fuente de datos *USGS*, composición propia.

El índice de vegetación de diferencia normalizada (*NDVI*) de *Landsat 8* se empleó para evaluar la salud y distribución de la vegetación en el área de estudio. El *NDVI* utiliza la reflectancia medida en las bandas del *NIR* y *Roja* para calcular el estado de la vegetación, ofreciendo un indicativo confiable sobre la vitalidad vegetal. En este estudio, el *NDVI* fue crucial para identificar áreas de vegetación densa y detectar cambios en la cobertura vegetal del complejo de páramo TBM.

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (3.1)$$

3.2.2 Modelo digital de elevación (*DEM*)

Un Modelo Digital de Elevación (*DEM*) es una representación tridimensional de la superficie terrestre, esencial para el análisis topográfico. Este modelo es crucial para evaluar las características altimétricas de un área, que son determinantes en estudios de geología, hidrología y planificación del uso del suelo. En este estudio, se utilizó el *DEM* de la *NASA SRTM Digital Elevation 30m* de la Misión Topográfica *Shuttle Radar (Google Earth Engine, 2007)*. Esta resolución facilita una evaluación detallada de las formas del relieve y la caracterización de elementos del terreno en el páramo TBM, como se ilustra en la **Figura 3-4**.

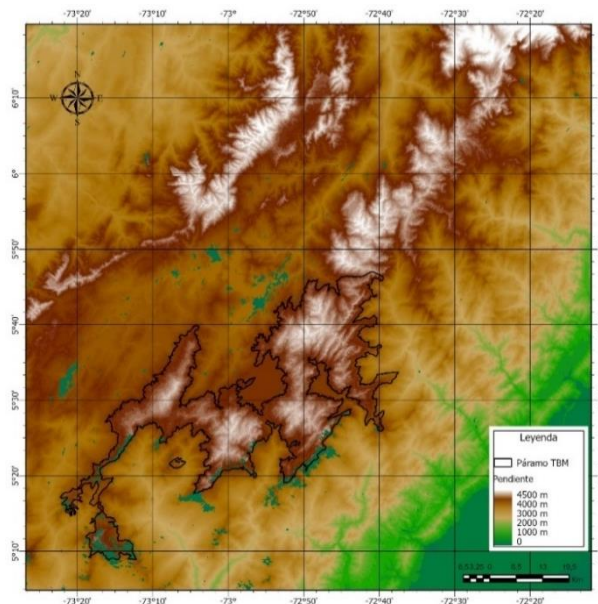


Figura 3-4: DEM del área de estudio. Fuente de datos NASA/USGS/JPL-Caltech, composición propia.

3.2.3 Etiquetas de referencia.

Para el entrenamiento de los modelos se construyeron las etiquetas de referencia a partir de la clasificación de coberturas del IDEAM del 2018. Las cuales mostraron que las principales coberturas en el área de estudio fueron *Pastos* (27%), *Áreas agrícolas heterogéneas* (23%), *Bosques* (20%) y *Herbazales* (17%). Además, un 9% corresponde a *Vegetación secundaria y transitoria* y *Arbustales*, mientras que el 6% restante incluye *Aguas continentales*, *Áreas abiertas, sin o con poca vegetación*, *Cultivos permanente*, *Cultivos transitorios* y *Áreas húmedas continentales*, (IDEAM et al., 2021). En la **Figura 3-5**, se muestra el detalle de la distribución de las coberturas.

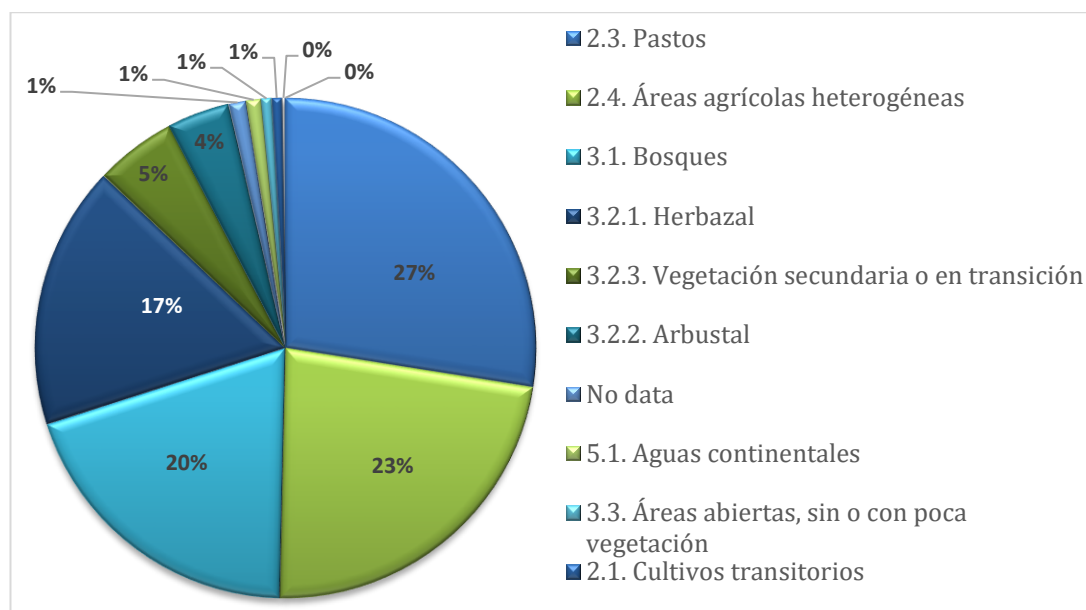



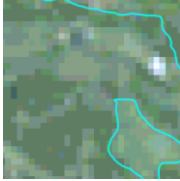








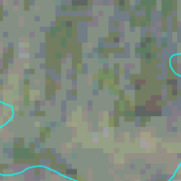

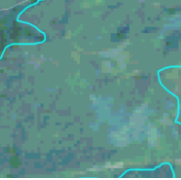




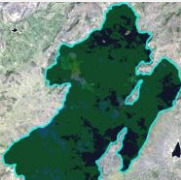




Figura 3-5 Coberturas de la tierra en el Área de estudio. Elaboración Propia basada en el estudio del IDEAM et al., (2021).

Esta distribución no solo refleja las coberturas naturales del páramo TBM, sino también los patrones de uso del suelo y prácticas agrícolas de la región. La comprensión de estas coberturas es fundamental para analizar la biodiversidad y dinámicas del páramo. Para obtener detalles específicos de cada cobertura, consulte la **Tabla 3-3**.

Tabla 3-3. Información de las Bandas de *Landsat 8*. Fuente (IDEAM, 2010).

Cobertura	Descripción	Fotografía	Composición RGB 432
2.1. Cultivos transitorios	Las áreas ocupadas por cultivos del ciclo vegetativo corto comprenden aquellos cuya duración es menor a un año, y en algunos casos, solo de unos pocos meses.		
2.2 Cultivos permanentes	Esta categoría abarca los territorios dedicados a cultivos cuyo ciclo vegetativo supera el año, lo que les permite producir varias cosechas sin necesidad de ser replantados.		
2.3. Pastos	Comprende tierras que están cubiertas con una densa capa de hierba, dominada principalmente por especies de la familia Poaceae, y que están dedicadas al pastoreo permanente durante un período de dos o más años.		
2.4. Áreas agrícolas heterogéneas	Son unidades que reúnen dos o más clases de coberturas agrícolas y naturales, en patrones geométricos complejos. Estos mosaicos presentan una disposición tan intrincada que dificulta la separación y clasificación en coberturas individuales.		
3.1. Bosques	Se refieren a zonas naturales o seminaturales caracterizadas por la presencia predominante de árboles, ya sean especies nativas o introducidas. Estas áreas están formadas por elementos arbóreos que pueden variar en densidad y tipo.		
3.2.1. Herbazal	Estas áreas están compuestas por comunidades vegetales dominadas por especies herbáceas, que se desarrollan de manera natural en distintos sustratos y con variadas densidades. Los herbazales pueden presentar una cobertura vegetal densa, con más del 70% del terreno		

Cobertura	Descripción	Fotografía	Composición RGB 432
	cubierto, o ser más abiertos, con una ocupación del 30% al 70%.		
3.2.2. Arbustal	Estas zonas incluyen territorios ocupados por vegetación de tipo arbustivo, que también se desarrollan naturalmente en diferentes densidades y sustratos. Un arbusto se define como una planta perenne, leñosa, que alcanza una altura de entre 0,5 y 5 metros, con ramificación intensa desde la base y sin una copa claramente definida.		
3.2.3. Vegetación secundaria o en transición	Se refiere a la cobertura vegetal que emerge como resultado del proceso de sucesión ecológica tras la intervención humana o la destrucción de la vegetación primaria. Este tipo de vegetación está en un proceso de recuperación, restableciéndose tras haber sido alterada o dañada diversas actividades		
3.3. Áreas abiertas, sin o con poca vegetación	Incluye territorios donde la vegetación es mínima o inexistente. Estas áreas se caracterizan por tener suelos desnudos, quemados, superficies arenosas o afloramientos rocosos. En algunos casos, también pueden abarcar regiones que están cubiertas de hielo y nieve, indicando una falta de vegetación activa debido a condiciones climáticas extremas o a la naturaleza del terreno.		
4.1. Áreas húmedas continentales	Se refieren a zonas que pueden ser inundables, pantanosas o terrenos anegadizos, donde el agua subterránea alcanza la superficie de manera temporal o permanente.		
5.1. Aguas continentales	Esta categoría abarca desde lagos, lagunas y ciénagas hasta embalses y cuerpos de agua en movimiento como ríos y canales. Estos cuerpos de agua pueden ser de origen natural o artificial y se caracterizan por ser de agua dulce, es decir, no salina.		

Cobertura	Descripción	Fotografía	Composición RGB 432
No data	se excluyeron específicamente las áreas cubiertas por nubes en las imágenes, así como las áreas artificiales que no eran relevantes para el estudio.		

Las coberturas del páramo TBM constituyen una valiosa fuente de datos para identificar y clasificar las características del terreno. La **Figura 3-6** ilustra la distribución de estas, proporcionando una visualización que facilita la comprensión de la relación entre la vegetación, las características topográficas y otros elementos ambientales.

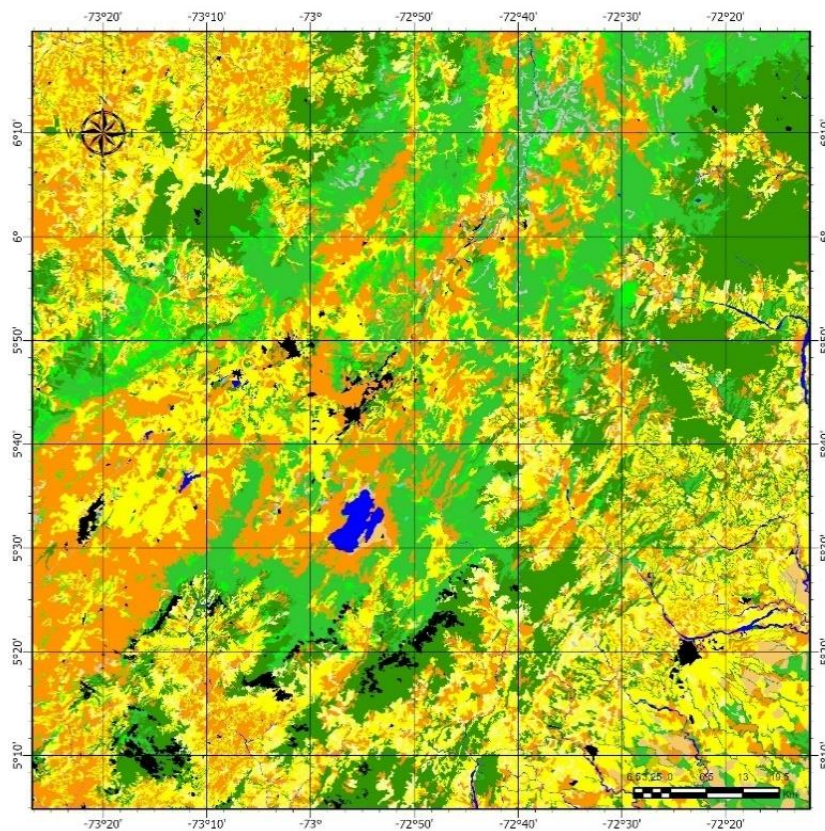


Figura 3-6: Mapa de referencia del que extraen las Etiquetas del área de estudio con los colores *RGB* para el entrenamiento. Fuente de datos IDEAM, composición propia

3.3 Métodos

La metodología aplicada en esta investigación se dividió en tres etapas claves, cada una diseñada para asegurar la exactitud temática y la eficiencia en la clasificación de coberturas del páramo TBM. A continuación, se describen las etapas, ilustradas en la **Figura 3-7**:

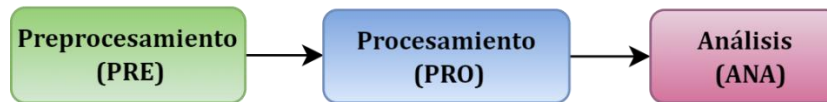


Figura 3-7: Diagrama de la metodología general de la investigación. Fuente propia.

- **Preprocesamiento (PRE):** la primera etapa se enfocó en la preparación y limpieza meticulosa de los datos. Las actividades principales incluyeron la selección cuidadosa de imágenes satelitales, corrección de errores, estandarización de imágenes y recortes, además de la generación de etiquetas. Esta fase buscó garantizar la fiabilidad y coherencia de los datos para su análisis posterior. Los detalles adicionales se muestran en la **Figura 3-8**.
- **Procesamiento (PRO):** Durante esta fase se aplicaron técnicas avanzadas de segmentación semántica a los datos ya preprocesados. Se utilizaron las arquitecturas *U-Net++* y *DeepLabV3+*, con las cuales se desarrollaron varios modelos que se fueron entrenando para optimizar las métricas de las coberturas del páramo TBM. Estos experimentos permitieron explorar y comparar la eficiencia de distintas configuraciones y ajustes en los modelos mencionados.
- **Análisis (ANA):** Esta última fase de la investigación consistió en un análisis detallado de los resultados obtenidos a través de los modelos de segmentación semántica. Se evaluó minuciosamente la exactitud de la clasificación mediante técnicas estadísticas.

3.3.1 Preprocesamiento (PRE)

Durante la fase inicial de preprocesamiento, se desarrollaron dos subetapas: **PRE1**, centrada en la preparación de las imágenes para los modelos, y **PRE2**, enfocada en la creación de etiquetas de entrada.

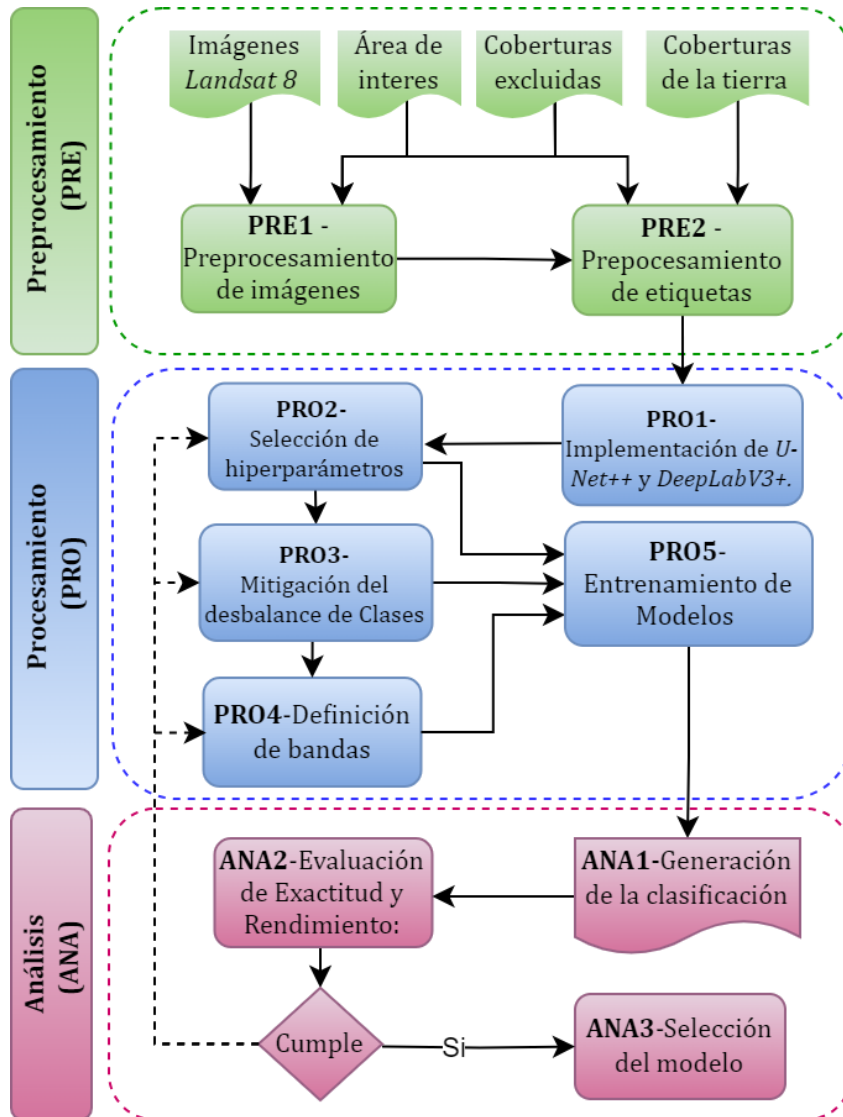


Figura 3-8: Diagrama de la metodología detallada de la investigación. Fuente propia.

Etapa *PRE1* - Preprocesamiento de Imágenes:

- Se adquirieron las imágenes Landsat en nivel 2 (reflectancia superficial) del área de estudio a través de *Google Earth Engine (GEE)*. Inicialmente, se seleccionaron imágenes al año 2018. Sin embargo, debido a la considerable presencia de nubes en el área del Páramo TBM, fue necesario ampliar el rango de fechas. Tras evaluar diferentes períodos, se determinó que las imágenes capturadas entre el 1 de diciembre de 2017 y el 31 de enero de 2019 presentaban la menor cobertura de nubes.

- Se creó una máscara específica para excluir zonas no relevantes de las imágenes y las etiquetas. Esta máscara incluyó áreas con nubes y sombras, así como superficies artificiales, clasificadas bajo el nivel 1 de CLC en las coberturas del IDEAM (2018). Una vez creada la máscara para limpiar los datos, se procedió a removerla de las imágenes, permitiendo así el análisis de las zonas de interés.
- Para evaluar el efecto de distintas combinaciones espectrales en la exactitud de la clasificación, se generaron cuatro imágenes compuestas a partir de la mediana utilizando las siguientes bandas.
 - ✓ Composición *img_432*: Utiliza las bandas *B4*, *B3* y *B2*. replicando el color verdadero para facilitar la identificación visual y el análisis de coberturas vegetales y otros elementos naturales.
 - ✓ Composición *img_765432*: Combina las bandas *B7* a *B2*, incluyendo el espectro visible e infrarrojo para una visión más amplia.
 - ✓ Composición *img_765432SE*: Añade variables topográficas como la pendiente y la elevación del terreno a la composición de las bandas *B7* a *B2*.
 - ✓ Composición *img_765432NSE*: Incorpora *NDVI*, pendiente y elevación del terreno a la combinación de bandas *B7* a *B2* para un análisis más detallado.
- Las imágenes fueron recortadas en ventanas de 256x256 y 128x128 píxeles. Las grillas (a) y (b) constan de 324 cuadrados, mientras que las grillas en (c) y (d) consta 1296 cuadrados. Además, las imágenes se dividieron de forma aleatoria para las fases de entrenamiento con un 70%, validación con un 15 % y prueba con un 15 %, como se detalla en la **Tabla 3-4**:

Tabla 3-4. Imágenes utilizadas en el entrenamiento, validación y pruebas. Fuente propia.

Imágenes	Entrenamiento 70%	Validación 15%	Prueba 15%
256*256	228	48	48
128*128	1036	129	131

Etapas PRE2 - Preprocesamiento de Etiquetas:

- Utilizando el mapa nacional de coberturas 1:100,000 del IDEAM de 2018, se recortó el área de estudio y se seleccionaron las coberturas correspondientes al

nivel 2 de *CLC*. Específicamente, para la cobertura de *3.2 Espacios de vegetación arbustiva y o herbácea*, donde se encuentran las coberturas naturales de páramo, se optó por tener un mayor detalle, utilizando las categorías del nivel 3, correspondiente a *3.2.1. Herbazal*, *3.2.2. Arbustal* y *3.2.3. Vegetación secundaria o en transición*.

- La máscara previamente creada, que incluye áreas de nubes y zonas artificiales, fue eliminada de las etiquetas para asegurar que tanto las imágenes como sus correspondientes etiquetas reflejen las mismas área.
- Cada cobertura se codificó con un valor RGB específico, en un rango de 0 a 255, para facilitar la identificación y visualización durante los análisis. Esta codificación ayuda a diferenciar claramente cada tipo de cobertura en los análisis visuales y técnicos.
- Las coberturas se transformaron a formato ráster y se normalizaron en una escala de 0 a 11, para alinearlas con las categorías de cobertura de manera consistente y prepararlas para los posteriores procesos. Los detalles de esta codificación se resumen en la **Tabla 3-5**.

Tabla 3-5. Coberturas, nivel de representatividad y el color asignado. Fuente propia

No	Cobertura	% Área	Composición RGB	Color
0	2.1. Cultivos transitorios	0,69%	250,200,100	Café claro
1	2.2 Cultivos permanentes	0,26%	200,200,0	Amarillo-verde
2	2.3. Pastos	27,50%	255, 255, 0	Amarillo
3	2.4. Áreas agrícolas heterogéneas	22,90%	255,150,0	Naranja
4	3.1. Bosques	18,37%	50, 150, 0	Verde oscuro
5	3.2.1. Herbazal	17,37%	50, 200, 50	Verde medio
6	3.2.2. Arbustal	4,28%	0, 255, 0	Verde
7	3.2.3. Vegetación secundaria o en transición	5,24%	250,250,150	Amarillo pálido
8	3.3. Áreas abiertas, sin o con poca vegetación	0,72%	200, 200, 200	Gris
9	4.1. Áreas húmedas continentales	0,01%	0, 255, 255	Cian
10	5.1. Aguas continentales	0,92%	0, 0, 255	Azul
11	No data	1,74%	0,0,0	Negro

- Las etiquetas se recortaron en ventanas de 256x256 y 128x128 píxeles, preparándolas para las fases de entrenamiento, validación y pruebas. Las etiquetas se numeraron secuencialmente de 0 a 323 y de 0 a 1295, correspondientes a los tamaños de las ventanas, facilitando su gestión y procesamiento. Su distribución se muestra en la **Figura 3-9**:

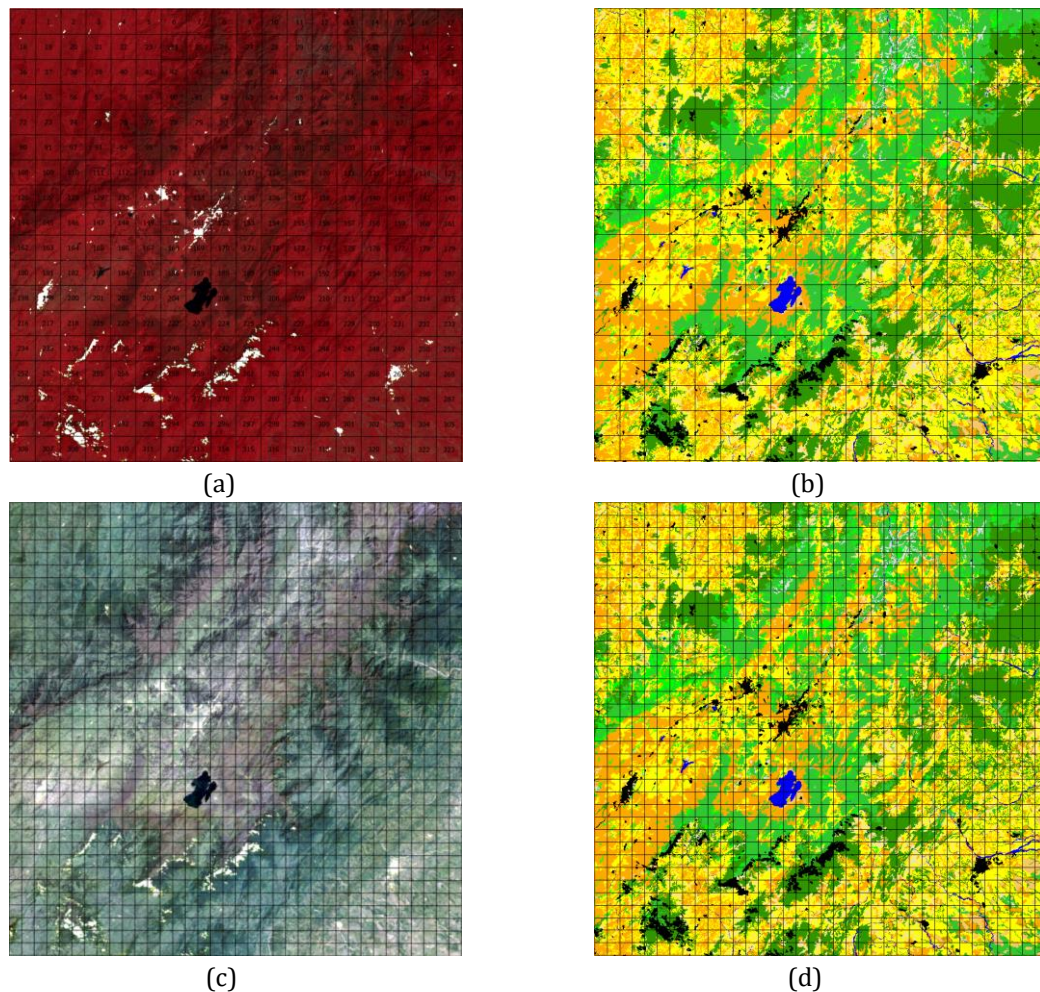


Figura 3-9: Distribución de las grillas: (a) Imagen de 256*256 píxeles en falso color *RGB-543*. (b) Etiquetas *CLC 2018* de 256*256 píxeles. (c) Imagen de 128*128 píxeles en verdadero color *RGB-432*. (d) Etiquetas *CLC 2018* de 128*128 píxeles.

3.3.2 Procesamiento (*PRO*)

Durante esta etapa crucial, se implementaron y optimizaron las arquitecturas de *CNN* para la clasificación de coberturas en el páramo TBM. La etapa se dividió en cinco procesos fundamentales: en ***PRO1*** se integró el código necesario para las arquitecturas

DeeplabV3+ y *U-Net++*, mientras que en **PRO2** se ajustaron los hiperparámetros de las redes. En **PRO3** se implementaron las técnicas para mitigar el desbalance de clases. Además, en **PRO4** se evaluó el comportamiento al aumentar el número de bandas en la imagen de entrada y en **PRO5** se centró en el entrenamiento de los modelos.

Etapa PRO1 - Implementación de U-Net++ y DeepLabV3+:

- Se configuraron las arquitecturas *DeepLabV3+* y *U-Net++* utilizando *PyTorch* en *Jupyter Lab*. Este entorno de desarrollo avanzado permite la integración eficiente de código y realización de pruebas en tiempo real, aprovechando los recursos de hardware avanzado de *NVIDIA*. Esta configuración facilita una iteración rápida y eficiente durante el proceso de desarrollo.
- Para el manejo de imágenes y procesamiento de datos, se emplearon bibliotecas especializadas como *scikit-image* para el manejo de imágenes, *numpy* para cálculos numéricos, y *matplotlib* para visualizaciones, facilitando el análisis y ajustes durante el desarrollo.

Etapa PRO2 - Selección de Hiperparámetros:

- Se optimizó la configuración de los hiperparámetros para mejorar la eficiencia del aprendizaje de los modelos. Variables críticas como la tasa de aprendizaje, función de pérdida, función de activación y la elección del optimizador fueron ajustadas para afinar la capacidad de los modelos de procesar y aprender eficientemente de los datos. Como resultado de los experimentos iniciales se seleccionaron los hiperparámetros que se muestran en la **Tabla 3-6**.
- Los experimentos iniciales indicaron problemas de sobreajuste, por lo que se implementaron técnicas de regularización para mejorar la generalización de los modelos en datos no vistos. Se aplicó el decaimiento de los pesos, que ajusta los pesos durante el entrenamiento para penalizar los valores de pesos excesivamente grandes.
- Se llevaron a cabo iteraciones utilizando variables críticas para evaluar y mejorar el rendimiento de los modelos. Esto incluyó el uso de diferentes codificadores como ResNet-18, ResNet-152, y ResNet-101, además de ajustar el número de épocas y el tamaño del lote. Estos elementos fueron fundamentales para determinar el mejoramiento en la exactitud de los modelos.

Tabla 3-6. Hiperparámetros usados para la segmentación semántica de imágenes. Fuente propia.

Hiper-parámetro	Valor	Descripción
Tasa de Aprendizaje	0.0001	Este valor controla la rapidez con la que el modelo aprende. Una tasa de aprendizaje baja permite ajustes más exactos en los pesos del modelo durante el entrenamiento, lo que puede mejorar la exactitud del modelo, pero también aumentar el tiempo de entrenamiento.
Función de Pérdida	Entropía cruzada ponderada	Es una función de pérdida común en problemas de clasificación. La ponderación de cada clase es particularmente útil en escenarios donde existe un desequilibrio significativo entre las clases, ayudando a evitar que el modelo se incline excesivamente hacia las clases más representadas.
Función de activación	<i>softmax2d</i>	Adecuada para problemas de clasificación con múltiples clases. Esta función convierte las salidas del modelo en probabilidades, facilitando la asignación de cada píxel a una de las coberturas.
Optimizador	<i>Adam</i> $1 e^{-3}$	El optimizador Adam combina las ventajas de dos extensiones del algoritmo de descenso de gradiente: <i>Momentum</i> y <i>RMSprop</i> . Adam ajusta la tasa de aprendizaje durante el entrenamiento para cada parámetro, mejorando la eficiencia y exactitud del modelo.
Regularización	Decaimiento de los pesos $1 e^{-4}$	La regularización <i>L2</i> o decaimiento de pesos agrega una penalización basada en la norma <i>L2</i> de los pesos a la función de pérdida. Esta técnica ayuda a prevenir el sobreajuste, asegurando que los pesos del modelo no se vuelvan demasiado grandes y que el modelo generalice mejor los datos nuevos.

- Se realizaron pruebas con imágenes de diferentes dimensiones, específicamente de 256x256 y 128x128 píxeles, junto con sus correspondientes etiquetas. Este análisis permitió evaluar cómo los diferentes tamaños de imagen impactan el rendimiento de los modelos. Este paso es crucial para determinar la configuración óptima que logre un equilibrio adecuado entre exactitud y eficiencia computacional.

Etapa PRO3 – Mitigación del desbalance de clases:

Para abordar el desbalance de clases en la clasificación de coberturas de páramo mediante imágenes satelitales, se implementaron tres enfoques estratégicos principales, cada uno orientado a optimizar la representatividad y exactitud de las coberturas en el modelo:

- Expansión del área de estudio: Se realizaron pruebas preliminares utilizando el área de influencia de 2 km alrededor del páramo TBM. Al no proporcionar una representación adecuada de todas las coberturas necesarias y proporcionar bajos $F1$ para las coberturas menos representadas, se amplió el área de estudio hacia el complejo de páramos en la zona nororiente, asegurando así una mayor diversidad y representatividad en el análisis. Esta expansión busca equilibrar el número de etiquetas por clase, generando un conjunto de datos más representativo y balanceado.

Ampliación del conjunto de datos: Inicialmente se duplicaron las imágenes con coberturas menos representativas en un intento de equilibrar el conjunto de datos. Sin embargo, este método no produjo los resultados esperados en las métricas de $F1$, lo que indica que duplicar los datos existentes no contribuía significativamente al aprendizaje de los modelos, ya que no introducía nuevas características que pudieran ayudar al modelo a generalizar mejor. Por lo que se optó por aplicar transformaciones aleatorias utilizando la librería *Albumentations*, generando 1036 nuevas imágenes y etiquetas. Este enfoque incluyó diferentes transformaciones como rotación, inversión horizontal, recortes aleatorios, ajustes de brillo y contraste, entre otras. Entre las técnicas utilizadas se encuentran:

- ✓ *RandomRotate90*: Rota aleatoriamente una imagen de 90 grados, introduciendo variaciones en la orientación.
- ✓ *HorizontalFlip*: Voltea horizontalmente una imagen, mejorando la capacidad de los modelos para generalizar a diferentes orientaciones de objetos.
- ✓ *RandomResizedCrop*: Recorta y cambia el tamaño de una imagen aleatoriamente.
- ✓ *RandomBrightnessContrast*: Ajusta aleatoriamente el brillo y el contraste de una imagen, simulando diferentes condiciones de iluminación.
- ✓ *ElasticTransform*: Aplica deformación elástica a una imagen, introduciendo distorsiones locales.
- ✓ *GaussNoise*: Añade ruido gaussiano a una imagen, mejorando la robustez de los modelos al ruido en imágenes del mundo real.
- ✓ *Blur*: Aplica un filtro de desenfoque a una imagen, útil para reducir ruido o suavizar detalles.
- ✓ *ChannelDropout*: Establece aleatoriamente en cero una fracción de canales de una imagen, simulando canales faltantes o dañados.

- Asignación de pesos específicos por cobertura: Se utilizó inicialmente la inversa de la frecuencia de las clases como método para ponderar los pesos durante el entrenamiento de los modelos. Este método destacó una sobrerrepresentación de ciertas coberturas, mientras que otras, con una presencia menor al 6%, no fueron adecuadamente estimadas. Al ajustar los pesos para aumentar la influencia de las coberturas menos representadas y reducir las más frecuentes, se observó una disminución en las métricas *F1* de las coberturas más comunes, indicando un desequilibrio en el modelo. Para abordar este problema, se realizaron múltiples iteraciones ajustando los valores de los pesos, en un esfuerzo por encontrar un equilibrio que mejorara tanto la representatividad de las clases menos frecuentes como el *F1* de todas las coberturas.

Etapa PRO4 - Definición de bandas:

En esta fase, se adaptaron las arquitecturas de *CNN* para que los modelos pudieran procesar imágenes multiespectrales.

- Las arquitecturas *U-Net++* y *DeepLabV3+*, se modificaron para aceptar múltiples bandas espectrales. Esto implicó cambios en la entrada de las redes para manejar datos más complejos y ricos en información.
- Los pesos preentrenados de *ImageNet* (reconocida base de datos que contiene 1.000 clases de objetos y más de un millón de imágenes) no se adaptan bien a las características específicas de las imágenes satelitales multiespectrales. Por lo que se calcularon nuevos pesos a partir de la media y la desviación estándar de cada banda y se utilizaron como valores iniciales para el entrenamiento de los modelos. Este paso es crucial para la normalización de los datos, asegurando que las entradas a la red neuronal tengan una distribución consistente, lo que facilita el aprendizaje.
- Se llevó a cabo la evaluación de la exactitud temática y tiempos de ejecución al incluir más bandas en el modelo. Esto permitió identificar el aporte de cada banda en la clasificación.

Etapa PRO5 - Entrenamiento de modelos:

- El proceso de entrenamiento incluye una serie de experimentos iterativos donde se definieron algunos hiperparámetros como la tasa de aprendizaje, el tamaño de

lote, y el número de épocas. Esto se realizó para encontrar la configuración más adecuada que maximiza la exactitud de los modelos sin comprometer demasiado los tiempos de ejecución.

- Los experimentos iniciales se realizaron sin pesos preentrenados, posteriormente se incluyeron los pesos preentrenados de ImageNet y por último se calcularon los pesos al usar las imágenes multiespectrales.
- Se ejecutaron los experimentos con diferentes tamaños de las imágenes de entrada y la selección de bandas espectrales. Estos experimentos ayudan a entender mejor cómo cada variable afecta el rendimiento de los modelos en términos de exactitud y eficiencia computacional.

3.3.3 Análisis (ANA)

En esta etapa se validó la eficiencia de los modelos entrenados en la clasificación de coberturas del páramo TBM y se seleccionó el más adecuado. Se realizaron tres procesos: en **ANA1**, se generó la clasificación de las coberturas del páramo TBM; en **ANA2** se evaluó la exactitud temática y el rendimiento de cada modelo. Finalmente, en **ANA3**, se seleccionó el modelo más adecuado.

ANA1 – Generación de la clasificación:

- Se realizó la clasificación de las coberturas del páramo TBM utilizando diferentes modelos *CNN*. Teniendo en cuenta los parámetros seleccionados previamente durante la etapa de entrenamiento.

ANA2 - Evaluación de Exactitud y Rendimiento:

- Se reportó los resultados de prueba de los modelos usando métricas generales como *OA* y *Kappa*, que ayudan a determinar cuán bien los modelos predijeron las coberturas en comparación con las etiquetas.
- Se utilizó el *F1* y *IoU* para evaluar la clasificación de cada tipo de cobertura. Estas métricas proporcionan una comprensión más detallada de la eficiencia de los modelos en clasificar correctamente las diversas coberturas, destacando áreas de fortaleza y aquellas que requieren mejoras.

- También se consideró la eficiencia de los modelos en términos de recursos computacionales necesarios y el tiempo requerido para el entrenamiento y la inferencia, asegurando que el modelo seleccionado sea viable para implementaciones prácticas.

ANA3 - Selección del Modelo:

- Los modelos seleccionados de cada arquitectura fueron comparados meticulosamente. Se revisaron las métricas de exactitud general y específica por cobertura, así como la eficiencia computacional de cada uno. Esto permitió identificar cuál de los modelos ofrecía un mejor equilibrio entre exactitud y tiempo.

3.3.4 Herramientas

Durante la investigación se utilizaron diferentes herramientas y software, cada uno desempeñando funciones esenciales en diferentes etapas del proceso. Estas herramientas incluyeron plataformas de procesamiento de datos, lenguajes de programación, bibliotecas de aprendizaje profundo, y aplicaciones para la manipulación y visualización de datos.

La integración de estas herramientas facilitó el manejo eficiente de los conjuntos de datos, permitió la implementación efectiva de las arquitecturas de CNN, y aseguró una evaluación eficiente de los modelos junto con una presentación clara y detallada de los resultados. En la **Tabla 3-7** se muestra el detalle.

Tabla 3-7. Herramientas y software y sus propósitos respectivos. Fuente propia.

Herramientas / Software	Actividad	Descripción
<i>Google Earth Engine -GEE</i>	<i>PRE1</i>	Utilizado para seleccionar y realizar el preprocesamiento de las imágenes. <i>GEE</i> es una plataforma de análisis geoespacial que facilita el acceso a un amplio conjunto de datos satelitales
<i>ArcGIS Pro</i>	<i>PRE1, PRE2</i>	<i>ArcGIS Pro</i> -Versión 3.0.5 fue empleado para construir los <i>shapefiles</i> necesarios para el proyecto, así como para crear máscaras de nubes y transformar las etiquetas en imágenes.
<i>Jupyter lab</i>	<i>PRO1, PRO2, PRO3, PRO4, PRO5, ANA1, ANA2.</i>	Este entorno de desarrollo interactivo permitió la creación y ejecución de notebooks, combinando código, texto explicativo y elementos gráficos, lo cual fue esencial para la experimentación y documentación de los procesos.

Herramientas / Software	Actividad	Descripción
Python	<i>PRO1, PRO2, PRO3, PRO4, PRO5, ANA1, ANA2.</i>	Se utilizó para procesar los conjuntos de datos de entrenamiento, validación y prueba, cortar las imágenes, aumentar imágenes, entrenar y evaluar los modelos con arquitecturas <i>CNN</i> .
PyTorch	<i>PRO1, PRO2, PRO3, PRO4, PRO5, ANA1, ANA2.</i>	Esta biblioteca de aprendizaje profundo fue utilizada para la implementación de arquitecturas de redes neuronales, proporcionando una base sólida para el desarrollo de los modelos. Desarrollada por el Laboratorio de Investigación de Inteligencia Artificial de Facebook (<i>FAIR</i>).
Segmentation-models-pytorch	<i>PRO1, PRO2, PRO3, PRO4, PRO5, ANA1.</i>	Esta biblioteca proporcionó modelos preentrenados y arquitecturas de segmentación de imágenes para facilitar la implementación de redes neuronales en <i>PyTorch</i> .
Scikit-learn - sklearn.metrics	<i>ANA2.</i>	Se utilizó para tareas de aprendizaje automático y evaluación de modelos, proporcionando funciones y métricas clave para determinar la calidad y el rendimiento de los modelos de aprendizaje automático.
NumPy	<i>PRO1, PRO2, PRO3, PRO5, ANA1</i>	Biblioteca fundamental para la manipulación de matrices en <i>Python</i> , esencial para operaciones numéricas y el preprocesamiento de datos.
Matplotlib	<i>PRO1, PRO5, ANA1</i>	Utilizada para la generación de gráficos y visualización de datos, facilitando la representación visual de los resultados obtenidos.
Albumentations	<i>PRO3</i>	Librería para aumentación de datos en imágenes, utilizada para enriquecer el conjunto de entrenamiento y mejorar la generalización de los modelos.
Servidor NVIDIA Quadro RTX™ 5000	<i>PRO1, PRO2, PRO3, PRO4, PRO5, ANA1</i>	Equipado con hardware potente, incluyendo núcleos <i>CUDA</i> equipada con 3072 núcleos <i>CUDA</i> , 384 núcleos Tensores, 48 núcleos <i>RT</i> y memoria <i>GDDR6</i> de 16 <i>GB</i> . este servidor fue crucial para ejecutar scripts de <i>Python</i> que requerían grandes recursos computacionales y tiempo de procesamiento.

4. RESULTADOS

Este capítulo detalla los experimentos realizados durante la investigación, el cual se está dividido en cuatro secciones. Comienza con la selección de los hiperparámetros, que establece la base para los experimentos subsiguientes. A continuación, se describen los experimentos enfocados en abordar el desbalance de clases, seguido por aquellos que buscan determinar la combinación óptima de bandas espectrales. Finalmente, se presentan las comparaciones de los resultados de los modelos seleccionado, destacando las ventajas y desventajas de cada uno.

4.1 Selección de Hiperparámetros

Durante la etapa de ajuste de hiperparámetros, se realizaron múltiples iteraciones en las cuales se variaron aspectos clave como los codificadores, los tamaños de las imágenes de entrada, los tamaños de lote y el número de épocas. Estos ajustes se hicieron con el objetivo de identificar la configuración más eficiente para el procesamiento de imágenes RGB, adaptada específicamente a las necesidades y características de los páramos. Cada iteración se evaluó rigurosamente para determinar cómo estas variables influían en la eficiencia de los modelos, buscando optimizar tanto la calidad de la clasificación como el tiempo de procesamiento.

A partir de los resultados obtenidos en las métricas generales, *OA* y *Kappa* se destacaron los modelos *U-Net++-4* y *DLV3+-4*. La evaluación detallada, presentada en la **Tabla 4-1**, reveló que *ResNet-101* presenta mejores resultados comparada con *ResNet-18* y *ResNet-152*. Se observó que imágenes de menor tamaño presentaron un mejor rendimiento en términos de *OA* y *Kappa*, indicando que dimensiones reducidas pueden facilitar un procesamiento más eficiente.

Reducir el tamaño del lote resultó en un incremento en el tiempo de entrenamiento sin mejorar significativamente las métricas generales. Utilizar modelos con pesos preentrenados de *ImageNet* redujo el tiempo de procesamiento, dado que aprovecha los conocimientos previos para acelerar la convergencia. Además, el número de épocas

demonstró tener un impacto significativo en el tiempo de entrenamiento, siendo un factor crucial para alcanzar la convergencia de los modelos sin incurrir en sobreajuste.

Tabla 4-1. Hiperparámetros de *U-Net++* y *DeepLabV3+*. En color verde se resaltan los modelos seleccionados. Fuente propia.

Modelo	Codificador	Tamaño imagen	Tamaño del Lote	Épocas	Pre - entrenado	OA	Kappa	Tiempo
<i>U-Net++-1</i>	<i>ResNet-152</i>	256x256	8	20	No	51,85%	0,37	36,12
<i>U-Net++-2</i>	<i>ResNet-18</i>	128x128	8	20	No	59,87%	0,48	38,48
<i>U-Net++-3</i>	<i>ResNet-152</i>	128x128	4	30	SI	62,83%	0,52	42,35
<i>U-Net++-4</i>	<i>ResNet-101</i>	128x128	8	40	Si	63,35%	0,52	48,35
<i>DLV3+ 1</i>	<i>ResNet-152</i>	256x 256	4	25	No	59,48%	0,49	35,12
<i>DLV3+ 2</i>	<i>ResNet-18</i>	128x128	8	20	No	59,56%	0,48	32,58
<i>DLV3+-3</i>	<i>ResNet-152</i>	128x128	4	20	No	60,25%	0,49	36,23
<i>DLV3+-4</i>	<i>ResNet-101</i>	128x128	8	40	Si	62,67%	0,52	42,18

Se identificó un desbalance significativo en varias coberturas particularmente en 2.2 *Cultivos permanentes*, 3.2.2. *Arbustal*, 3.2.3. *Vegetación secundaria o en transición*, 3.3. *Áreas abiertas, sin o con poca vegetación*, 4.1. *Áreas húmedas continentales* y 5.1. *Aguas continentales*. Estas coberturas presentaron una representatividad limitada en los datos recolectados, lo que resultó en una clasificación inadecuada por parte de los modelos. Los resultados de estos análisis se muestran en la **Tabla 4-2**.

Tabla 4-2. Métricas de las coberturas de *U-Net++-4* y *DLV3+-4*. Fuente propia.

Cobertura / Métricas de modelo	<i>U-Net++-4</i>		<i>DLV3+-4</i>	
	<i>F1</i>	<i>IoU</i>	<i>F1</i>	<i>IoU</i>
2.1. Cultivos transitorios	0,20%	0,001	0,20%	0,001
2.2 Cultivos permanentes	0%	0	0%	0
2.3. Pastos	68,20%	0,52	67,00%	0,50
2.4. Áreas agrícolas heterogéneas	54,80%	0,38	60,70%	0,44
3.1. Bosques	74,00%	0,59	72,40%	0,57
3.2.1. Herbazal	73,50%	0,58	72,50%	0,57
3.2.2. Arbustal	0%	0	0%	0
3.2.3. Vegetación secundaria o en transición	0%	0	0%	0
3.3. Áreas abiertas, sin o con poca vegetación	0%	0	0%	0
4.1. Áreas húmedas continentales	0%	0	0%	0
5.1. Aguas continentales	0%	0	0%	0

4.2 Mitigación del desbalance de clases

Para mitigar el desbalance de clases en la clasificación de coberturas de páramos, se implementaron tres enfoques estratégicos. Se amplió el área de estudio, se aumentaron de datos y se asignaron de pesos por cobertura. Estas medidas mejoraron significativamente la capacidad de los modelos para equilibrar el desbalance de clases, reflejado en un aumento de la de AO y $Kappa$ para las coberturas menos representadas. Los resultados detallados del modelo $U-Net++_4$ se muestran en la **Tabla 4-3**, donde se especifican los pesos seleccionados para cada cobertura.

Tabla 4-3. Métricas generales de las iteraciones de los pesos de las coberturas en el modelo $U-Net++_4$. Fuente propia.

Modelo	$U-Net++_4_1$	$U-Net++_4_2$	$U-Net++_4_3$	$U-Net++_4_4$
Parámetros				
Datos aumentados	No	No	No	Si
Pesos de clases	1-1-0,0 4-0,04-0,05-0,06-0,26-0,19-1-1-1-0,60	1-1-0,1-0,1-0,1-0,15-0,5-0,4-1-1-1-0,6	1-1-0,2-0,2-0,2-0,25-0,6-0,5-1-1-1-0,7	1-1-0,2-0,2-0,25-0,3-0,6-0,5-1-1-1-0,7
Píxeles	2146304	2146304	2146304	2146304
OA	55,30%	58,04%	62,44%	62,56%
$Kappa$	0,47	0,48	0,53	0,53

El modelo $U-Net++_4_4$ demostró un mejor rendimiento en la clasificación de coberturas de páramos. En particular, este modelo logró los mejores resultados de $F1$ para las coberturas de 3.1. *Bosques* con un 70,10%, y 3.2.1. *Herbazal* con un 75,05%. No obstante, se encontraron dificultades en la clasificación de coberturas más heterogéneas, como con las coberturas de 2.4. *Áreas agrícolas heterogéneas*, con un 52,18%, y 3.2.3. *Vegetación secundaria o en transición*, con un 40,47% y particularmente coberturas como 2.1. *Cultivos transitorios*, 2.2 *Cultivos permanentes*, 3.3. *Áreas abiertas, sin o con poca vegetación* y 4.1. *Áreas húmedas continentales* no se obtuvieron muestra suficientes para que el modelo pudiera clasificarlas.

Es importante destacar que, aunque el modelo $U-Net++_4_4$ muestra un mejor rendimiento en la segmentación de áreas extensas y homogéneas, enfrenta mayores dificultades al tratar con coberturas heterogéneas. Estas dificultades fueron especialmente evidentes en áreas donde las coberturas varían dependiendo de la temporalidad de la toma de la imagen. Para más detalles, consulte la **Tabla 4-4**.

Tabla 4-4. Métricas por coberturas de las iteraciones de los pesos de las coberturas en el modelo *U-Net++_4*. Fuente propia.

Modelo		<i>U-Net++_4_1</i>	<i>U-Net++_4_2</i>	<i>U-Net++_4_3</i>	<i>U-Net++_4_4</i>
Métrica					
F1	2.1	13,08%	0,56%	0,25%	0,00%
	2.2	0,00%	0,00%	0,00%	0,00%
	2.3.	62,17%	65,78%	67,81%	69,54%
	2.4.	51,70%	49,68%	58,72%	52,18%
	3.1.	66,94%	67,95%	71,35%	70,10%
	3.2.1.	69,08%	72,22%	73,41%	75,05%
	3.2.2.	36,75%	37,09%	40,63%	41,10%
	3.2.3.	32,33%	32,58%	32,64%	40,47%
	3.3.	0,00%	0,00%	0,00%	0,00%
	4.1.	0,00%	0,00%	0,00%	0,00%
	5.1.	46,85%	26,46%	0,00%	31,25%
	No data	83,37%	85,00%	86,73%	94,97%

En la **Figura 4-1**, los recuadros rojos destacan variaciones de cobertura donde el modelo no clasificó de manera correcta, para facilitar la interpretación, en la etiqueta de cada área se agregó el código de cobertura según *CLC*, y su correspondiente inferencia. Estos recuadros destacan las limitaciones al clasificar las delimitaciones detalladas de algunas coberturas, y la incorporación de pequeñas coberturas dentro de grandes áreas.

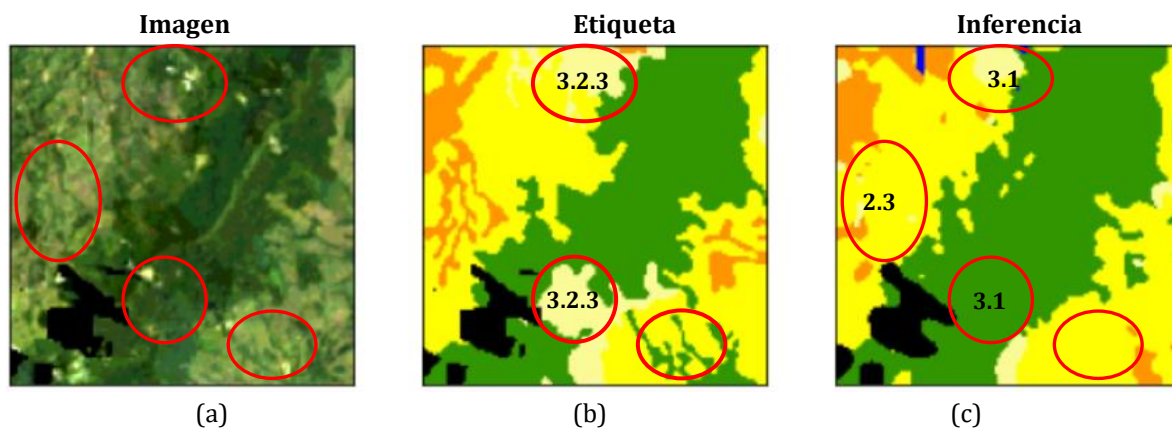


Figura 4-1. Resultados de la clasificación de *U-Net++_4_4* con la imagen *img_432*. (a) Imagen, (b) Etiqueta y (c) Inferencia. Fuente propia.

Los resultados detallados del modelo *DLV3+_4* se resumen en la **Tabla 4-5**, mostrando los pesos seleccionados para cada cobertura. El modelo *DLV3+_4* se destacó particularmente en la clasificación de coberturas de páramos, alcanzando los mejores resultado en coberturas de *3.1. Bosques*, con un 73,69%, y *3.2.1. Herbazal*, con un 80,40%. y particularmente coberturas como *2.1. Cultivos transitorios*, *2.2 Cultivos permanentes* y *4.1. Áreas húmedas continentales* no se obtuvieron muestra suficientes para clasificarlas. Para más detalles, consulte la **Tabla 4-6**.

Tabla 4-5. Métricas generales de las iteraciones de los pesos de las coberturas en el modelo *DLV3+_4*. Fuente propia.

Modelo Parámetros	<i>DLV3+_4_1</i>	<i>DLV3+_4_2</i>	<i>DLV3+_4_3</i>
Datos aumentados	No	No	Si
Pesos de clases	1-1-0,1-0,1- 0,1- 0,15-0,5- 0,4- 1- 1- 1- 0,6	1-1-0,2-0,2- 0,2- 0,25-0,6- 0,5- 1- 1- 1- 0,7	1- 1- 0,2-0,2- 0,25- 0,3-0,6- 0,5- 1- 1- 1- 0,7
Pixeles	2146304	2146304	2146304
OA	58%	62,14%	63,55%
Kappa	0,49	0,53	0,54

Tabla 4-6. Métricas por coberturas de la iteración de los pesos de las coberturas en el modelo *DLV3+_4*. Fuente propia.

Métrica	Modelo	<i>DLV3+_4_1</i>	<i>DLV3+_4_2</i>	<i>DLV3+_4_3</i>
F1	2.1	1,02%	0,25%	0,00%
	2.2	0,00%	0,00%	0,00%
	2.3.	62,69%	67,18%	70,15%
	2.4.	59,21%	58,44%	53,70%
	3.1.	65,12%	72,37%	73,69%
	3.2.1.	71,71%	73,26%	80,40%
	3.2.2.	37,61%	43,02%	27,17%
	3.2.3.	32,19%	31,20%	29,93%
	3.3.	31,07%	0,00%	18,35%
	4.1.	0,00%	0,00%	0,00%
	5.1.	28,81%	32,51%	35,44%
	No data	78,63%	74,98%	97,62%

Aunque el modelo *DLV3+_4_3* mostró un rendimiento similar al modelo *U-Net++_4_4* y se enfrentó a retos significativos al clasificar coberturas más heterogéneas, como es el caso de la cobertura de 3.2.2. *Arbustal*, donde el *F1* resultó bajo a pesar de las variaciones en la ponderación de los pesos por cobertura. Estos desafíos fueron particularmente evidentes en áreas con coberturas heterogéneas y patrones de vegetación complejos, lo que impactó negativamente el desempeño general del modelo, resultando en un *AO* de 63,55%. En la **Figura 4-2**, los recuadros rojos ilustran variaciones de cobertura donde el modelo no logro clasificar correctamente.

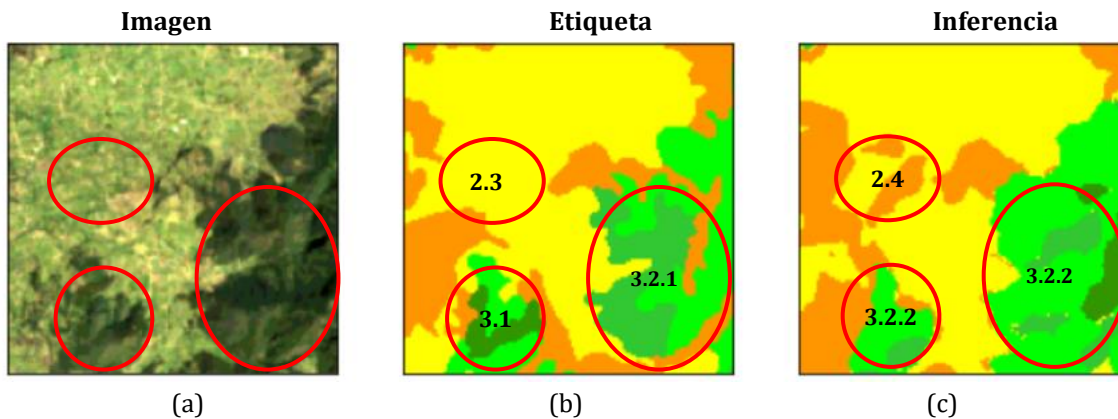


Figura 4-2: Resultados de la clasificación del modelo *DLV3+_4_3* con la imagen *img_432*. (a) Imagen, (b) Etiqueta y (c) Inferencias. Fuente propia.

Basándose en los experimentos realizados, la selección de pesos para los modelos *U-Net++_4* y *DLV3+_4* se orientó a balancear la habilidad del modelo para identificar coberturas con diferentes grados de representatividad, los cuales se resumen en la Tabla 4-7, es importante mencionar que estos valores se establecieron para lograr un equilibrio entre las clases más y menos representativas. Este proceso tomó como referencia los resultados de *F1* por cobertura del modelo inicial donde todos los pesos eran iguales. Además, este enfoque consideró el comportamiento específico de las coberturas de páramos, tratando de evitar que el modelo se sesgue hacia las coberturas más representativa.

Tabla 4-7. Pesos asignados a las diferentes coberturas para afrontar el desbalance de clases. Fuente propia.

Cobertura	Pesos	Cobertura	Pesos
2.1. Cultivos transitorios	1,0	3.2.2. Arbustal	0,6
2.2 Cultivos permanentes	1,0	3.2.3. Vegetación secundaria o en transición	0,5
2.3. Pastos	0,2	3.3. Áreas abiertas, sin o con poca vegetación	1,0
2.4. Áreas agrícolas heterogéneas	0,2	4.1. Áreas húmedas continentales	1,0
3.1. Bosques	0,3	5.1. Aguas continentales	1,0
3.2.1. Herbazal	0,3	No data	0,7

Se identificó que algunas coberturas no tenían suficiente representatividad, por lo que se evaluaron dos puntos claves. Primero, el área de estas coberturas dentro del estudio no superaba el 1%, siendo menor su porcentaje dentro del Páramo TBM y algunas, como 4.1. *Áreas húmedas continentales*, ni siquiera estaban presentes. El segundo aspecto que se consideró fue que estas coberturas no eran predominantes en los páramos colombianos, según lo establece el IDEAM et al., (2014), donde los *Herbazales* (33,9%), *Arbustos* (23,2%) y *Bosques* (20,7%) eran más comunes. Por estas razones, las de 2.1. *Cultivos transitorios*, 2.2 *Cultivos permanentes*, 3.3. *Áreas abiertas, sin o con poca vegetación* y 4.1. *Áreas húmedas continentales*, no se consideraron en los posteriores análisis. La **Figura 4-3** compara el *F1* de las coberturas incluida en la clasificación de los modelos *U-Net++-4* y *DLV3+-4*.

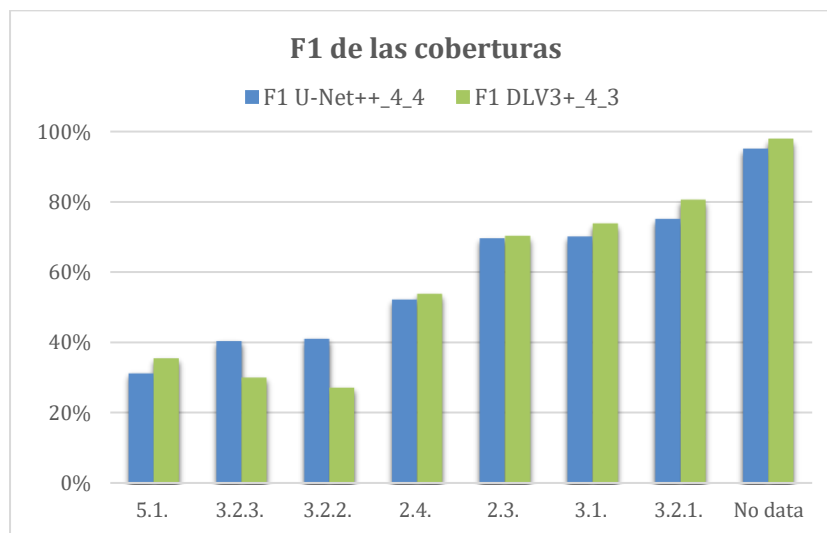
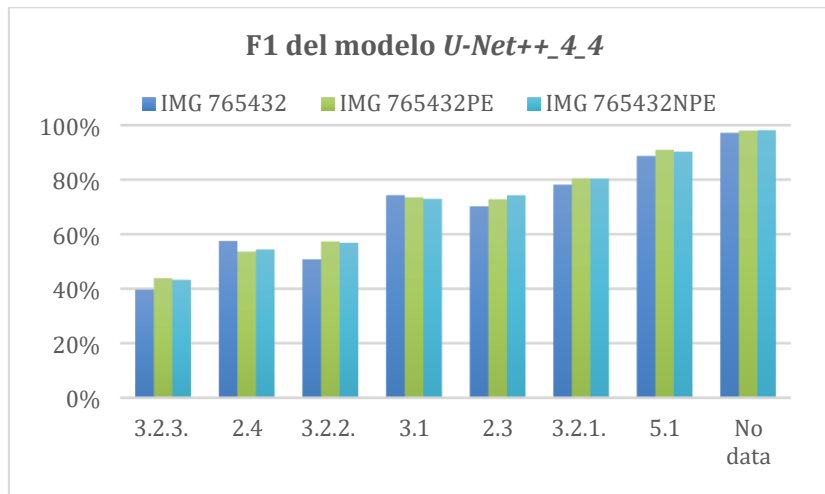


Figura 4-3 F1 de las coberturas representativas usando los modelos *U-Net++-4* y *DLV3+-4*. Fuente propia.

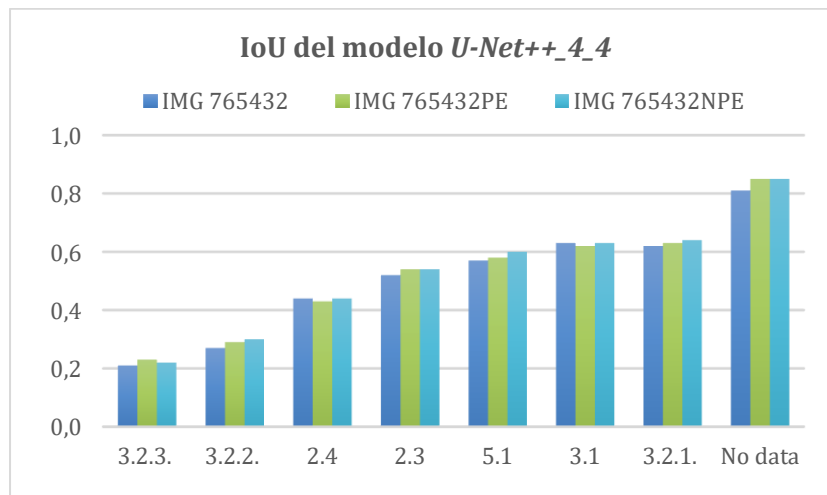
4.3 Definición del número de bandas

Para la adaptación de los modelos al uso de imágenes multiespectrales, se realizaron varios cambios significativos, debido a que las arquitecturas originales estaban diseñadas principalmente para procesar bandas *RGB*. Se ajustaron los pesos de las bandas de entrada calculando la media y la desviación. Adicionalmente, se incrementó el número de épocas de entrenamiento de 40 a 50, permitiendo un aprendizaje más exhaustivo por parte del modelo. Y se utilizaron de imágenes *img_765432*, *img_765432PE* e *img_765432NPE*, para evaluar el impacto de aumentar el número de bandas en las métricas de desempeño de los modelos.

El modelo *U-Net++_4_4*, mostró un incremento progresivo en el *F1* al aumentar el número de bandas, excepto en las coberturas *2.4. Áreas agrícolas heterogéneas* y *3.1. Bosques*, donde disminuyeron. La coberturas con el mejor *F1* fueron *3.1. Bosques*, *3.2.1. Herbazal* y *5.1. Aguas continentales*. En cuanto al *IoU*, esta métrica penaliza más las inferencias, con valores más bajos, especialmente en las coberturas de *5.1. Aguas continentales*. La mayoría de las coberturas experimentaron un aumento gradual de *IoU*, a excepción de *2.4. Áreas agrícolas heterogéneas* y *3.1. Bosques*, cuyos valores permanecieron constantes. Las coberturas con las inferencias más destacadas según el *IoU* fueron *5.1. Aguas continentales*, *3.1. Bosques*, *3.2.1. Herbazal*, respectivamente, como se puede ver en la **Figura 4-4**.



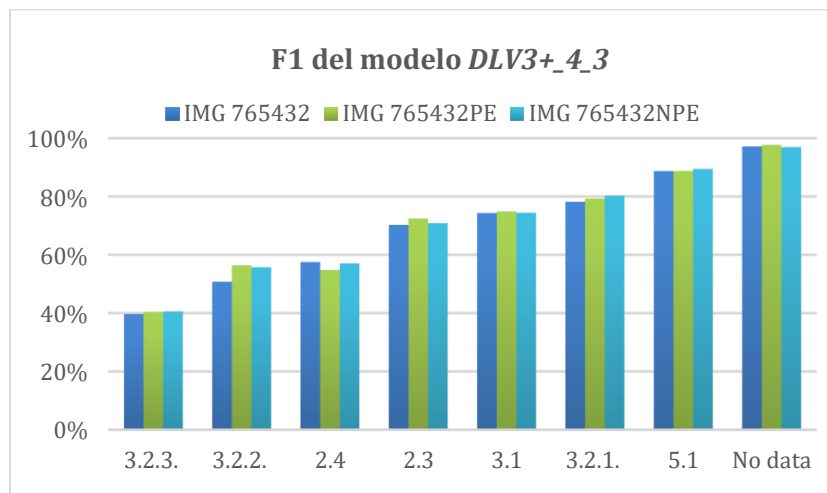
(a)



(b)

Figura 4-4: Métricas de las coberturas del modelo *U-Net++_4_4* (a) *F1* y (b) *IoU*. En el eje x se indican los códigos de las diferentes coberturas. Fuente propia.

En el modelo *DLV3+_4_3*, se observó un aumento progresivo en el *F1* con el incremento de bandas, aunque se registraron disminuciones en las coberturas *2.4. Áreas agrícolas heterogéneas* y *3.1 Bosques*. En contraste, las coberturas que mostraron la mejor exactitud en las inferencias fueron *3.1. Bosques*, *3.2.1. Herbazal* y *5.1. Aguas continentales*. En cuanto al *IoU*, este mostró una penalización más rigurosa de las inferencias, con valores más bajos, especialmente en *5.1. Aguas continentales*. A excepción de *2.4. Áreas agrícolas heterogénea* que disminuyó las otras coberturas presentaron un incremento gradual. Las coberturas con las mejores inferencias fueron, en este orden, *5.1. Aguas continentales*, *3.1. Bosques* y *3.2.1. Herbazal*, como se puede ver en la **Figura 4-5**.



(a)

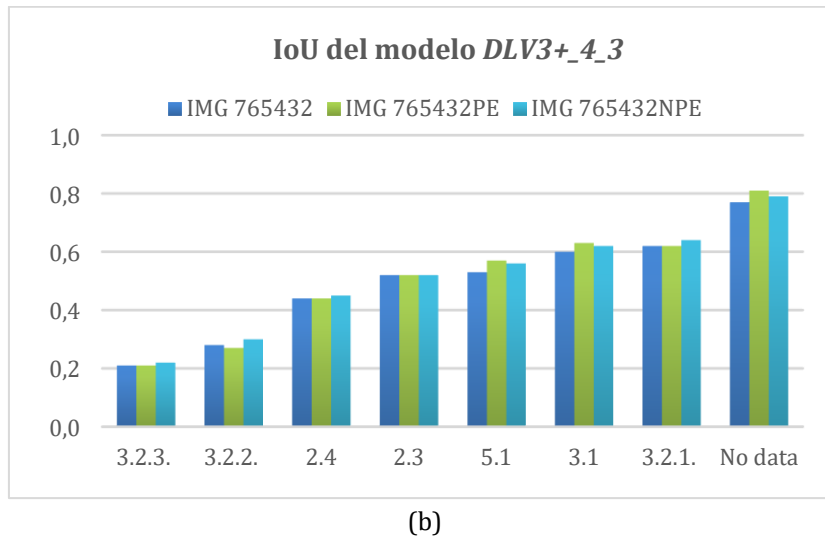


Figura 4-5: Métricas de las coberturas del modelo *DLV3+_4_3* (a) *F1* y (b) *IoU*. En el eje x se indican los códigos de las diferentes coberturas. Fuente propia

Se registró un incremento progresivo en los tiempos de ejecución a medida que se incrementaban las bandas espectrales durante el entrenamiento e inferencia del modelo. Este aumento en los tiempos de procesamiento es atribuible al aumento de cálculos computacionales necesaria para manejar un mayor volumen de datos. Cada banda adicional intensifica la carga en las operaciones de convolución y las transformaciones requeridas durante el entrenamiento. Este fenómeno fue sistemáticamente documentado en las pruebas, estableciendo una relación directa entre el número de bandas y el aumento en el tiempo de ejecución. Estos resultados se detallan en la **Tabla 4-8**.

Tabla 4-8. Desempeño computacional de *U-Net++_4_4* y *DLV3+_4_3* utilizando diferentes imágenes. Fuente propia

Imagen \ Modelo	<i>U-Net++_4_4</i> (min)	<i>DLV3+_4_3</i> (min)
<i>img_765432</i>	65,24	56,46
<i>img_765432NPE</i>	75,15	64,55
<i>img_765432NPE</i>	80,21	69,36

4.4 Selección del modelo

Los resultados de los modelos *U-Net++_4_4* y *DLV3+_4_3* mostraron métricas generales muy similares. *U-Net++_4_4* alcanzó una *OA* del 68,07% y un *kappa* de 0,6, mientras que *DLV3+_4_3* registro un *OA* del 67,29% y un *kappa* de 0,591. Las diferencias entre

los modelos fueron mínimas, siendo de solo un 0,78% en OA y 0,009 en Kappa, Estos resultados, no permiten seleccionar claramente un modelo sobre el otro, mostrando comportamientos similares en la clasificación de coberturas de páramos. Los detalles se presentan en la **Tabla 4-9**.

Tabla 4-9. Comparación de las métricas de *U-Net++_4_4* y *DLV3+_4_3* usando la imagen *img_765432NPE*. Fuente propia.

Modelo Métricas	<i>U-Net++_4_4</i>	<i>DLV3+_4_3</i>
Pixeles	2124330	2146304
OA	68,07%	67,29%
Kappa	0,6	0,591

Un análisis más detallado de las métrica por coberturas muestra que el modelos *U-Net++_4_4* alcanzo las mejores métricas de *F1* en 5.1. *Aguas continentales* con un 90,17%, 3.2.1. *Herbazal* con un 78,43% y 3.1. *Bosques* con un 75%. Sin embargo, la variabilidad en los valores de *kappa* para estas coberturas fue de 0,6, 0,64 y 0,54 respectivamente, indicando ciertas inconsistencias en la clasificación. Además, este modelo enfrente dificultades al clasificar las coberturas de 2.3. *Vegetación secundaria o en transición*, en las que se obtuvo un *F1* de 43,26% y un *IoU* de 0,22.

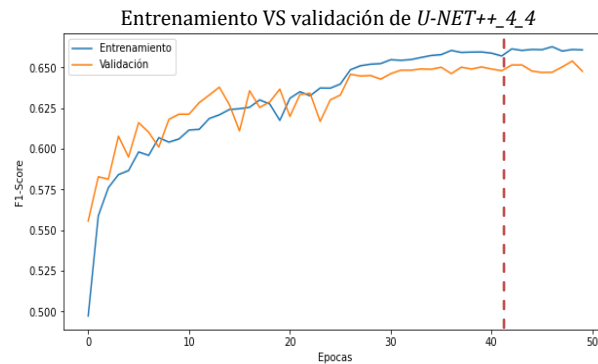
Por otro lado, el modelo *DLV3+_4_3* mostró capacidades similar en la clasificación de las cobertura de 5.1. *Aguas continentales*, 3.2.1. *Herbazal* y 3.1. *Bosques*, con valores de *F1* de 89,07%, 79,22% y 74,27% y *IoU* de 0,56, 0,64 y 0,62 respectivamente. Este modelo también encontró dificultades en clasificar a 2.3. *Vegetación secundaria o en transición*, con un *F1* de 42,08% y un *IoU* de 0,22. Ambos modelos reportaron resultados similares en las coberturas naturales de páramos, 3.2.1. *Herbazal* y 3.1. *Bosques*, como se documenta en la **Tabla 4-10**.

Tabla 4-10. Métricas por coberturas de *U-Net++_4_4* y *DLV3+_4_3* con la imagen *img_765432NPE*. Fuente propia.

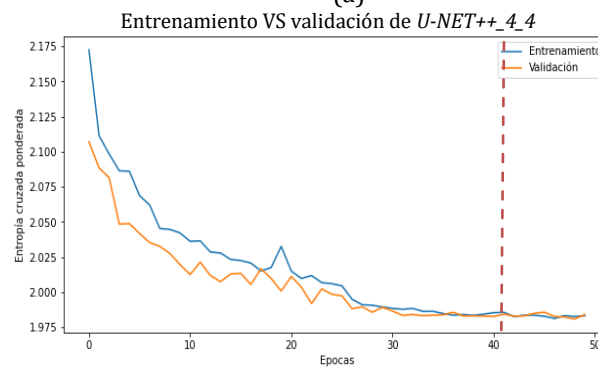
Cobertura	<i>U-Net++_4_4</i>		<i>DLV3+_4_3</i>	
	<i>F1</i>	<i>IoU</i>	<i>F1</i>	<i>IoU</i>
2.3.	72,85%	0,54	73,14%	0,53
2.4.	55,12%	0,44	52,70%	0,44
3.1.	75%	0,63	74,27%	0,61
3.2.1.	78,43%	0,64	79,22%	0,63
3.2.2.	54,24%	0,29	55,72%	0,28

Cobertura	<i>U-Net++_4_4</i>		<i>DLV3+_4_3</i>	
	<i>F1</i>	<i>IoU</i>	<i>F1</i>	<i>IoU</i>
3.2.3.	46,40%	0,23	42,08%	0,22
5.1.	90,17%	0,6	89,07%	0,56
No data	98,22%	0,85	97,43%	0,79

Un aspecto destacable de ambos modelos, *U-Net++_4_4* y *DLV3+_4_3*, es la estabilización del *F1* y de la función de pérdida después de aproximadamente 35 épocas de entrenamiento. La proximidad entre las métricas de entrenamiento y validación, claramente ilustrada en la **Figura 4-7** y **Figura 4-7**, sugiere una implementación efectiva de regularización. Esta es fundamental para minimizar el riesgo de sobreajuste y garantizar una buena generalización de los modelos a nuevos datos, lo cual es crucial para aplicaciones prácticas en la clasificación de coberturas en entornos de páramos.



(a)



(b)

Figura 4-6: Comparación del entrenamiento y la validación (a) Exactitud (b) Entropía cruzada ponderada para *U-Net++_4_4*. Fuente propia

Alcanzar y mantener este nivel de rendimiento constante es crucial para asegurar que los modelos proporcionen resultados consistentes y fiables. Además, la estabilidad observada permite considerar la reducción del número de épocas de entrenamiento de 50 a 35, lo cual podría reducir significativamente los tiempos de entrenamiento sin comprometer la eficiencia de la clasificación.

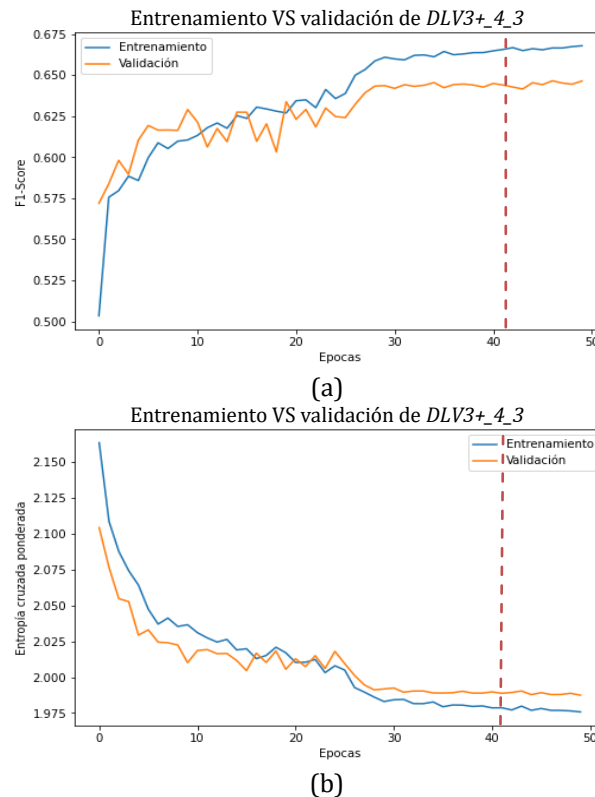


Figura 4-7: Comparación del entrenamiento y la validación (a) Exactitud (b) Entropía cruzada ponderada para *DLV3+_4_3*. Fuente propia

En la **Figura 4-8**, la matriz de confusión normalizada proporciona una visión global del rendimiento de los modelos, permitiendo un análisis detallado de su comportamiento en la clasificación específica de cada cobertura. Las intensidades de color en la diagonal principal reflejan la exactitud de las inferencias para cada cobertura. Un tono más intenso de azul indica una mayor concordancia entre las inferencias y las coberturas reales, mientras que tonos más claros indican áreas de mejora o mayor variabilidad en la clasificación.

Los resultados obtenidos con el modelo *U-Net++_4_4*, demuestran un mejor desempeño en identificar coberturas de gran extensión y homogeneidad en las imágenes, lo que facilita el reconocimiento de coberturas tales como *5.1. Aguas*

continentales, 3.2.1. Herbazal y 3.1. Bosques donde la uniformidad del terreno ayuda a mejorar la exactitud del modelo. Sin embargo, el modelo presenta limitaciones en áreas más pequeñas o donde las características fueron menos definidas en las imágenes utilizadas.

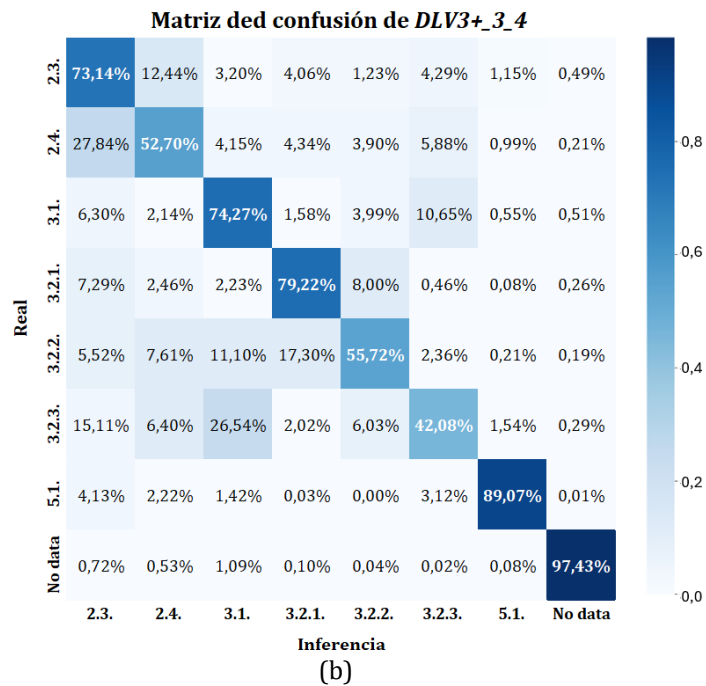
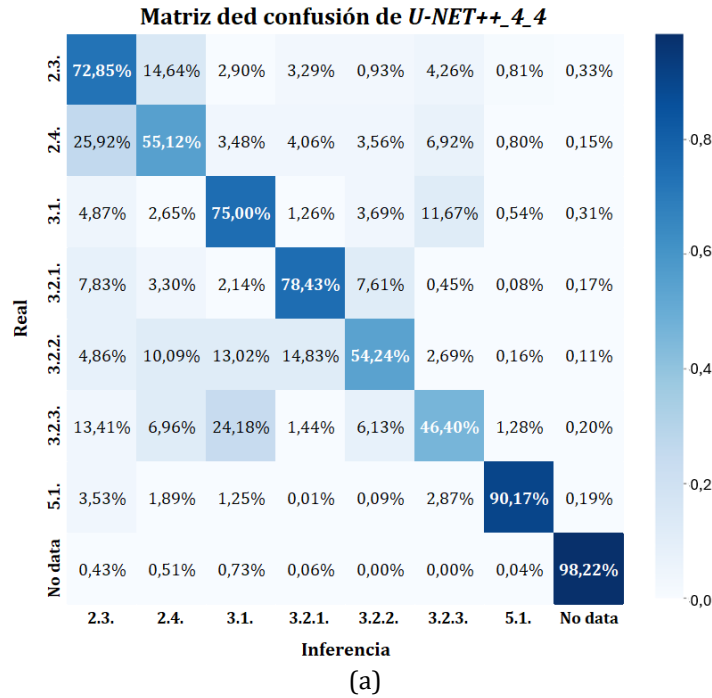
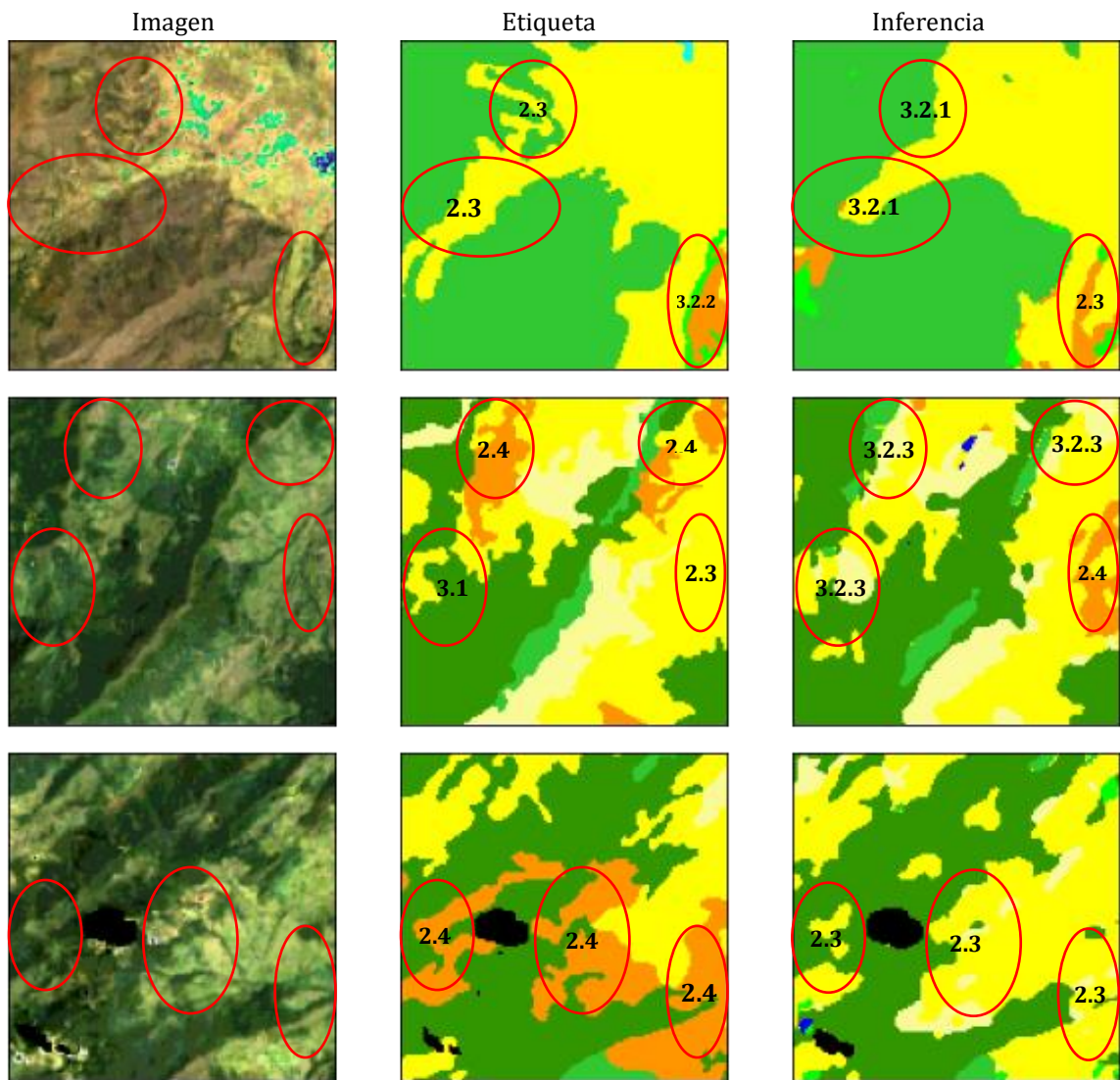


Figura 4-8: Matriz de confusión del modelo (a) U-Net++_4_4 (b) DLV3+_4_3. Fuente propia.

En la **Figura 4-9**, los recuadros rojos destacan las áreas con discrepancias significativas entre las imágenes originales, las etiquetas y las inferencias del modelo, ilustrando los retos en la clasificación final, especialmente en zonas con transiciones complejas entre diferentes tipos de vegetación. Por ejemplo, la cobertura de 2.4. *Áreas agrícolas heterogéneas* que presentan una alta variabilidad, fueron clasificadas incorrectamente en ocasiones como 2.3. *Pastos* o 3.2.3. *Vegetación secundaria o en transición*. Estas inexactitudes pueden estar relacionadas con los cambios visuales causados por las prácticas agrícolas y la similitud de las coberturas en diferentes etapas de desarrollo de los cultivos, que pueden alterar significativamente la apariencia de estas áreas en las imágenes satelitales. En el **Anexo A** se muestra el resultado de la clasificación del área de estudio usando el modelo *U-Net++_4_4*.



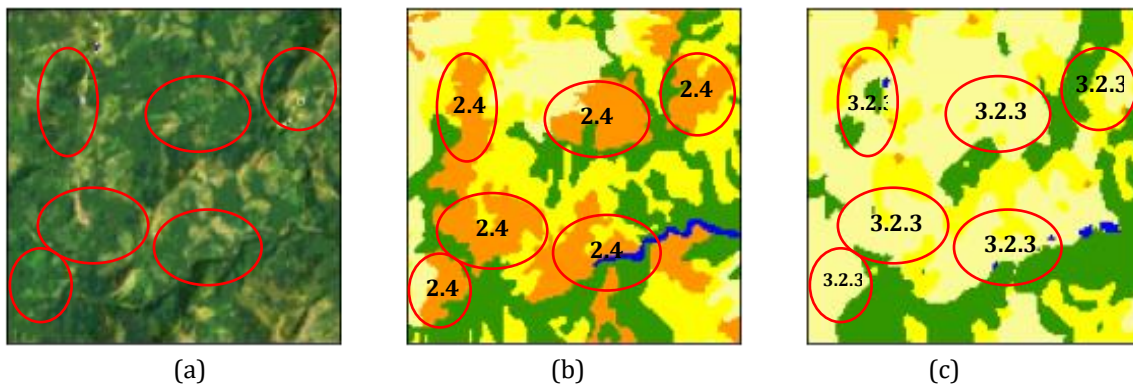


Figura 4-9: Resultados de la clasificación del modelo U-Net++_4_4. (a) Imagen, (b)Etiqueta y (c) Inferencia. Fuente propia.

Los resultados obtenidos con el modelo *DLV3+_4_3*, muestran un mejor desempeño en identificar coberturas de gran extensión y homogeneidad en las imágenes. Esta capacidad del modelo facilita el reconocimiento de las coberturas de *5.1. Aguas continentales*, *3.2.1. Herbazal* y *3.1. Bosques*, donde la uniformidad del terreno contribuye a mejorar la exactitud de la clasificación. Sin embargo, el modelo muestra limitaciones al delimitar coberturas en áreas pequeñas e irregulares, donde los límites visibles en las imágenes no corresponden con las etiquetas. Estos desafíos se ilustran en la **Figura 4-10**, mediante recuadros rojos que resaltan discrepancias significativas entre las imágenes originales, las etiquetas, y las inferencias del modelo.

En particular, el modelo enfrenta dificultades en zonas que exhiben mosaicos de coberturas o transiciones entre diferentes tipos de vegetación. Un ejemplo de la clasificación errónea de las *2.4. Áreas agrícolas heterogéneas* y *2.3. Pastos*, que a menudo fueron clasificados erróneamente como *3.1 Bosques* y *2.3. Pastos*. Estas inexactitudes están relacionadas con los cambios visuales causados por prácticas agrícolas y la similitud de las coberturas en diferentes etapas del desarrollo de los cultivos, lo que puede alterar significativamente la apariencia de estas áreas en las imágenes satelitales. Los resultados detallados de la clasificación del modelo *DLV3+_4_3* se presentan en el **Anexo A**.

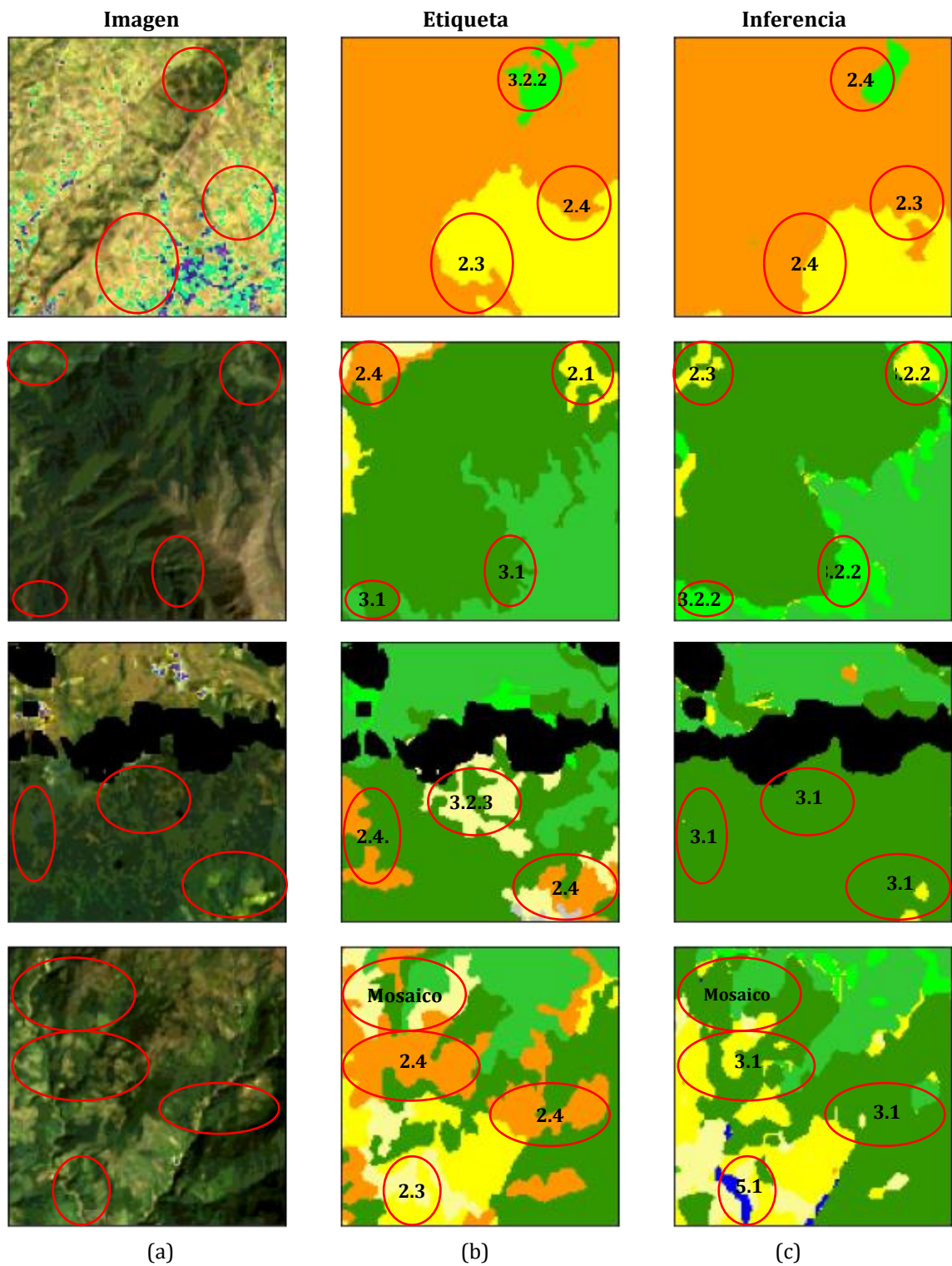


Figura 4-10: Resultados de la clasificación del modelo *DLV3+_4_3*. (a) Imagen, (b) Etiqueta y (c) Inferencia. Fuente propia.

En términos de desempeño computacional, se observa una diferencia en los tiempos requeridos para completar el entrenamiento e inferencia entre los modelos *U-Net++_4_4* y *DLV3+_4_3*. Específicamente, el modelo *U-Net++_4_4* requiere aproximadamente 80,21 minutos para completar ambas fases, mientras que el modelo *DLV3+_4_3* requiere solo 69,36 minutos. Esta diferencia representa un 13,53% menos de tiempo para *DLV3+_4_3*. Los detalles los tiempos de entrenamiento e inferencia se encuentran en la **Tabla 4-11** **Tabla 4-9**.

Tabla 4-11. Desempeño computacional de *U-Net++_4_4* y *DLV3+_4_3*. Fuente propia.

Modelo	Tiempo (Minutos)
<i>U-Net++_4_4</i>	80,21
<i>DLV3+_4_3</i>	69,36

En los experimentos realizados, se encontró que el tiempo necesario para procesar grandes volúmenes de datos no aumenta de manera lineal conforme crece el tamaño de los datos. A pesar de mantener constantes el modelo, el hardware y el tamaño de los lotes, se observó que la eficiencia en la ejecución del modelo y la convergencia de los algoritmos pueden variar considerablemente con un incremento en la cantidad de datos. Debido a la falta de datos precisos sobre las áreas exactas de los contornos y los páramos, se optó por estimar el área continental de Colombia, que abarca 114.174.800 hectáreas según datos del IGAC. Basándose en los tiempos registrados durante los experimentos en el páramo TBM, se realizó una extrapolación que sugiere que el modelo *U-Net++_4_4* requeriría aproximadamente 79.70 horas para la clasificación completa, mientras que el modelo *DLV3+_4_3* necesitaría alrededor de 68.92 horas. Estas estimaciones proporcionan una estimación aproximada del tiempo necesario al considerar áreas más extensas.

Tabla 4-12. Comparación del desempeño computacional entre el área de estudio y Colombia. Fuente propia.

	Área (ha)	<i>U-Net++_4_4</i> (h)	<i>DLV3+_4_3</i> (h)
Área de estudio	1.914.973,76	1,34	1,16
Colombia	114.174.800	79,70	68,92

5. DISCUSIÓN

5.1 Preguntas de investigación

La investigación se centró en determinar **¿Cuál de los modelos de segmentación semántica de imágenes satelitales, *U-Net++* o *DeepLabV3+*, demuestra el mejor rendimiento en la clasificación de cobertura de la tierra en páramos utilizando imágenes satelitales?**

La selección y optimización de hiperparámetros resultaron ser factores decisivos para el rendimiento de ambos modelos. Variaciones en el tipo de codificador y en las dimensiones de las imágenes tuvieron un impacto significativo en la capacidad de generalización. Reducir el tamaño del lote incrementó el tiempo de entrenamiento, pero facilitó una mejor generalización, estableciendo un marco de referencia para continuar configurando los hiperparámetros.

En la evaluación de la capacidad de clasificación de coberturas del Páramo TBM, se utilizaron indicadores claves como el *OA* y *kappa* como métricas generales y el *F1* y *IoU* para las diferentes coberturas. Estos indicadores proporcionan una medida cuantitativa de la exactitud y consistencia de la clasificación (Foody, 2001). Ambos modelos, *U-Net++4_4* y *DLV3+_4_3* mostraron rendimientos similares. *U-Net++* alcanzó un *OA* de 68,07% mientras que *DLV3+_4_3* logró un *OA* de 67,29% indicando que ambos modelos ofrecen una utilidad moderada, pero aún presentan un margen de mejora considerable. Respecto a los valores de *kappa*, *U-Net++4_4* registró un 0,6, y *DLV3+_4_3* un 0,591, mostrando métricas ligeramente más bajas, lo que refleja la necesidad de optimizaciones adicionales para mejorar la exactitud en la clasificación de los modelos.

Los modelos *U-Net++4_4* y *DLV3+_4_3* tuvieron las mejores métricas en la clasificación de coberturas de páramo, como *Herbazal* y *Bosques*. *U-Net++4_4* alcanzó valores de *F1* de 78,43% y 79,22% respectivamente, lo que indica un desempeño óptimo. Estos resultados fueron ligeramente superiores a los de *DLV3+_4_3* que logró un *F1* de 75% y 74,27% para las mismas coberturas.

Sin embargo, ambos modelos enfrentaron desafíos significativos al clasificar coberturas más heterogéneas o mosaicos de vegetación. En 2.4. *Áreas agrícolas heterogéneas* 3.2.3. *Vegetación secundaria o en transición*, *U-Net++4_4* registró un *F1* de 55,12% y 52,70% y un *IoU* de 0,44 en ambas, mientras que *DLV3+_4_3* alcanzó un *F1* de 46,40% y 42,08% con un *IoU* de 0,23 y 0,22 respectivamente, reflejando resultados bajos y poco confiables. Esta tendencia se repitió en la clasificación de 3.2.2. *Arbustos* donde la baja representatividad de la cobertura disminuyó significativamente los resultados de *F1* e *IoU*, lo que resultó en mediciones poco confiables para ambas métricas en ambos modelos.

Los resultados de este estudio acerca del desbalance de clases se alinean con hallazgos de investigaciones previas (Z. Chen et al., 2022; Iyer et al., 2021; Yuan et al., 2021), que también documentaron desafíos similares en la clasificación de coberturas heterogéneas. Esto sugiere que tales limitaciones podrían ser inherentes a la tecnología actual de clasificación y a las características intrínsecas de las imágenes satelitales utilizadas en el estudio. Aunque las imágenes empleadas corresponden al mismo período que las utilizadas por el IDEAM, no se incorporaron fuentes secundarias. Esta divergencia en las fuentes de datos podría influir en la comparabilidad y los resultados de la clasificación, subrayando la necesidad de considerar tanto la calidad como la congruencia de las fuentes de datos en futuras investigaciones.

En términos de eficiencia computacional, *DLV3+_4_3* reportó un menor tiempo, reduciendo el tiempo de entrenamiento e inferencia en un 13,53% en comparación con *U-Net++4_4*. Sin embargo, esta ventaja, aunque beneficiosa en escenarios con un volumen de datos más grande, no resulta significativa para el área de estudio específica de esta investigación. Basándonos en estos resultados, ambos modelos presentaron resultados similares tanto en exactitud temática como en tiempo de ejecución. Esta similitud en los desempeños implica que la selección de uno sobre el otro no puede justificarse claramente basándose únicamente en los datos actuales.

5.1.1 ¿Cómo abordar el desbalance de las diferentes clases de coberturas de páramos presentes en el área de estudio?

El desbalance de clases en la segmentación semántica usando *CNN* ha demostrado tener un impacto adverso en las inferencias, como en las coberturas del Páramo TBM, donde la distribución de coberturas es heterogénea. Esta región presenta áreas irregulares con múltiples coberturas derivadas tanto de ecosistemas naturales de páramos como en áreas de transición derivadas de actividades antropogénicas,

incluyendo la agricultura y la ganadería. La **Figura 5-1** muestra la representatividad de cada clase, evidenciando la disparidad en la representación de clases y los retos asociados al desbalance.

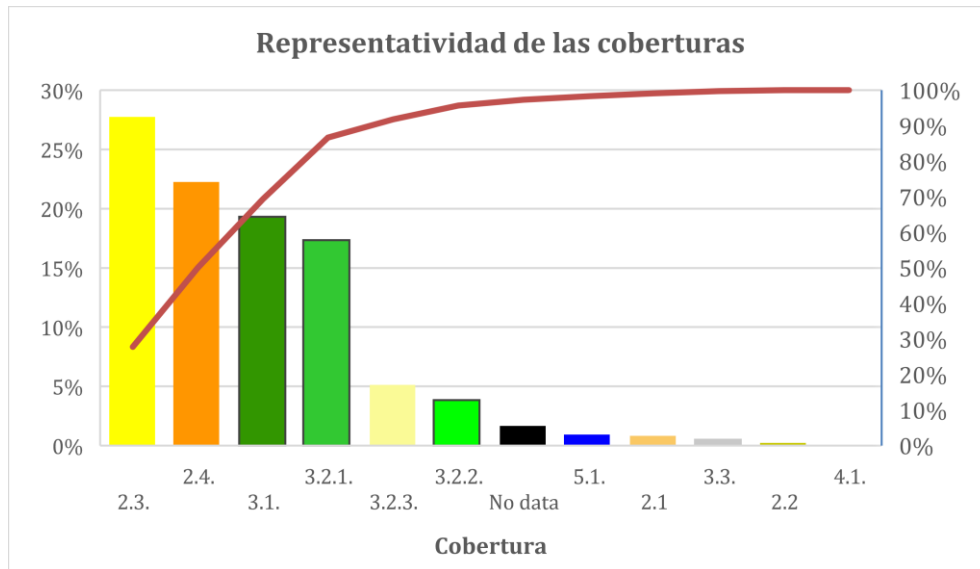


Figura 5-1: Representatividad de las coberturas en el área de estudio. Fuente propia.

Para abordar el desbalance de clases en la clasificación de coberturas de páramos, se implementaron tres enfoques estratégicos, cada uno contribuyendo de manera diferente a la mejora de la representatividad de las clases menos frecuentes. Los experimentos iniciales demostraron que la expansión del área de estudio y el aumento del volumen de datos fueron eficientes para incrementar el número de muestras disponibles de todas las coberturas. Sin embargo, la estrategia que más impacto tuvo en la mitigación del desbalance fue la asignación de pesos diferenciados por cobertura dentro del modelo.

La **Figura 5-2** muestra cómo la representatividad de las coberturas influye en la exactitud temática de los modelos *U-Net++4_4* y *DLV3+_4_3* al aplicar pesos diferenciados a cada clase. Se puede observar que, incluso con pesos ajustados, las clases menos representadas no aumenta significativamente la exactitud dado que no se proporcionan suficientes muestras para realizar inferencias efectivas. Asimismo, se destaca que una alta representatividad no siempre conduce a una mayor exactitud. Esto se observa en coberturas como *2.3. Pastos* y *2.4. Áreas agrícolas heterogéneas*.

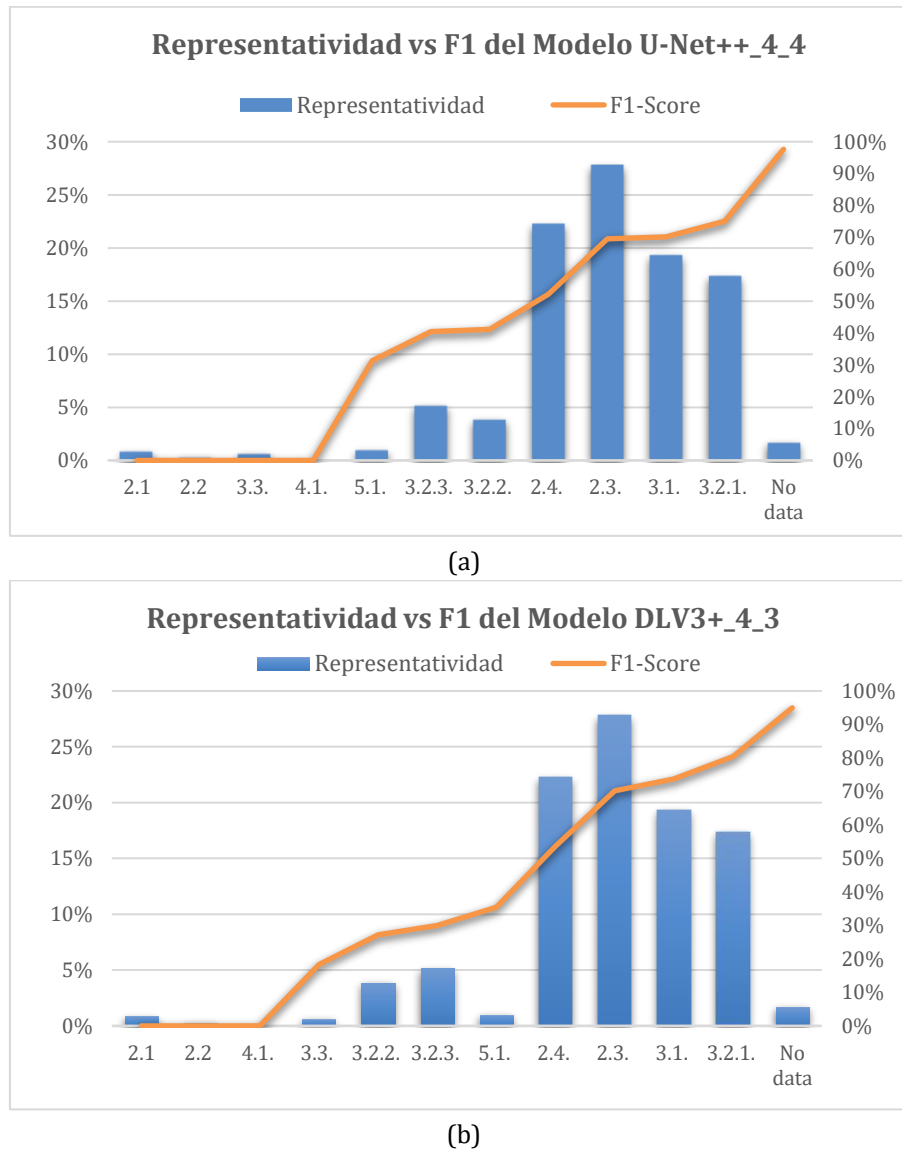


Figura 5-2: Representatividad de las coberturas vs F1 de los resultados. (a) Modelo *U-Net++_4_4* y (b) Modelo *DLV3+_4_3*. Fuente propia.

5.1.2 ¿Cuál es el impacto de incrementar el número de bandas utilizadas como variables explicativas en la exactitud temática de los modelos?

El aumento en el número de bandas espectrales utilizadas en los modelos *U-Net++* y *DeepLabv3+* resultó en una mejora notable en su capacidad de inferencia, reflejada en un incremento de aproximadamente el 5% en la métrica *F1* tras añadir más bandas de *Landsat 8*. Este resultado enfatiza la importancia de utilizar un espectro más amplio de información espectral para la segmentación de imágenes satelitales.

La (b)

Figura 5-3 ilustra mediante un análisis comparativo cómo el incremento en el número de bandas mejora la exactitud de los modelos. Aunque variables adicionales como la pendiente, la elevación y el *NDVI* contribuyen a la exactitud, el impacto de las bandas espectrales es más significativo. Este hallazgo está en consonancia con investigaciones previas, como la de Castelo-Cabay et al., (2022), subrayando que la inclusión de más bandas espectrales es crucial debido a su capacidad para capturar una gama más variada de información sobre las características del terreno.

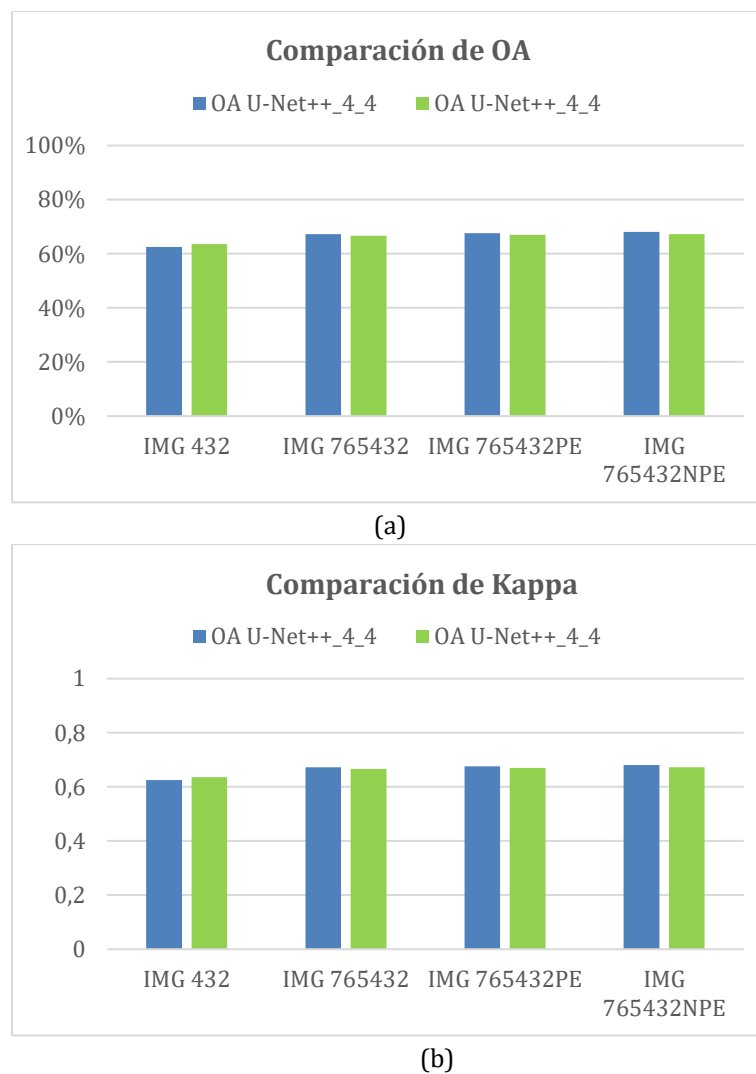
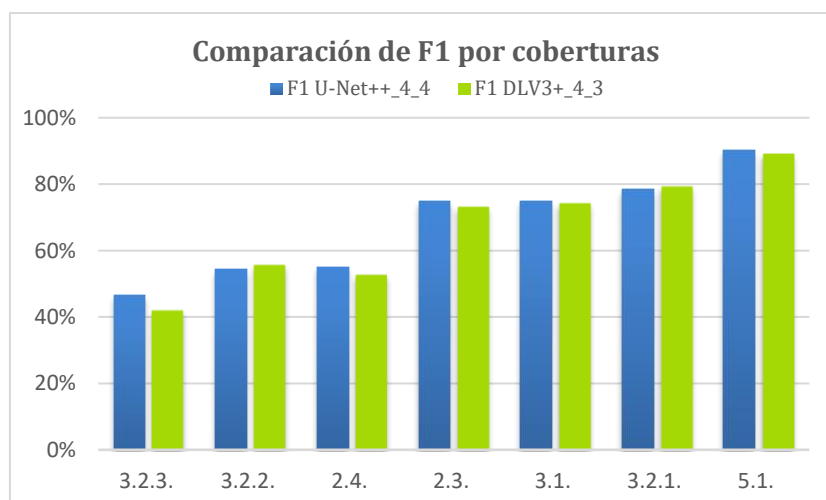


Figura 5-3: Comparación de las métricas generales de (a) *OA* y (b) *kappa* de *U-Net++_4_4* y *DLV3+_4_3*. Fuente propia.

5.1.3 ¿Cuáles coberturas de páramos muestran una segmentación de mayor exactitud temática mediante el uso de los modelos U-Net++ y DeepLabv3+?

Los modelos *U-Net++* y *DeepLabV3+* han mostrado un mejor desempeño en las coberturas de páramo, específicamente en *3.1 Bosques* y *3.2.1 Herbazal*, donde la alta representatividad de estas coberturas fue un factor determinante en la exactitud temática de la clasificación. Sin embargo, la cobertura *3.2.2 Arbustal*, a pesar de ser propia de los páramos, no logró buenos resultados debido a la escasa presencia en el área de estudio, lo que impactó negativamente en su inferencia. En cuanto a otras coberturas, como *5.1 Aguas continentales* destacó por su alta exactitud, mientras que *3.3. Áreas abiertas, sin o con poca vegetación*, *2.4. Áreas agrícolas heterogéneas* mostraron un desempeño más bajo, subrayando el desafío de clasificar coberturas con mayor heterogeneidad y menor representatividad.

La **Figura 5-4** presenta una comparación de las métricas *F1* y *IoU* para cada cobertura, evidenciando diferencias significativas entre ellas. Altos valores de *F1* indican buena exactitud y recuperación eficiente de las clases, mientras que los valores *IoU* sugieren variabilidad en la consistencia de estas clasificaciones a través de diferentes conjuntos de datos. Este contraste enfatiza la importancia de considerar ambas métricas para una evaluación integral de la calidad de la segmentación, especialmente en coberturas como *5.1 Aguas continentales*, *3.2.3. Vegetación secundaria o en transición* y *3.2.2. Arbustos*, donde se observan marcadas diferencias. Por otro lado, las coberturas de páramos *3.1. Bosques* y *3.2.1. Herbazal* muestran las menores diferencias.



(a)

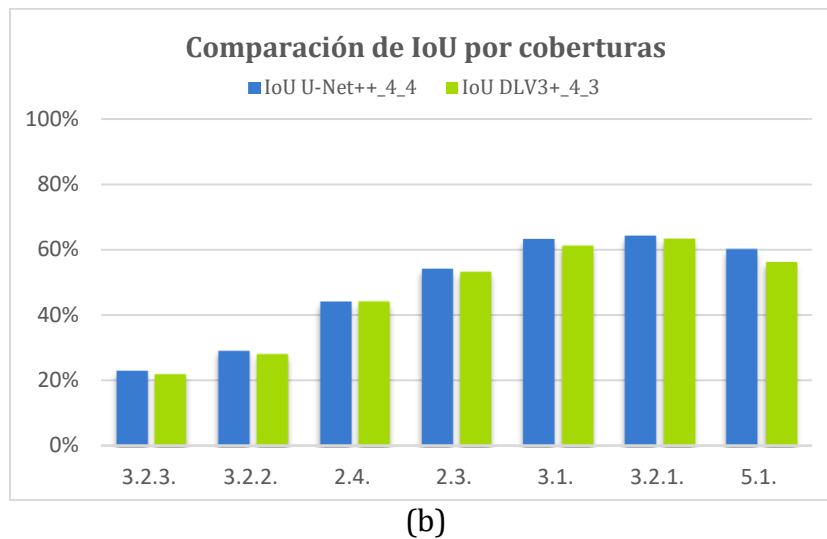


Figura 5-4: Comparación de las métricas por coberturas de (a) $F1$ y IoU de $U-Net++_4_4$ y de (b) $F1$ y IoU de $DLV3+_4_3$. Fuente propia.

5.2 Aportes a la geomática

Esta investigación contribuye significativamente a la geomática mediante la implementación y validación de metodologías avanzadas en CNN para la clasificación de cobertura en páramos, utilizando la segmentación semántica de imágenes satelitales. Se destacan contribuciones fundamentales en la aplicación de modelos avanzados como $U-Net++$ y $DeepLabV3+$ en entornos complejos, el manejo del desbalance de clases, y la contribución al progreso del aprendizaje profundo en Colombia enfocado en páramos. Los principales aportes se detallan a continuación:

- **Aplicación efectiva de modelos avanzadas:** Basándose en estudios previos, como Yuan et al., (2021), esta investigación establece una línea base fundamental para la aplicación de técnicas avanzadas en áreas de páramo, caracterizadas por una limitada disponibilidad de datos actualizados. La aplicación de modelos como $U-Net++$ y $DeepLabV3+$, ha facilitado una clasificación más eficiente en tiempo para las coberturas de páramos.
- **Abordaje de desafíos en las CNN :** esta investigación implementó diferentes estrategias para mitigar el desbalance de clases y la insuficiencia de datos etiquetados, como la expansión del área de estudio y la aplicación de técnicas de aumento de datos, de acuerdo con los retos estudiados por Yuan et al., (2021). Estas medidas han mejorado el tratamiento del desbalance de clases y han aumentado la representatividad de las coberturas menos representadas en el

estudio, optimizando así la capacidad de las *CNN* para clasificar las coberturas en entornos de páramo

- **Contribución al aprendizaje profundo en Colombia:** Esta investigación destaca la implementación del aprendizaje profundo en Colombia está en su fase inicial, particularmente en la aplicación de técnicas digitales. Estudios previos, como el de Suárez Londoño et al., (2017), estableció una base sobre la cual se pueden construir futuros estudios. Este investigación promueve el desarrollo del aprendizaje profundo en Colombia, impulsando la creación de metodologías que pueden ser adaptadas y optimizadas para abordar retos geográficos y ambientales locales, enriqueciendo así la capacidad del país en geomática y teledetección.
- **Comparación con otros estudios en Páramos:** A diferencia del estudio de Castelo-Cabay et al., (2022) que investigó páramos ecuatorianos empleando agrupaciones generales de coberturas (Nivel I), esta investigación se centra en una clasificación más detallada (Niveles II y III) de las coberturas naturales de páramos. Además, se emplearon modelos específicos de *CNN* como *U-Net++* y *DeepLabV3+*. Se utilizaron datos de Landsat 8 correspondientes a los años 2018 y 2019, proporcionando una comparativa con las imágenes Sentinel-2 del año 2020 usadas en el estudio ecuatoriano. Este enfoque resalta la importancia de utilizar conjuntos de datos completos y etiquetas detalladas, fundamentales para mejorar la precisión en la clasificación de coberturas en áreas de páramo. Esto es crucial para una gestión y conservación efectivas de estos ecosistemas críticos.
- **Investigaciones Futuras:** Los resultados obtenidos destacan la necesidad de adoptar técnicas avanzadas de aprendizaje profundo para mejorar los tiempos de clasificación de coberturas de tierra en los páramos. Técnicas como las *redes adversarias generativas (GANs)*, estrategias de *transferencia de aprendizaje*, *enmascaramiento de las clases complejas* o *ensamble de modelos* muestran un gran potencial para potenciar la generalización de nuevos modelos para una amplia variedad de coberturas. La exploración y validación de estos modelos avanzados en todos los páramos de Colombia constituye una dirección prometedora para investigaciones futuras. Esta aproximación no solo verificará la generalización y adaptabilidad de los modelos a diversos ecosistemas de páramo, sino que también evidenciará su aplicabilidad en programas nacionales de monitoreo y conservación ambiental.

5.3 Ventajas y limitaciones de esta investigación

5.3.1 Ventajas:

Las ventajas de esta investigación fueron:

- **Mejora en modelos de segmentación semántica:** La investigación destacó la necesidad de optimizar los modelos de segmentación semántica de imágenes satelitales utilizando *CNN*. Los modelos *U-Net++* y *DeepLabV3+* se ajustaron significativamente para enfrentar desafíos en la clasificación de coberturas menos representadas, como las coberturas de *3.1 Arbustos* y *5.1 Aguas continentales*. A través de ajustes en los modelos, se logró detectar y clasificar estas coberturas, evidenciado particularmente un aumento en *5.1 Aguas continentales*, obteniendo un *F1*, de 90.26%. Este avance demuestra la capacidad de los modelos para interpretar y clasificar con exactitud imágenes satelitales complejas.
- **Clasificación de coberturas de todos los páramos:** La implementación de técnicas avanzadas de clasificación digital para todos los páramos de Colombia representa un enfoque estratégico para la adquisición de datos actualizados y detallados en tiempos más cortos. La utilización de modelos como *U-Net++* y *DeepLabV3+* a gran escala permitiría tener una mayor representatividad de las clases, enfoque que facilita el aprendizaje del modelo sobre las características específicas de diversas coberturas, proporcionando muestras más representativas y mejorando la exactitud de las clasificaciones.

5.3.2 Limitaciones:

Las limitaciones de esta investigación fueron:

- **Ausencia de etiquetas actualizadas y detalladas:** Una de las principales limitaciones enfrentadas en este estudio es la falta de etiquetas actualizadas y detalladas para el páramo TBM. La última actualización de etiquetas disponibles es de 2014, a pesar de ser la de mayor detalle, no se pudo utilizar debido a la ausencia de clasificación de las áreas limítrofes del páramo dado que se requiere la información de las áreas con coberturas en transición. Esta carencia impactó significativamente la capacidad del modelo para aprender y generalizar con una mayor exactitud temática.

- **Dependencia de imágenes y coberturas:** En el proceso de clasificación de coberturas mediante el uso de *CNN*, es crucial contar con imágenes y etiquetas correspondientes para que el modelo pueda aprender correctamente a asociar características espectrales con categorías específicas de cobertura. Las coberturas del IDEAM se basaron principalmente en imágenes de *Landsat 8*, aunque también se integraron otras fuentes como Sentinel-2 y Google Maps. Sin embargo, la falta de detalle suficiente para reconstruir exactamente las imágenes llevó a utilizar exclusivamente imágenes de Landsat 8 en esta investigación. Durante una prueba sustituyendo las imágenes de Landsat 8 por imágenes de Sentinel-2, se observaron resultados inferiores tanto en la exactitud general como en la clasificación de coberturas específicas, debido a discrepancias con las etiquetas establecidas
- **Estrategias alternativas para el desbalance de Clases:** A pesar de implementar técnicas efectivas como la ponderación por clase y el aumento de datos para abordar el desbalance de clases en la clasificación de coberturas de páramos, esta investigación no exploró opciones como la generación de datos sintéticos. Lo que abre una oportunidad para futuras investigaciones, enfocadas en examinar y validar otras estrategias que puedan mejorar aún más el rendimiento del modelo.

6. CONCLUSIONES

Los modelos *DLV3+_4_3* y *U-Net++_4_4*, demostraron resultados similares en términos de *OA* y *Kappa*, sin que se evidencie una diferencia significativa que favorezca a uno modelo sobre el otro. Aunque *DLV3+_4_3* fue un 13.53% más eficiente en tiempo de procesamiento, esta ventaja no justifica una superioridad clara. Por lo tanto, la elección entre estos modelos debe considerar no solo la eficiencia computacional, sino también la exactitud, capacidad de generalización y aplicabilidad, siendo ambos igualmente viables para estudios de coberturas en páramo.

La selección y optimización de hiperparámetros han demostrado ser decisivos para el rendimiento de los modelos *U-Net++* y *DeepLabV3+*. El ajuste de variables como el tipo de codificador y el tamaño de las imágenes impacta significativamente en la capacidad de generalización de los modelos. Esto subraya la necesidad de una calibración meticulosa de los hiperparámetros para mejorar la exactitud y eficiencia de la clasificación de coberturas en entornos complejos como los páramos.

El desbalance de clases es uno de los retos más significativo para la clasificación efectiva de coberturas de páramo utilizando *CNN*. A pesar de la implementación de varias estrategias, las clases con menor presencia no se pudieron clasificar correctamente. Esto genera la necesidad de seguir investigando enfoques más sofisticados que permitan abordar este problema.

El incremento en el número de bandas espectrales utilizadas como variables explicativas mejoró la *OA* de los modelos *U-Net++* y *DeepLabV3+*. Específicamente, el uso de bandas espectrales del visible de Landsat 8 demostraron tener el mayor impacto positivo en el desempeño de los modelos. Este hallazgo destaca la relevancia de emplear un espectro más amplio de información espectral, lo cual es crucial para capturar las características detalladas de la superficie terrestre en la segmentación de imágenes satelitales.

Los modelos seleccionados demostraron un rendimiento óptimo en coberturas de páramo como *Bosques* y *Herbazal*, dada su mayor representatividad y homogeneidad en el área de estudio. Sin embargo, los modelos enfrentaron desafíos significativos al

clasificar los *Arbustales* una cobertura caracterizada por su heterogeneidad y baja representatividad en el páramo TBM. Aunque se observaron buenos resultados en *Aguas continentales*, los modelos no lograron un rendimiento satisfactorio con *Vegetación secundaria o en transición y Áreas agrícolas heterogéneas*. Este patrón de resultados subraya la necesidad de desarrollar estrategias adicionales, como el enmascaramiento de clases complejas o el ensamble de modelos, que permitan una adaptación efectiva a la complejidad y diversidad de las coberturas de páramo.

La implementación de técnicas avanzadas de *CNN* en la clasificación de coberturas de páramos marca un avance significativo en el campo de la geomática. Esta investigación ha abordado desafíos técnicos significativos mediante la introducción de métodos innovadores que mejoran la eficiencia en el monitoreo y gestión de ecosistemas frágiles.

Este estudio ha marcado un progreso en la aplicación del aprendizaje profundo en Colombia, destacando la importancia de utilizar imágenes ópticas multiespectrales y datos etiquetados detallados para una clasificación efectiva de coberturas terrestres, sentando así las bases para futuras investigaciones. Los resultados obtenidos en esta tesis establecen un punto de partida para investigaciones adicionales en el aprendizaje profundo aplicado a la clasificación de coberturas terrestres, con un enfoque específico en el desarrollo y optimización de modelos que manejen el desbalance de clases, un desafío recurrente en estudios ambientales.

Adicionalmente, se recomienda la expansión de la aplicación de estos modelos a todos los páramos de Colombia para evaluar su adaptabilidad y eficiencia en diversos contextos ambientales. Estas áreas de investigación son cruciales no solo para avanzar en la ciencia de datos aplicada al medio ambiente, sino también para contribuir de manera significativa al manejo y conservación de ecosistemas claves en Colombia.

A. Anexo: Clasificación de las áreas de estudio.

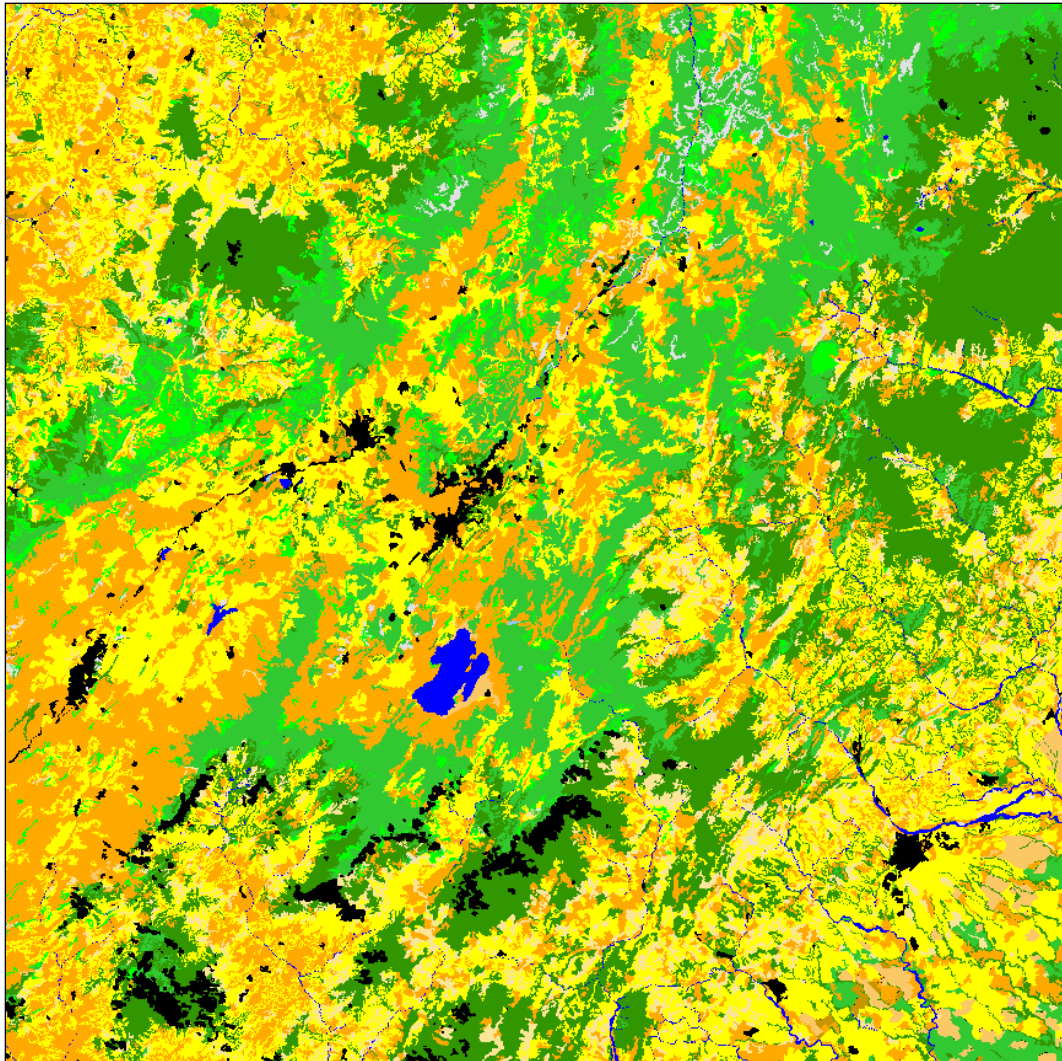


Figura 6-1: Etiquetas del área de estudio con los colores *RGB* para el entrenamiento. Adapta del IDEAM et al., (2021)

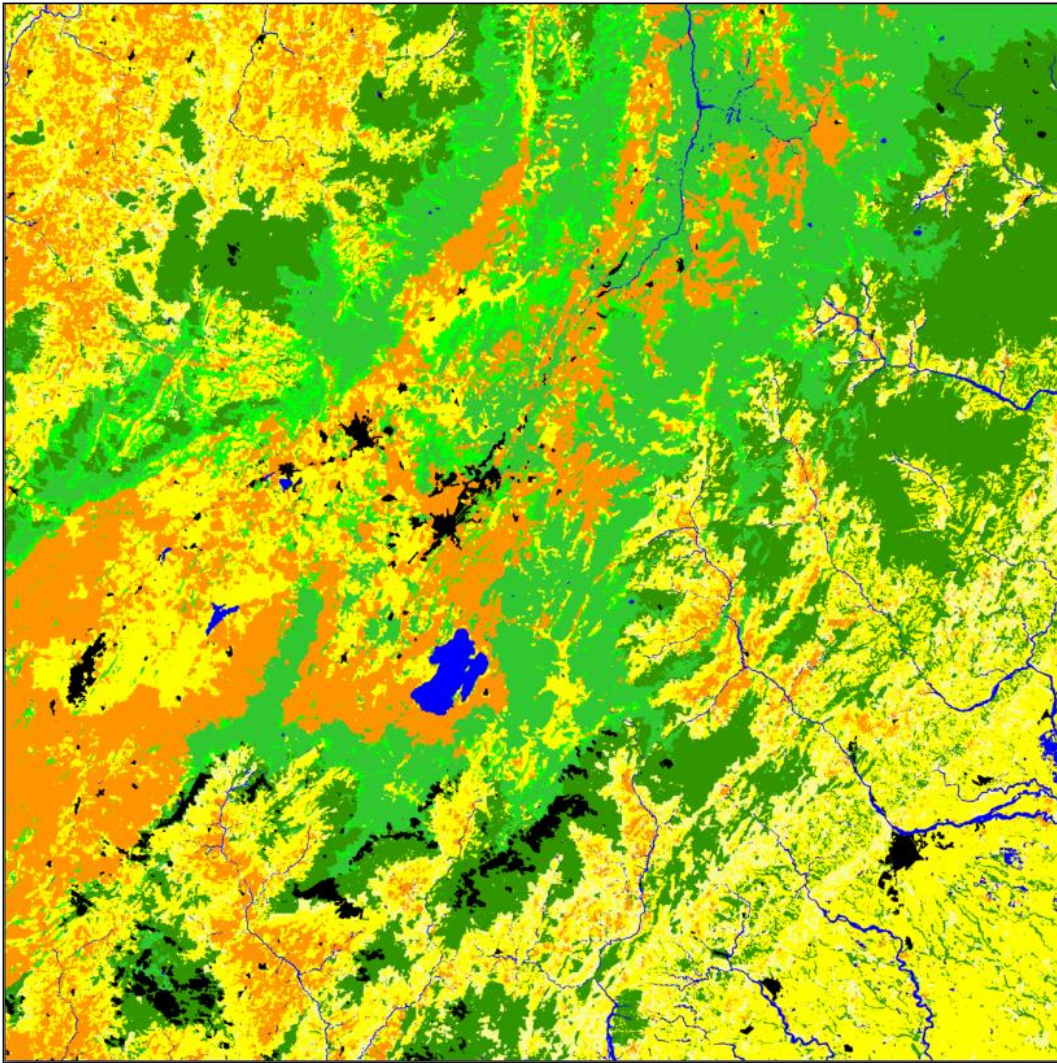


Figura 6-2: Resultados de la clasificación del modelo *U-Net++4_4*. Fuente propia.

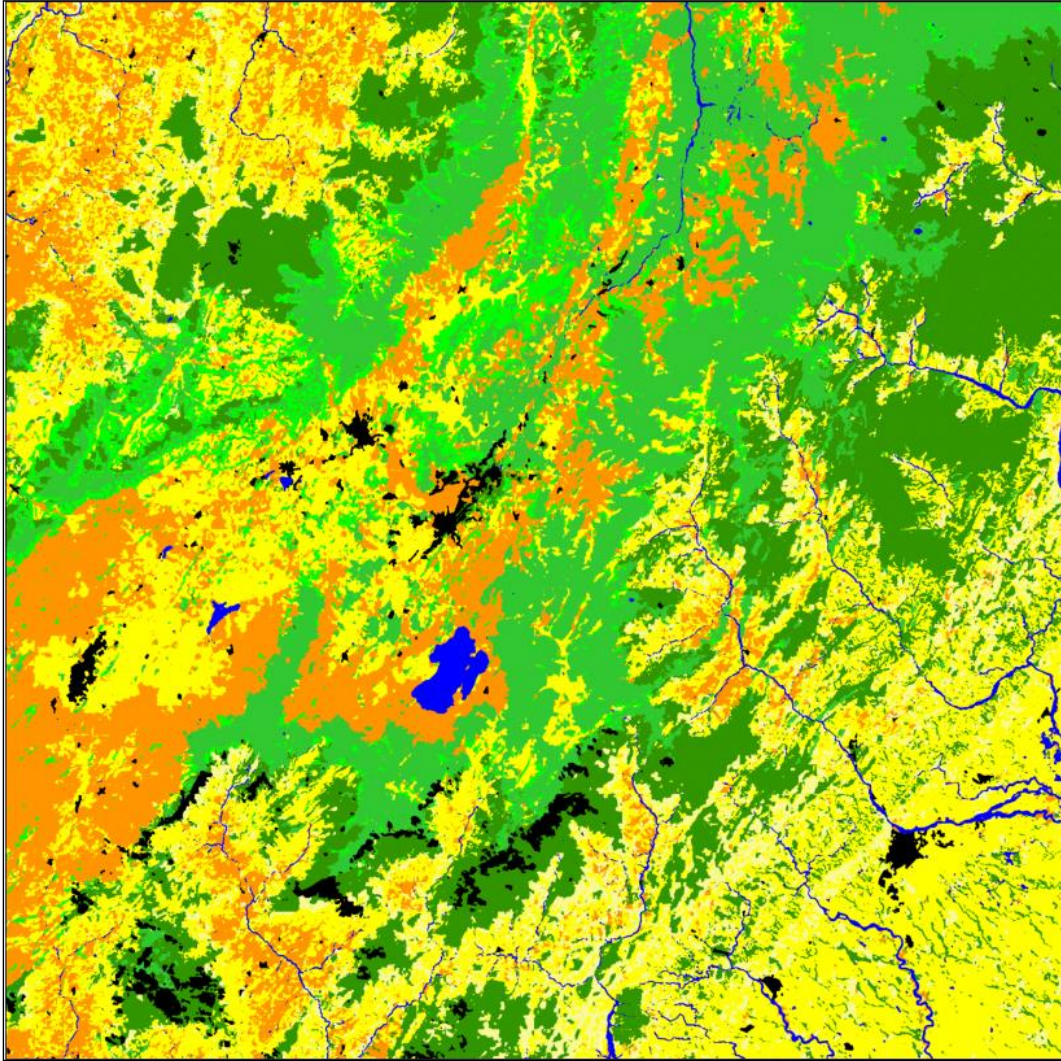


Figura 6-3: Resultados de la clasificación del modelo *DLV3+4_3*. Fuente propia.

B. Anexo: Glosario

- **Dropout:** Este término se refiere a “apagar” algunas neuronas de forma aleatoria durante el entrenamiento de un modelo para evitar que la red memorice los datos de ejemplo y no aprenda de manera general.
- **Épocas:** Imagina que estás leyendo un libro para aprender algo. Cada vez que lees todo el libro de principio a fin es una época. En el aprendizaje de máquina, significa *pasar* por todos los datos de entrenamiento varias veces.
- **Función de Activación:** Es como una regla que decide cuánto debe activarse una neurona. Piensa en ello como un filtro que transforma los datos de entrada de una neurona en una salida más útil.
- **GPU:** Son como super calculadoras especializadas que ayudan a realizar muchos cálculos a la vez, muy útiles para tareas como el procesamiento de imágenes o el entrenamiento de modelos de inteligencia artificial.
- **Tasa de Aprendizaje (*Learning Rate*):** Es como ajustar qué tan grandes son los pasos que da el modelo al aprender. Un paso muy pequeño puede hacer que el aprendizaje sea lento, y uno muy grande puede hacer que se salte la mejor solución.
- **Sobreajuste (*Overfitting*):** Sucede cuando un modelo aprende los detalles y ruido de los datos de entrenamiento tan bien que no funciona correctamente con datos nuevos. Es como estudiar tanto un tema que luego no puedes aplicar bien ese conocimiento a situaciones reales.
- **Padding:** Es como agregar bordes extra alrededor de una imagen antes de procesarla. Esto se hace para mantener el tamaño original de la imagen después de aplicar filtros.
- **Stride:** Piensa en esto como el tamaño del paso que da un filtro cuando se mueve a través de la imagen. Un paso más grande significa que el filtro cubre la imagen más rápidamente.
- **Número de Canales:** En el despliegue de imágenes por computador, cada banda se asigna a alguno de los colores primarios como rojo, verde y azul (*RGB*).

- **Número de Clases:** Es la cantidad de diferentes categorías o tipos de objetos en los que puedes clasificar algo. Por ejemplo, en un modelo que reconoce frutas, las clases podrían ser manzana, banana, naranja, etc.
- **Tamaño de Imágenes:** Se refiere al ancho (número de columnas) y alto (número de filas) de una imagen en píxeles. El tamaño representa cuánta información tiene la imagen y cuánto trabajo necesita el modelo para procesarla.
- **Tamaño del Lote (*Batch Size*):** Indica cuántas muestras de datos se procesan juntas en un paso durante el entrenamiento de un modelo basado en una red convolucional. Un lote más grande puede hacer que el entrenamiento sea más rápido, pero necesita más memoria del computador.

7. BIBLIOGRAFÍA

- Altamirano P, L. (2021). *Deep Learning aplicado a la Segmentación Semántica de Imágenes aéreas*. Universidad de Valparaíso, Chile.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1). <https://doi.org/10.1186/s40537-021-00444-8>
- Arcos Serrano, R. (2021). *Segmentación objeto-fondo mediante redes convolucionales* [Universidad Autónoma De Madrid]. https://calidad.uniovi.es/c/document_library/get_file?p_l_id=2535716&folderId=4003749&name=DLFE-54735.pdf
- Audebert, N., Le Saux, B., & Lefèvre, S. (2017). Semantic segmentation of earth observation data using multimodal and multi-scale deep networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10111 LNCS, 180–196. https://doi.org/10.1007/978-3-319-54181-5_12
- Ayala-Izurieta, J. E., Márquez, C. O., García, V. J., Recalde-Moreno, C. G., Rodríguez-Llerena, M. V., & Damián-Carrión, D. A. (2017). Land cover classification in an ecuadorian mountain geosystem using a random forest classifier, spectral vegetation indices, and ancillary geographic data. *Geosciences (Switzerland)*, 7(2). <https://doi.org/10.3390/geosciences7020034>
- Ayhan, B., Kwan, C., Budavari, B., Kwan, L., Lu, Y., Perez, D., Li, J., Skarlatos, D., & Vlachos, M. (2020). Vegetation detection using deep learning and conventional methods. *Remote Sensing*, 12(15). <https://doi.org/10.3390/RS12152502>
- Belyadi, H., & Haghghat, A. (2021). Model evaluation. In *Machine Learning Guide for Oil and Gas Using Python*. <https://doi.org/10.1016/b978-0-12-821929-4.00009-3>
- Bocarejo, D., Cairo, C. Del, Ojeda, D., Montenegro, I. E., Instituto Alexander von Humboldt, & ERIGAIE, F. (2014). Caracterización socioeconómica y cultural del Complejo de Páramos Tota-Bijagual-Mamapacha en jurisdicción de Corpoboyacá con énfasis en caracterización de actores, análisis de redes y de servicios ecosistémicos. CONTRATO 15-13-014-041PS.

- Repositorio Institucional de Documentación Científica Humboldt Digital*, 014(15), 93. <http://repository.humboldt.org.co/handle/20.500.11761/9578>
- Brück, S. A., Torres, B. D. M., & de Moraes Polizeli, M. de L. T. (2023). The Ecuadorian paramo in danger: What we know and what might be learned from northern wetlands. In *Global Ecology and Conservation* (Vol. 47). Elsevier B.V. <https://doi.org/10.1016/j.gecco.2023.e02639>
- Cabrera, M., & Ramírez, W. (2014). Restauración Ecológica de los páramos de Colombia: Transformación y herramientas para su conservación. In *Restauración Ecológica de los páramos de Colombia: Transformación y herramientas para su conservación*.
- Castelo-Cabay, M., Piedra-Fernandez, J. A., & Ayala, R. (2022). Deep learning for land use and land cover classification from the Ecuadorian Paramo. *International Journal of Digital Earth*, 15(1), 1001–1017. <https://doi.org/10.1080/17538947.2022.2088872>
- Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834–848. <https://doi.org/10.1109/TPAMI.2017.2699184>
- Chen, L.-C., Papandreou, G., Schroff, F., & Adam, H. (2017). *Rethinking Atrous Convolution for Semantic Image Segmentation*. <http://arxiv.org/abs/1706.05587>
- Chen, Z., Duan, J., Kang, L., & Qiu, G. (2022). Class-Imbalanced Deep Learning via a Class-Balanced Ensemble. *IEEE Transactions on Neural Networks and Learning Systems*, 33(10), 5626–5640. <https://doi.org/10.1109/TNNLS.2021.3071122>
- Coca Castro, A., Zaraza Aguilera, M. A., Benavides Miranda, Y. T., Montilla Montilla, Y. M., Posada Fandiño, H. B., Avedaño-Gomez, A. L., Hernández-Hamon, H. A., Garzón Martínez, S. C., & Franco Prieto, C. A. (2021). Evaluación de algoritmos de clasificación en la plataforma Google Earth Engine para la identificación y detección de cambios de construcciones rurales y periurbanas a partir de imágenes de alta resolución. *Revista de Teledetección*, 58, 71. <https://doi.org/10.4995/raet.2021.15026>
- Davins Jovells, M. (2021). *Deep learning for semantic segmentation of remote sensing imaging*.
- Deepan, P., & Sudha, L. R. (2020). Object Classification of Remote Sensing Image Using Deep Convolutional Neural Network. In *The Cognitive Approach in Cloud Computing and Internet of Things Technologies for Surveillance Tracking Systems* (pp. 107–120). Elsevier. <https://doi.org/10.1016/b978-0-12-816385-6.00008-8>
- Di Gregorio, A. (2016). *Land Cover Classification System (LCCS): Classification Concepts and User Manual version 3*. Environmental and Natural Resources Series, FAO. <https://www.fao.org/3/i5232e/i5232e.pdf>

- Espejo O. (2017). Desarrollo de una metodología para estimación de la deforestación mediante el análisis multitemporal de imágenes multiespectrales en un entorno de análisis basado en objetos geográficos (GEOBIA). *Tesis de Maestría. Universidad Distrital Francisco José de Caldas, Colombia.*
- Fan, Z., Zhan, T., Gao, Z., Li, R., Liu, Y., Zhang, L., Jin, Z., & Xu, S. (2022). Land Cover Classification of Resources Survey Remote Sensing Images Based on Segmentation Model. *IEEE Access, 10*, 56267–56281. <https://doi.org/10.1109/ACCESS.2022.3175978>
- Foody, G. M. (2001). *Status of land cover classification accuracy assessment.* www.elsevier.com/locate/rse
- Garavito, N. (2016). Los páramos en Colombia, un ecosistema en riesgo. *Ingeniare, 19*, 127–136. <http://www.unilibrebaq.edu.co/ojsinvestigacion/index.php/ingeniare/article/view/704>
- Gil, J. (2017). *Briófitos del complejo de páramos Tota-Bijagual-Mamapacha y fitogeografía.* 69.
- Gómez Sánchez, C., Rincón Romero, M., & Sánchez López, R. (2002). Caracterización de los suelos páramos de Colombia: Génesis de una transformación. *Páramos y Ecosistemas Alto Andinos de Colombia En Condiciones HotSpot & Global Climatic Tensor*, 210–333. <https://www.researchgate.net/publication/264416399>
- Hoeser, T., & Kuenzer, C. (2020). Object detection and image segmentation with deep learning on Earth observation data: A review-part I: Evolution and recent trends. In *Remote Sensing* (Vol. 12, Issue 10). MDPI AG. <https://doi.org/10.3390/rs12101667>
- Huang, B., Zhao, B., & Song, Y. (2018). Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sensing of Environment, 214*, 73–86. <https://doi.org/10.1016/j.RSE.2018.04.050>
- Iaar. (2020). *Visión por computadora - Libro online de IAAR.* <https://iaarbook.github.io/vision-por-computadora/%0Ahttps://iaarbook.github.io/vision-por-computadora/#redes-neuronales-convolucionales%0Ahttps://iaarbook.github.io/vision-por-computadora/>
- IDEAM. (2010). Leyenda nacional de coberturas de la tierra. Metodología CORINE Land Cover adaptada para Colombia, escala 1:100.000. In *Area: Vol. TH-62-04-1* (Issue 257).
- IDEAM. (2015). *COBERTURAS NACIONALES.* IDEAM - MONITOREO DE BOSQUES Y RECURSO FORESTAL. <http://www.ideam.gov.co/web/ecosistemas/coberturas-nacionales>
- IDEAM, IAVH, MADS, & MINHACIENDA. (2014). *TIERRA EN LOS PÁRAMOS PRIORIZADOS A ESCALA ADAPTADA PARA COLOMBIA) Contrato de Cooperación No. 13-13-014-093CE*

- Instituto Humboldt - Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM)* (Vol. 000, Issue 13).
- IDEAM, SINCHI, & PNN. (2021). *MAPA NACIONAL DE COBERTURAS DE LA TIERRA, ESCALA 1:100.000, PERIODO 2018 METODOLOGÍA CORINE LAND COVER ADAPTADA PARA COLOMBIA MEMORIA TÉCNICA Y RESULTADOS*.
- Instituto Geográfico Agustín Codazzi (IGAC). (2015). *IGAC culminó estudio de suelos en los páramos colombianos* / Instituto Geográfico Agustín Codazzi. <https://igac.gov.co/es/noticias/igac-culmino-estudio-de-suelos-en-los-paramos-colombianos>
- Iyer, P., A, S., & Lal, S. (2021). Deep learning ensemble method for classification of satellite hyperspectral images. *Remote Sensing Applications: Society and Environment*, 23, 100580. <https://doi.org/10.1016/J.RSASE.2021.100580>
- Kemker, R., Salvaggio, C., & Kanan, C. (2018). Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 60–77. <https://doi.org/10.1016/J.ISPRSJPRS.2018.04.014>
- Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters*, 14(5), 778–782. <https://doi.org/10.1109/LGRS.2017.2681128>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Liu, M., Fu, B., Xie, S., He, H., Lan, F., Li, Y., Lou, P., & Fan, D. (2021). Comparison of multi-source satellite images for classifying marsh vegetation using DeepLabV3 Plus deep learning algorithm. *Ecological Indicators*, 125, 107562. <https://doi.org/10.1016/j.ecolind.2021.107562>
- Loaiza-Usuga, J. C., Lis-Gutiérrez, M., & Rubiano-Sanabria, Y. (2021). Land use and environmental changes in the Andean Paramo soils. In *Climate and Land Use Impacts on Natural and Artificial Systems* (pp. 105–134). Elsevier. <https://doi.org/10.1016/B978-0-12-822184-6.00014-4>
- Machine Learning en la Web con Angular y TensorFlow*. (n.d.). Retrieved October 20, 2023, from <https://saludelectronica.com/machine-learning-en-la-web-con-angular-y-tensorflow/>
- MADS. (2021). *“Tenemos el 50 % de los páramos del mundo”*: Ministro de Ambiente en el Día de las Montañas - Ministerio de Ambiente y Desarrollo Sostenible. <https://www.minambiente.gov.co/bosques-biodiversidad-y-servicios->

- ecosistemas/tenemos-el-50-de-los-paramos-del-mundo-ministro-de-ambiente-en-el-dia-de-las-montanas/
- Martin-Abadal, M., Guerrero-Font, E., Bonin-Font, F., & Gonzalez-Cid, Y. (2018). Deep Semantic Segmentation in an AUV for Online Posidonia Oceanica Meadows Identification. *IEEE Access*, 6, 60956–60967. <https://doi.org/10.1109/ACCESS.2018.2875412>
- Ministerio del Medio Ambiente (MADS). (2002). *RESOLUCIÓN NÚMERO 769 DE 2002*.
- Monserud, R. A., & Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. In *Ecological Modelling* (Vol. 62).
- Morales M, Otero J, Van der Hammen T, Torres A, Cadena C, Pedraza C, Rodríguez N, Franco C, Betancourth J.C., Olaya E., P. E. y C. L. (2007). Atlas de páramos de Colombia. In *Instituto de Investigación de Recursos Biológicos Alexander von Humboldt (IAVH)*. (Vol. 16, Issue 1). <https://doi.org/10.1111/j.1365-2303.2004.00206.x>
- Navab, N., Hornegger, J., Wells, W. M., & Frangi, A. F. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9351(Cvd), 12–20. <https://doi.org/10.1007/978-3-319-24574-4>
- Oscó, L. P., Marcato Junior, J., Marques Ramos, A. P., de Castro Jorge, L. A., Fathollahi, S. N., de Andrade Silva, J., Matsubara, E. T., Pistori, H., Gonçalves, W. N., & Li, J. (2021). A review on deep learning in UAV remote sensing. In *International Journal of Applied Earth Observation and Geoinformation* (Vol. 102). Elsevier B.V. <https://doi.org/10.1016/j.jag.2021.102456>
- Poveda-Sotelo, Y., Bermúdez-Cella, M. A., & Gil-Leguizamón, P. (2022). Evaluation of supervised classification methods for the estimation of spatiotemporal changes in the Merchán and Telecom paramos, Colombia. *Boletín de Geología*, 44(2), 51–72. <https://doi.org/10.18273/revbol.v44n2-2022002>
- Prince, S. J. D. (2023). *Understanding Deep Learning*. <http://udlbook.com>.
- Radoux, J., & Bogaert, P. (2017). Good practices for object-based accuracy assessment. *Remote Sensing*, 9(7). <https://doi.org/10.3390/rs9070646>
- Retavisca G., S. (2021). *La defensa de un complejo de páramos*. La Cola de Rata. <https://www.lacoladerata.co/cultura/relatos/la-defensa-de-un-complejo-de-paramos/>
- Saca, F. L., & Moreno, A. (2019). Clasificación de imágenes usando redes neuronales convolucionales en Python. *Universidad de Sevilla*, 80.

- Sahu, B. (2019, July 12). *The Evolution of Deeplab for Semantic Segmentation | by Beeren Sahu | Towards Data Science*. Towards Data Science. <https://towardsdatascience.com/the-evolution-of-deeplab-for-semantic-segmentation-95082b025571>
- Suárez Londoño, A. S., Jiménez López, A. F., Castro Franco, M., & Cruz Roa, A. A. (2017). Clasificación y mapeo automático de coberturas del suelo en imágenes satelitales utilizando Redes Neuronales Convolucionales. *Orinoquia*, 21(1 Sup), 64–75. <https://doi.org/10.22579/20112629.432>
- Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y. A., & Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations-A review. In *Remote Sensing* (Vol. 12, Issue 7). MDPI AG. <https://doi.org/10.3390/rs12071135>
- Torralba, P. P. (2022). Qué son las Redes Neuronales Convolucionales. *Thinking for Innovation*. <https://www.iebschool.com/blog/redes-neuronales-convolucionales-big-data/>
- Vásquez, A., Buitrago, A. C. (2011). El gran libro de los páramos. 2011. *Instituto de Investigación de Recursos Biológicos Alexander von Humboldt. Proyecto Páramo Andino*, 208.
- Yuan, X., Shi, J., & Gu, L. (2021). A review of deep learning methods for semantic segmentation of remote sensing imagery. In *Expert Systems with Applications* (Vol. 169). Elsevier Ltd. <https://doi.org/10.1016/j.eswa.2020.114417>
- Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). *UNet++: A Nested U-Net Architecture for Medical Image Segmentation*. <http://arxiv.org/abs/1807.10165>
- Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2019). *UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation*. <http://arxiv.org/abs/1912.05074>
- Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). *Deep learning in remote sensing: a review*. *October*. <https://doi.org/10.1109/MGRS.2017.2762307>